JOURNAL AerospaceLab

ONERA
THE FRENCH AEROSPACE LAB

Issue 12
December 2016
Special Issue celebrating the 70th birthday of ONERA

DOI : 10.12762/2016.AL12

Testing in Aerospace Research

Publisher
Stéphane Andrieux

Editor in Chief
Francis Dupoirieux

Production
ONERA Scientific
Information Department

On line
www.aerospacelab-journal.com
Webmaster ONERA

Contact
E-mail: aerospacelab@onera.fr

Produced by
ONERA - BP 80100
Chemin de la Hunière
et des Joncherettes
91123 PALAISEAU CEDEX
France
www.onera.fr

ISSN: 2107-6596

Testing in Aerospace Research: Introduction

Stéphane ANDRIEUX
(ONERA)
Chief Scientist

DOI : 10.12762/2016.AL12-00

The ONERA's 70th anniversary is an opportunity to ask ourselves about the place and evolution of research in the field of aerospace, especially in terms of needs and means. A key topic of these reflections is obviously the contribution of testing activities. Originally involved in all of the developments, they progressively had to face more complex requirements, as well as the emergence of competing approaches, mainly based on computation capacities and engineering methods. In this context, the 12th issue of *AerospaceLab Journal* offers an overview of the evolution of testing in various scientific domains concerned by aerospace developments: facilities, methods, instrumentation, processing, performance, exploitation of data, new fields of research involved, etc. The purpose is to assess the future use of testing, either as the best or as the only way to meet a specific requirement, such as equipment certification for instance, or as a complementary approach to other methods, including their validation or hardware in the loop analysis.

As quoted by Henri Poincaré in *La Science et l'hypothèse* (1908): "*The man of science must work with method. Science is built up of facts, as a house is built of stones; but an accumulation of facts is no more a science than a heap of stones is a house*". Such a quote is a good introduction to this special issue, which strives to illustrate both the primary need for experiments in aerospace research and also the need to stand back from the experiences: in conceiving it and in exploiting its results as well. This also encompasses all of the efforts and imagination that have to be deployed in order to build measurement or observation instruments, and to gain more benefit from the experiments by using or deriving improved methods for data or signal processing. Contributions using the inverse problem approach must also be recognized. Beyond the response to the need for understanding, modeling, establishing reference physical parameter values, etc., the development of more accurate instruments, more precise data processing techniques and even new physically-based devices has led over the last years to a great transfer towards industry.

This special issue dedicated to testing in aerospace contains fifteen papers gathering the contributions of more than ninety researchers. It covers the research and development activities of the four branches at ONERA: Materials and Structures, Fluid Mechanics and Energetics, Physics and Information Processing, and Systems, through numerous research projects and overviews covering several scientific Branches.

The first paper [*Inverse Problems and Experiments: a Fruitful Symbiosis*] strives to identify the mutual benefits that experiments and inverse problems have gained by working together, and exemplifies it with general linear problems in signal processing and two examples of image-based identification problems.

Then, a set of four papers addresses the field of mechanics. In the first one [*On Recent Advances in Microstructural Characterization and In-Situ Testing Techniques to Study Material Behavior*], an overview is provided of recent experimental techniques to characterize the microstructure and to measure the deformation of metallic materials. Some of the most promising characterization, measurement and identification techniques and methods used to develop, calibrate and validate physics-inspired constitutive models are discussed, with examples. Notably, the individual use of traditional interferometry is now clearly on the way to be replaced by the combined use of techniques based on image correlation (2D and 3D), *in-situ* and *ex-situ* scanning electronic microscopy (SEM), or EBSD, among others. The great potential offered by the combination of in-situ techniques at different scales with real-time computations is presented. The next paper [*Combined Experimental and Modeling Approaches for Strength Analysis of 3D Woven Composites: From Elementary Coupons to Complex Aeronautical Structures*] illustrates the combined approach of experimentation and modeling for strength analysis of composite materials and structures. Namely, dealing with new generation woven composites; it shows how multi-instrumented tests at various scales can be used in order to understand, model and validate the various damage mechanisms encountered both at the material level and at the structure level, up to the scale of industrial components. The thermal and mechanical fatigue of single crystal superalloys is addressed in the third paper [*Complex Thermo-Mechanical Approaches to Study the Behavior of High-Temperature Alloy*], based on complex torsional thermomechanical fatigue testing, including thermal gradient, which enables the severe thermomechanical in-service conditions acting on thermal barriers and modern air cooling gas turbine blade technologies to be reproduced. Thanks to a precise control of induction heating, it is possible to validate a complex constitutive model for thermal and mechanical fatigue,

including overheating effects, and to address the creep-fatigue interaction mechanism. The last paper [*Aircraft Ground Vibration Testing at ONERA*] describes the ground vibration testing techniques, past, present and future, deeply involved in the development of new aircraft, especially for avoiding (or determining the flight domain without) flutter.

This leads naturally to aerodynamics, again with four papers. In [*Testing in Aerodynamics Research at onera: the Example of the Transonic Buffet*], thirty years of combined experiments, modeling and computations at ONERA are detailed. It recalls the progress made in the deep understanding of the fundamental phenomenon of transonic buffet and the development of simulation capabilities. Among the different perspectives, the use of data assimilation techniques is identified as the most promising for the experiment-computation dialogue. Then, the paper [*Planar Particle Image Velocimetry for Aerospace Research at ONERA*] addresses the technique with the most striking progress for the study of aerodynamic flows in experimental setups over these last years. However, environmental conditions are also of primary importance to manage, and Lidar is a very precious tool to understand and measure atmospheric hazards, such as turbulence, wind shear or wind gust, and also induced hazards such as wake vortices. The paper [*Doppler LIDAR Developments for Aeronautics*] exposes recent developments at ONERA, giving some insight into the state of the art, as well as perspectives for future applications in aeronautics, such as the optimization of the separation distance between aircrafts in runways. The last paper [*Experimental Fluid Mechanics goes 3D: New Numerical Methods for Quantitative Instantaneous 3D Imagery of Fluids*] emphasizes the pointwise character of particles in PIV, in contrast with classical Tomo-PIV approaches, namely on the modeling of the imaging process, reconstruction of the 3D volume of particles and estimation of the 3D motion field between two time instants. Next, the 3D-BOS (Background Oriented Schlieren), which is aimed at estimating the instantaneous 3D density field of a moving fluid, is addressed in a new one-step numerical approach. Finally, important perspectives are drawn, especially for the study of compressible flows by combining both techniques

Turning now to physics, the first of a sequence of three papers [*Ro-Vibrational Spectroscopy in Hybrid fs/ps-CARS for N2 Thermometry*] addresses the coherent anti-Stokes Raman scattering methodology for the thermometry of N_2 in combustion chambers of aeronautics gas-turbines. The method and instruments have been enhanced, both with regard to robustness, which now enables real-life combustion chambers to be dealt with, and with regard to the repetition rate, which now reaches 1 kHz, a magnifying factor of 10 to 100 with respect to existing devices. The second paper [*Gravitation and Geodesy with Inertial Sensors, from Ground to Space*] reminds us of the dramatic change in the perception of Earth's gravity that has taken place since the beginning of this century. Thanks to past spatial missions like CHAMP, GRACE and GOCE, we have changed our mind in considering the Earth's gravity field and its measurement, using dedicated sensors and adequate data processing, revealing the changes of the Earth's field as a true signal, rather than perturbing terms added to the geostatic reference field. The development of ultraprecise instruments at ONERA has given rise to fundamental science applications like in the MICROSCOPE mission, where Einstein's Principle of Equivalence will be tested with an unprecedented accuracy, or to ground applications with the atomic interferometer for positioning marine vessels within the gravity field on the globe. The last article of this series [*Radiation Testing of Electronics Systems: How Can Simulation Tools Help in the Definition and Optimization of Test Plans in Labs?*] addresses a major concern related to on-board space systems: exposure to harsh environmental conditions, including radiation, extreme temperatures and a high vacuum. The wide diversity of observed disruptions, potential anomalies (and thus physical-chemical mechanisms behind them) and the large set of technologies involved, require the development of modeling tools to support experimentation, in order to achieve a representative simulation of device response and a reliable prediction of life mission, and to limit costly ground testing.

The interaction between modeling and experiment, between sensing the environment and acting on systems, can be exemplified by the last set of three papers. The paper [*Dealing with Complexity through Advanced Control Techniques*] deals with the control of complex dynamic systems encountered in aerospace, both with regard to their behavior and to their mission, through the implementation of multi-loop control architectures based on information on the system internal state and on its environment. In order to design "task" level control primitives, which take place just above the previous control loops, it is essential to combine sensor-based robust and non-linear control techniques, together with information on the environment, extracted from exteroceptive sensors in order to adapt system behavior to uncertainties and perturbations. The paper [*Surrogate Models for Aircraft Flight Control: Some Off-Line and Embedded Applications*] focuses on the need for surrogate models for applications that cannot afford high-fidelity simulation, and the various techniques that enable these models to be built from HF simulation results or sensing data. Beyond the environment, the last paper [*Component-Based Simulation for Real-Time Experiments of Advanced Aerospace Systems*] addresses concerns about the role of the human arising in human-in-the-loop system architecture.

After this quick overview of the articles in this issue, the extent of the spectrum of the applications and the importance of experiments in the aerospace domain now appears clear, although it cannot claim to be exhaustive ■

S. Andrieux
(ONERA)

E-mail: stephane.andrieux@onera.fr

DOI: 10.12762/2016.AL12-01

Inverse Problems and Experiments: a Fruitful Symbiosis

This paper offers a partial overview of some techniques and results emanating from inverse problems or signal processing with application to the development or enhancement of the exploitation of data obtained by experiments on physical systems pertaining to the aerospace domain. The first part deals with linear problems encountered in various problems of signal processing, whereas the second one addresses two specific image-based identification techniques. Finally, the conclusion strives to identify some future trends.

Introduction: Experiments, Simulation and Inverse Problems

From its very beginning, experiment has been indissolubly related to measurements and instrumentation, but it is only since a few decades ago that numerical processing, simulation and modeling have been deeply incorporated into the field. Symmetrically, the development of computational physics and mathematics in the mid-fifties opened the way to signal processing, and ten years later to addressing what was known as inverse problems (IP) but was studied until the eighties by a very small scientific community only due to a lack of computational power.

Since that period, a continuous development of the "accompaniment" of experiments by computation and modeling has led to new experiments, unprecedented accuracy, and measurements of "hidden" quantities. At the same time, the volume of experimental data has given rise to new capabilities for the identification of physical or geometrical parameters, and also to the emergence of validation procedures for the simulation means. Table 1 illustrates some of the mutual benefits that experiments, simulation and inverse problems have gained by collaborating.

In this paper, a partial overview will be given of some techniques and results emanating from the signal processing theory and the inverse problem community, which have enhanced, made possible or increased the performance level of experiments. Emphasis will also be placed on the notion of *a priori* information in inverse problems, this information being extracted from all of the knowledge accumulated through experiments. First, linear problems arising essentially from signal processing are addressed. Then, two problems coping with image capturing are investigated, since they are considered to be of utmost importance in aeronautical applications: the first deals with Particle Image Velocimetry for fluid flows, and the second one deals with full-field displacement measurement by digital image correlation in mechanics.

Benefits for Experiments gained from Simulation and IP	Benefits for IP and Simulation gained from Experiments
• Regularization, interpolation (super-resolution) and extrapolation of measurements • Extraction of information from highly noisy measured data • Optimization of instrumentation and location of sensors • Access to non-directly measured quantities or unreachable zones • Consideration of uncertainties • Enhancement of NDT-NDE processes and observation instruments • Hybrid experiments	• Definition of new inverse and identification problems from experimental data and procedures • Identification of model parameters, boundary conditions, and internal sources • Forecasting • Enrichment of a priori information • Validation • New NDT- NDE processes and new observation instruments • Hybrid experiments

Table 1 – Cross-fertilization between Experimentation, Simulation and Inverse Problems

Linear problems

Many examples can be found of linear inverse problems arising from experimental instrumentation or, more directly, from the goal aimed at in the investigation of the reality. Most of the examples fall into the field of Image and Signal Processing: signal deconvolution, image restoration, computed tomography, microwave imaging, and fluorescence imaging, as well as adaptive optics [24] and SAR [65] imaging. The general form of linear inverse problems is written as:

$$g = \boldsymbol{H}f + \varepsilon \tag{1}$$

where g is the measured data (output signal, blurred image, etc.), f is the unknown (input signal, original image) $\boldsymbol{H}$ can be called the forward operator (as it appears in the usual "direct problem"), and lastly ε is the noise or the error. Whereas seeking g when the pair $(f, \boldsymbol{H})$ is given is a direct or forward problem, seeking f with given $(g, \boldsymbol{H})$ is called the estimation or inverse problem, and seeking $\boldsymbol{H}$ (or parameters describing it) with given (f, g) is the identification problem.

As an example, in the deconvolution problem, $\boldsymbol{H}$ is a convolution with kernel h:

$$\boldsymbol{H}(f) = \int h(t-\tau) f(\tau) d\tau \tag{2}$$

$\boldsymbol{H}$ involves the impulse response of an instrument (called point spread function (PSF) in imaging, it is the spatial domain version of the imaging system transfer function). When the PSF is totally unknown one speaks of blind deconvolution, whereas when some measurements or estimations of the PSF are available (but the PSF is not fully known), we have the myopic deconvolution problem, in which both the object and the PSF have to be restored [51].

The operator $\boldsymbol{H}$ generally arises from the modeling of the underlying physics, and can sometimes be a matter of design, like in radar where the $\boldsymbol{H}$ operator is designated as the sensing operator. Note that the general form (1) also encompasses the Bayesian inference approach ([17], [18], [47]) where the pdf of the signal is now the unknown.

Once discretized (or sometimes directly resulting from the physics), $\boldsymbol{H}$ is a (rectangular) matrix, a Toeplitz matrix for deconvolution for example, and (f, g) will be vectors of respective lengths N and M (where generally $N>M$).

Ill-posedness and regularization

Inverse problems are generically ill-posed. From the continuous viewpoint, the ill-posedness arises either from the non-existence of a solution, the non-(finite) uniqueness of the solution or the non-continuity of the solution with respect to the data ([33], [44]). Box 1 gives some details about the ill-posedness of the Fredholm integral equation of the first kind, frequently arising in inverse problems. In the finite dimensional context, these concepts move to a non-invertible (or rectangular with $N>M$) matrix $\boldsymbol{H}$, and severe bad conditioning.

In order to deal with this difficulty, the general procedure is regularization, which was introduced in the sixties ([53], [63]); that is, loosely speaking, accepting to modify the operator or the model in such a way that well-posedness is recovered. The price to pay is a loss of accuracy.

Basically, regularization is performed on a variational form of the inverse problem, in which f is sought as the minimum of a functional (usually chosen as a norm of the gap with respect to the measured data):

$$f = \underset{p}{ArgMin}\, J(p) \tag{3}$$

The regularization is achieved by adding a stabilizing functional $S(p)$ with a (small) scalar regularization parameter α, which is a compromise parameter between the fidelity to the data and the desired properties of the function f that are enforced via the stabilizing functional.

$$J_\alpha(p) = J(p) + \alpha S(p) \tag{4}$$

Again, the choice of the stabilizing functional relies on *a priori* knowledge about the underlying physics. Usual choices for S are either based on regularity requirements on f, or guided by the vicinity of a given value f_0 of f. Historically, the stabilizing functional was the quadratic norm of the gradient of f, in order to smooth out the spatial or temporal oscillations of the signal frequently encountered in ill-posed problems, where the noise in the data is dramatically amplified in the solution. However, since this turns out to be too smoothing, and especially for edge-preserving applications, the total variation regularization or L1 stabilizing functional are used ([21], [54]), or even more sophisticated L1-L2 functionals: quadratic for small gradients and linear for large ones ([14],[23]).

Box 1 – Ill-posedness of the Fredholm integral of first kind

The equation is a typical example of an ill-posed problem and appears in numerous inverse problems:

$$g(x) = \int_J h(x,y) f(y) dy, x \in I$$

The Hadamard ill-posedness occurs under the following situations [3]:

- When the kernel h is continuous, g will be continuous for any measurable function f, therefore, no reasonable function f exists if g exhibits discontinuities,
- If the sign of the kernel is not constant, a non-finite set of solutions can exist. Simply by assuming that $I=J=[0, \pi]$, $h(x,y)=x \sin y$, $g(x)=x$, it is easy to verify that the set $\left\{ f_n(y) = \frac{1}{2} + \sin ny, n \in \mathbb{N} \right\}$ is a family of solutions,
- Thanks to the Lebesgue lemma (if $p(y)$ is continuous, then $\int_0^\pi p(y) \sin nydy \to 0$ when n tends to infinity), the perturbation of the solution can be set to be arbitrarily large whatever the norm of the data perturbation, so no continuity with respect to the data can be achieved.

Although the choice of the regularization parameter is often made on the basis of trial and error procedures, it can also rely on supplementary information about the measurement errors. The discrepancy principles ([10], [50]), for example, propose roughly speaking to determine the parameter α just by choosing the smallest value that makes the fidelity term of the functional be of the same order of magnitude as that of the measurement error: that is, not requiring more fidelity than the error level for the data. A last call to experience for the solution of inverse problems can be the choice of the minimization algorithm starting point, usually called the initial guess.

However, the most striking example of the use of *a priori* information in solving inverse problems is the Bayesian inference approach, where the philosophy is to consider the experiment as a means for adapting the *a priori* knowledge that we have about the situation at hand. More precisely, the Bayes theorem is used to build the posterior probability distribution of the parameters, given the prior distribution and the result of the experiment. In the linear case, the formulation obtained is very similar to the deterministic regularized least-square formulation (Gaussian prior corresponding to the weighted L2 fidelity norm, and the Laplacean prior to L1) with the noticeable difference being that there is no regularization parameter *per se*. The balance between the corresponding fidelity term and regularization term is provided naturally, given by the covariance matrices.

An additional benefit of the probabilistic approaches is that the question of the model used for extracting information or parameters out of the measured data can be addressed. Indeed, using models leads inevitably in an approximation or a systematic error, and to quoting G. Box [15]: "*The most that can be expected from any model is that it can supply a useful approximation to reality: All models are wrong; some models are useful*". This is the reason why some authors plead for the incorporation of uncertainties (and thus for a probabilistic, generally Bayesian, approach [39]): *A model that is wrong can only be useful if we acknowledge the fact that it is wrong […] model discrepancy is an important part of uncertainty quantification and must not be ignored, even though it may be hard to account for* [17].

Compressed sensing

Modeling and using a priori information also play a prominent role in the recent approach of compressed sensing [19]. One of its salient features is the identification of an orthonormal basis, say $\{\Psi_i\}$, where the unknown f has a S-sparse expansion, that is, a low number S of non-zero (or significant) components:

$$f = \sum_{i=1}^{S} f_i \psi_i \qquad S \text{ "small"} \tag{5}$$

The second fundamental premise of compressed sensing is incoherence, because it enables the possibility of performing the determination of f_i with very few samples of the data g. More precisely, if the basis for "representation" of the data g is denoted by $\{\varphi_i\}$, the coherence in IR^n between the "sensing" basis $\{\Psi_i\}$ and the representation basis $\{\varphi_i\}$, is

$$\mu(\varphi,\psi) = \sqrt{n} \max_{1\le i,j\le n} \left|\varphi_i.\psi_j\right| \tag{6}$$

A lower value of μ (note that $1 \le \mu \le \sqrt{n}$) corresponds to a lower correlation between any two elements of each basis [28]. Classical examples of pairs with low coherence are the *Fourier basis-Spikes basis* ($\mu = 1$), which represents maximal incoherence, or the *Noiselets basis-Haar wavelets basis* ($\mu = \sqrt{2}$). Coherence is involved in a quite strong result stating that the number of data sufficient for a good identification of f can be significantly lower than the Nyquist rate given by the Shannon theory: for an IR^n f signal that is S-sparse in the basis $\{\Psi_i\}$ and a set of m measurements selected in the basis uniformly at random, then

$$m \ge C\mu^2 S \log n \quad \Rightarrow \quad \{f^*\}_{i\le S} \text{ is exact} \tag{7}$$

where f^* is the solution of the following l_1 minimization problem, constrained by the m measurements of (1):

$$\min_{\{f\}\in IR^n} \left|\{f\}\right|_1 \quad \text{subject to } g_i = <\boldsymbol{H}f, \varphi_i> \; i = 1,\ldots,m \tag{8}$$

the result (6) being true within an "overwhelming probability" [19]. It is readily seen that for a coherence equal to one, the number m of needed data is of the order $S \log n$. If we now turn to noisy data, the question of robustness with respect to noise becomes central. The version of (8) is now a relaxed version, known as the LASSO or l_1-l_2 regularization ([29], [62], [51], [24], [27], [67]), where advantage is taken of the knowledge of the noise level ε:

$$\min_{\{f\}\in IR^n} \left|\{f\}\right|_1 \text{ subject to } \left\|g - \boldsymbol{H}f\right\|_2 < \varepsilon \tag{9}$$

Using the concept of the isometry constant δ_S for the matrix $\boldsymbol{H}$ and any integer S defined by the smallest number such that:

$$1 - \delta_S \le \frac{\left\|\boldsymbol{H}x\right\|_2^2}{\left\|x\right\|_2^2} \le 1 + \delta_S \quad \forall\, S - sparse\ x \tag{10}$$

[20] showed that the solution f^* of (8) has the following error inequality, when $\delta_{2S} < \sqrt{2} - 1$:

$$\left\|f^* - f\right\|_2 \le \frac{C_0}{\sqrt{S}} \left\|f - f_S\right\|_1 + C_1\varepsilon \tag{11}$$

This means that the error in the reconstruction is just supplemented, in the case of noise, with a term directly proportional to the noise level. Furthermore, [20] claims that the constants C_0 and C_1 are "small" (typically, for $\delta_{2S} = ¼$, C_0 and $C_1 < 6$).

Sparsity can also be enforced within the Bayesian approach to linear inverse problems [47], and also has a wide field of application in the identification of 3D vector fields ([25], [48]).

Completion and interpolation of velocity fields obtained by PIV

According to [64], "*The great challenges of flow dynamics require the validation of physical models with relevant length ranges spanning dozens of orders of magnitude, from a few Angstroms, at which atoms and molecules collide and bond (particularly in combustion), to a meter range (the size of a combustion chamber or a wing). Optical diagnostics are ideal for addressing this enormous challenge as they can span these different scales, offering both a global view through imaging and a microscopic view through spectroscopy. Moreover, these techniques are non-intrusive, so that they do not disturb the system or bias the measurements as other probing techniques do.*"

Unlike Particle Tracking Velocimetry [2], dealing with low seeding density and sparse estimation of the displacement, PIV is based either on direct correlation techniques between two successive images (including fast algorithms on GPUs, like the iterative Lucas-Kanade optical flow algorithm FOLKI [22]), or on image correlation techniques using the optic flow conservation equation. This last equation just expresses the fact that, under the assumption of conservation of the luminance of the particles, the image (gray level) $I(x,t)$ is only transformed by the velocity flow $\boldsymbol{v}$ between two time instants:

$$\partial_t I + \boldsymbol{v}.\nabla I = 0 \tag{12}$$

It is clear that this equation only gives access to limited information about the velocity field, namely the component of $\boldsymbol{v}$ parallel to the gradient of the function I; $\boldsymbol{v}^{\perp}$ remains undetermined (note in passing that no information is gained from a part of the flow with a uniform gray level in the image):

$$\boldsymbol{v} = -\partial_t I \frac{\nabla I}{\|\nabla I\|} + \boldsymbol{v}^{\perp} \ , \ \boldsymbol{v}^{\perp}.\nabla I = 0 \tag{13}$$

Thus, a naïve least-square formulation like the following, striving to minimize the residual of the equation (12), is doomed to fail.

$$\boldsymbol{v}^{opt} = \underset{\boldsymbol{v}}{ArgMin}\, J(\boldsymbol{v}) \equiv \int_{\Omega_m} \left(\partial_t I + \boldsymbol{v}.\nabla I\right)^2 dV \tag{14}$$

For two dimensional domains Ω, [59] showed that regularized functionals consisting of two terms (the first measuring the closeness to the data and the second measuring desired smoothing properties of the solution):

$$J_\alpha(\boldsymbol{v}) = \int_\Omega \left(\partial_t I + \boldsymbol{v}.\nabla I\right)^2 dV + \alpha \int_\Omega \|\nabla \boldsymbol{v}\|^2 dV \tag{15}$$

or

$$\begin{aligned} J_\alpha(\boldsymbol{v}) &= \int_\Omega \left(\partial_t I + \boldsymbol{v}.\nabla I\right)^2 dV + \alpha \int_\Omega \|\nabla \boldsymbol{v}\|^2_{W_I} dV \\ \|\nabla \boldsymbol{v}\|^2_{W_I} &= \nabla v_1.W_I.\nabla v_1 + v_2.W_I.\nabla v_2 \,, W_I \\ &= \frac{1}{|\nabla I|^2 + 2\gamma}\left[\nabla I^{\perp} \otimes \nabla I^{\perp} + \gamma Id_2\right], \nabla I^{\perp} \\ &= e_3 \wedge \nabla I \end{aligned} \tag{16}$$

are $H^1(\Omega)$-elliptic (hence convex) for $\alpha > 0$ and have a unique minimizer, which depends continuously on the data. The second regularized functional appears to be more physically grounded (since it turns out that it actually regularizes only the vector field component in the direction parallel to the intensity I, a component which is loosely controlled by the optical flow equation). Nevertheless, the first order Euler conditions for stationarity involve far more complicated boundary conditions for the solution field than for the first functional.

Estimates of fluid flow velocity fields are often corrupted, however, due to various deficiencies of the imaging process, making the physical interpretation of the measurements questionable. Some authors have proposed to deal with vector field estimates from any method and return a "physically plausible denoised version" thereof [66], whereas some more physical considerations can be introduced in the regularization added functional ([36] and the review paper [37]). By designing variational PIV methods, [55] showed how physically consistent flows should be estimated from PIV image sequences utilizing a distributed-parameter control approach. This has been extended in ([56], [57]) to a dynamic setting based on the vorticity transport

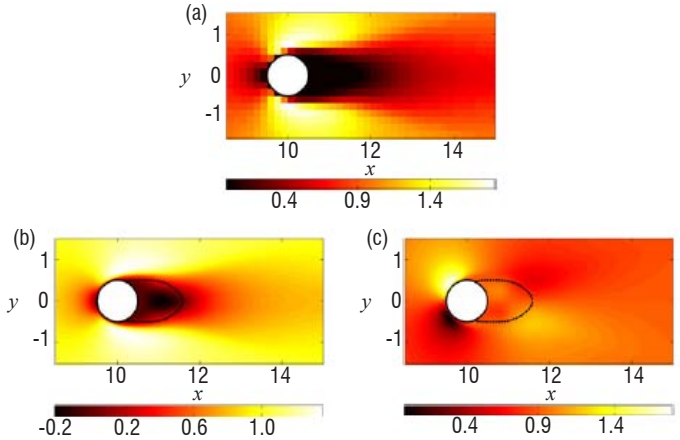


Figure 1 – Reynolds-averaged Navier-Stokes-driven mean-flow reconstruction around a cylinder: *(a)* velocity magnitude measurements, *(b)* and *(c)* reconstructed components of the mean velocity field (from [31])

equation formulation of the Navier-Stokes equation. [68] proposed to achieve spatial and temporal super-resolution of a measured flow. Applications are numerous in aerodynamics, for example in [26] for the analysis of turbulent jets, or for the interaction between shocks and boundary layer, [58].

Apart from the question of noise, which cannot be avoided, one can also be interested in extending the identification of the velocity field within the flow away from a "measurement box" or "measurement slice" Ω_m and in identifying other parameters describing the flow, or involved in its modeling. Two approaches can be proposed, sometimes sharing the final form of the algorithms, but with a different reasoning behind them: the optimal control approach [46] and the data assimilation approach [45]. Figure 1 describes the kind of results that can be achieved.

An optimal control approach for a model problem

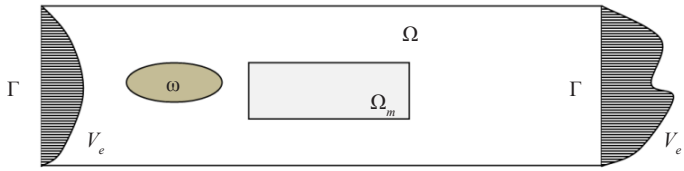


Figure 2 – Sketch of the geometry of the flow data regularization/extension problem

Let Ω be the total volume (or surface) where the velocity field v has to be identified, and let Ω_m be the subdomain where the velocity measurement is made. The inflow and outflow boundaries are denoted by Γ and let us denote the obstacle around which the flow is studied by ω (Figure 2).
In order to then enhance the velocity identification or estimation $\boldsymbol{v}$, and to extend the estimation outside of Ω_m, the following functional is built involving two auxiliary fields as control variables: a volume force density $\boldsymbol{f}$ in Ω and a surface velocity field g on Γ

$$F(\boldsymbol{f}, \boldsymbol{g}) \equiv H\left(\boldsymbol{v}(\boldsymbol{f}, \boldsymbol{g}), \boldsymbol{f}, \boldsymbol{g}\right) \tag{17}$$

with the velocity field $\boldsymbol{v}$ over the entire domain $\Omega\backslash\omega$ being related to the pair $(\boldsymbol{f}, \boldsymbol{g})$ by a state equation $E(\boldsymbol{v}, \boldsymbol{f}, \boldsymbol{g}) = 0$; for example, considering the simplest one:

$$\begin{cases} -\nu\Delta \boldsymbol{v} = \boldsymbol{f}\ in\,\Omega \\ \boldsymbol{v}\big|_{\Gamma} = \boldsymbol{V}^{e} + \boldsymbol{g} \\ \boldsymbol{v}\big|_{\partial\Omega\setminus\Gamma\cup\partial\omega} = 0 \end{cases} \tag{18}$$

This can be viewed as the Stokes equation without the incompressibility constraint div $\boldsymbol{v}=0$, which can hardly be fulfilled in a 2D setting where particles can exhibit a significant out-of-plane velocity component. $\boldsymbol{V}^{e}$ is an estimation of the inflow and outflow condition. The recovery algorithm is then:

$$\begin{aligned} (\boldsymbol{f}^{opt}, \boldsymbol{g}^{opt}) &= Arg\ Min\ F(\boldsymbol{f}, \boldsymbol{g}) \\ \text{and } \boldsymbol{v}^{opt} &= \boldsymbol{v}(\boldsymbol{f}^{opt}, \boldsymbol{g}^{opt}) \end{aligned} \tag{19}$$

where the velocity field $\boldsymbol{v}^{opt}$ is computed through the state equation with the optimal values of the forcing terms. The functional $F_i = H_i + S$ contains two terms. The first (fidelity term) can be the usual optical flow conservation residual, sometimes called the brightness change contrast equation, involving the intensity $I(\boldsymbol{x},t)$ (particle image):

$$H_1(\boldsymbol{v}) = \int_{\Omega_m} \left(\partial_t I + \boldsymbol{v}.\nabla I\right)^2 dV \tag{20}$$

if one deals directly with the analyzed image sequences, or a least-square error over the measured area:

$$H_2(\boldsymbol{v}) = \int_{\Omega_m} \left(\boldsymbol{v} - \boldsymbol{v}_m\right)^2 dV \tag{21}$$

if one deals with a first (rough) identified velocity field $\boldsymbol{v}_m$. The second term is a regularization (and bringing convexity) term, involving the stabilization of semi-norms on the control variables $\boldsymbol{f}$ and $\boldsymbol{g}$:

$$S_{\alpha\beta}(\boldsymbol{v}) = \alpha \|\boldsymbol{f}\|_{\Omega}^2 + \beta \|\boldsymbol{g}\|_{\Gamma}^2 \tag{22}$$

where (α, β) is a pair of (small) positive scalar parameters of the method. Whereas the semi-norm for the regularization of the velocity $\boldsymbol{g}$ can involve its tangential gradient $\nabla_s \boldsymbol{g}$ on Γ (in order to control more the spatial variation of the inlet and outlet velocities rather than their amplitude), the standard L^2 norm should be preferred for the regularization of the source term $\boldsymbol{f}$, as it stands for all of the irregular terms in the Navier-Stokes equation in the state equation.

Let us make a few remarks related to this formulation of the identification problem. First, the state equation is chosen in such a way that, given a pair of (regular) functions $(\boldsymbol{f}, \boldsymbol{g})$, a unique velocity field $\boldsymbol{v}(\boldsymbol{f}, \boldsymbol{g})$ is solution of the equation (18), thanks to the coerciveness of the Laplace operator and to the Lax-Milgram theorem. Second, the regularization term is essential in order to select a solution $\boldsymbol{v}$, just because the degree of freedom inserted in the formulation by the introduction of the unknown pair $(\boldsymbol{f}, \boldsymbol{g})$ is very large. Indeed, one can easily convince oneself that given a pair $(\boldsymbol{f}_0, \boldsymbol{g}_0)$, and accordingly a velocity field $\boldsymbol{v}_0$ solution of the state equation, the following family of triplets $(\boldsymbol{f}+\Delta\boldsymbol{\varphi}, \boldsymbol{g}, \boldsymbol{v}+\boldsymbol{\varphi})$ satisfies the state equation and leads to the same values of the functionals H_i (i=1or 2) for every regular $\boldsymbol{\varphi}$ with compact support in $\Omega\setminus(\Omega_m \cup \omega)$ .The reason why is simply that for every $\boldsymbol{v}$ in the family: $\boldsymbol{v} \equiv \boldsymbol{v}_0$ in Ω_m. Thus, for $\alpha = \beta = 0$, the convex functionals F_i cannot be strictly convex.

On the contrary, for $\alpha, \beta > 0$, the functionals F_i are strictly convex and positive, hence they have a unique minimum (depending on the pair (α,β)). Thus, the regularized identification problem has a unique solution. It is readily seen, thirdly, that the functionals F_i are quadratic (as $\boldsymbol{v}(\boldsymbol{f}, \boldsymbol{g})$ has an affine dependence on $(\boldsymbol{f}, \boldsymbol{g})$). Nevertheless, the implicit dependence of $\boldsymbol{v}$ on $(\boldsymbol{f}, \boldsymbol{g})$) prevents any use of the first order Euler condition for determining the minimum. The solution method is then the direct minimization and, due to the ill-posedness of the problem and to the cost of the computation of the functional F for a given pair $(\boldsymbol{f}, \boldsymbol{g})$, it is essential to have a precise computation of its gradient. The more efficient way to compute it and to take into account the implicit dependence of F with respect to $(\boldsymbol{f}, \boldsymbol{g})$, is the adjoint field approach. This can be achieved by using the following Lagrangian, involving two Lagrange multipliers $(\boldsymbol{w}, \boldsymbol{\mu})$:

$$\begin{aligned} L(\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{\mu}; \boldsymbol{f}, \boldsymbol{g}) = H(\boldsymbol{v}, \boldsymbol{f}, \boldsymbol{g}) - \int_{\Omega} \nu\nabla\boldsymbol{v}.\nabla\boldsymbol{w} dV + \\ \int_{\Omega} \boldsymbol{f}.\boldsymbol{w}\, dV + \int_{\Gamma} \boldsymbol{\mu}.(\boldsymbol{v} - \boldsymbol{V}^{p} - \boldsymbol{g}) dS \end{aligned} \tag{23}$$

The equation for the adjoint field $\boldsymbol{w}$ is obtained by stationarity of the Lagrangian with respect to $\boldsymbol{v}$, and is written as:

$$\begin{cases} -\nu\Delta\boldsymbol{w} = H'_{v}\left(\boldsymbol{v}(\boldsymbol{f}, \boldsymbol{g}), \boldsymbol{f}, \boldsymbol{g}\right)\ in\,\Omega \\ \nu\nabla\boldsymbol{w}.\boldsymbol{n}\big|_{\Gamma} = -\nu\nabla\boldsymbol{v}(\boldsymbol{f}, \boldsymbol{g}).\boldsymbol{n}\big|_{\Gamma} \\ \boldsymbol{w}\big|_{\partial\Omega\setminus\Gamma} = 0 \end{cases} \tag{24}$$

where H'_{v} is the partial derivative of the H functional with respect to $\boldsymbol{v}$; that is, for the two possible choices described previously:

$$\begin{aligned} H'_{v}\left(\boldsymbol{v}(\boldsymbol{f}, \boldsymbol{g}), \boldsymbol{f}, \boldsymbol{g}\right) &= 2\left(\partial_t I + \boldsymbol{v}.\nabla I\right)\nabla I\, \chi_{\Omega_m} \quad for\ H_1 \\ H'_{v}\left(\boldsymbol{v}(\boldsymbol{f}, \boldsymbol{g}), \boldsymbol{f}, \boldsymbol{g}\right) &= 2\left(\boldsymbol{v} - \boldsymbol{v}_m\right)\chi_{\Omega_m} \quad for\ H_2 \end{aligned} \tag{25}$$

where χ_{Ω_m} is the characteristic function of the set Ω_m. Equipped with the initial and adjoint fields, the gradients of the functional are simply computed from the equalities, which are none other than the partial derivative of L:

$$\begin{aligned} D_f F(\boldsymbol{f}, \boldsymbol{g}).\delta\boldsymbol{f} &= \int_{\Omega} \left(2\boldsymbol{f} + \boldsymbol{w}\right).\delta\boldsymbol{f} dV \\ D_g F(\boldsymbol{f}, \boldsymbol{g}).\delta\boldsymbol{g} &= \int_{\Omega} \left(2\nabla_t \boldsymbol{g}.\nabla_t \delta\boldsymbol{g} + \boldsymbol{w}.\delta\boldsymbol{g}\right) dV \end{aligned} \tag{26}$$

The choice of the state equation results clearly from a compromise between the simplicity of the state equation and the physics that it contributes to the problem; here, the simplest state equation has been retained (no inertial terms, no incompressibility constraint): the arbitration will be returned by the experience.

The (variational) data assimilation approach

Taking advantage of the measurements performed on a system, the general idea behind the data assimilation is to mix the information gained from the experiment, including the related uncertainties, the information gained from the modeling of the underlying physics, possibly including model uncertainties or imperfections and, lastly, *a priori* information on the system, substantially on its initial state (background information). The objective was initially to build a more precise estimation of the state, present or future, of the system. There are broadly two main families of methodologies for achieving this goal: filtering methods based on the (statistical) estimation theory [40] and variational methods [45] originally designed for meteorological forecasting applications. The general form of the functional to be minimized in data assimilation along a time interval of duration D is the following [32]:

$$J(\boldsymbol{g}, u) = \int_0^D \|H\boldsymbol{v}(\boldsymbol{g}, \boldsymbol{u}) - m\|_R^2\, dt + \int_0^D \|\boldsymbol{g}\|_F^2 dt + \int_{\Omega} \|\boldsymbol{u} - \boldsymbol{u}_0\|_B^2\, dV \tag{27}$$

where the following ingredients appear:

- v is the flow, governed by a state equation: $\partial_t \boldsymbol{v} + M(\boldsymbol{g}, \boldsymbol{u}; \boldsymbol{v}) = 0$, $\boldsymbol{v}(x,0) = \boldsymbol{v}(x)$ involving the initial state u, taken here as a control variable, and g, the perturbation of the inflow, taken as a control variable (like in Equation 18) as well;

- H is the observation operator, which maps the flow v to the measurements m, so that the first term of the functional J is simply a fidelity term to the data. The norm $\|\cdot\|_R$ used to measure the gap to the data in the data assimilation process, is built form the knowledge of measurement errors and is a covariance-based norm. In a finite dimensional setting, it reduces to the Manahobis norm generated by the inverse of the covariance matrix R: $\|x\|^2 = \langle R^{-1}x, x \rangle$.

- The second term appearing in the functional term is a regularization term on the control variable g. Again, the norm $\|\cdot\|_F$ is a covariance based norm.

- The last term involves the control variable u constituted by the initial condition of the flow, and here the regularization is related to the closeness of a background (*a priori*) information $\boldsymbol{u}_0$. The norm $\|\cdot\|_B$ is grounded in a finite dimensional setting to the covariance matrix B in the initial state.

As mentioned before, the probabilistic approach of variational data assimilation can be viewed as a regularization variational least-square identification method, but with a different reasoning behind it in the building of the regularization terms, which are based on *a priori* information (covariance matrix, backgrounds) and the peculiar feature that the regularization parameters are directly incorporated into the stabilizing functional terms via the choice of the covariance matrices. This last choice is nevertheless a delicate matter, with major concerns about cross-correlation and dependency.

In the previous example, a control variable has been added to the initial condition u, which is the only control variable in the so-called 4D-Var data assimilation method. For several years, there has been a growing demand for simultaneous identification of parameters and the estimation of state variables. Beyond the interest of exploiting large amounts of data for enhancing the simulation models (identification of model parameters) or the description of the experiments (identification of boundary conditions), a great potential can be expected either in having simulations or performing analysis for systems that do not yet exist, or even in revisiting the process of simulation itself. Joint state and parameter estimation can be simply addressed by augmenting the state variable space with the parameters $\{p\}$ and correspondingly the state equation on the parameters with the simple equation $\dot{p} = 0$ ([52], [60]). Apart from the issue of defining the related covariance operators, the adaptation of the previous formulation is straightforward. Nevertheless, this approach can be questionable on two points. First, the augmented state equation mixes quantities and equations of a very different nature, and the parameter evolution equation is somehow artificial given that the parameters are, by nature, fixed parameters. However, considered now as a function of time, it simply has to be expected that it will tend to a stationary value p^{∞}, in order for this value to be able to be recognized as the "true value" of the parameter. Secondly, for situations where it is impossible to have significant *a priori* statistical information about the parameters, it will be impossible to build the covariance operator and to calibrate the level of uncertainties in relation to one of the other control variables.

This is the reason why some other approaches have been proposed ([34], [3]), based on a hierarchical identification process (with a first deterministic step and a related functional energy error functional *E*), and with the choice of the covariance operator as the Hessian of *E*. As an extension of the parameter identification within a predetermined model, [1] proposed to estimate conjointly a flow and the parameters of its POD-reduced model.

Expanding, inside a body, measured surface fields

Advances in the development of digital cameras, image correlation techniques (DIC) and infrared cameras now make it possible to have measurement means for full-field surface displacements or temperatures that are cheap and relatively easy to manage and, more importantly, leading to very large amounts of information [61]; an example of estimated surface displacements on a cracked mock-up is given in Figure 3. Nevertheless, the use of these surface data is still largely restricted either to qualitative estimation or to quantitative analysis based on a plane mechanical or thermal state, or on homogeneous-through-the-thickness assumptions [12]. Aimed at a true 3D quantitative imaging process, the problem of reconstruction of the fields inside the solid from images obtained on parts of its boundary must be addressed.

One approach in dealing with this problem in mechanics is to first reformulate it within the continuous framework, taking advantage of the fact that the amount and spatial density of the information obtained using the digital image correlation techniques make it possible to consider that the complete displacement field is available on a part of the boundary and is not reduced on it to pinpoint data only. Then, it is possible secondly to formulate it as a Cauchy or Data Completion Problem, taking into account the fact that an overspecified data pair is given on a part Γ_m of the boundary: the capture of displacement fields via the DIC on a stress-free boundary with unit normal n gives access to the Dirichlet-Neumann pair $(U, \sigma.n)$. The Cauchy Problem is an archetypal ill-posed problem ([13], [33]).

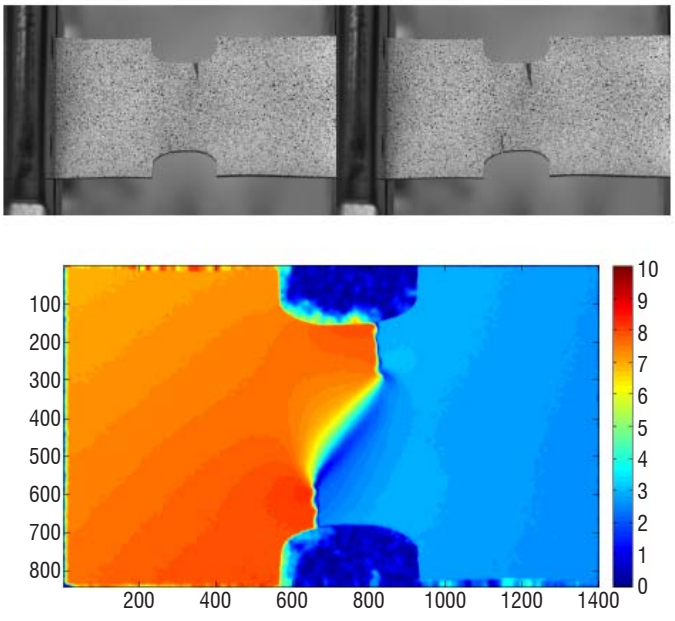


Figure 3 – Surface tangential displacement field obtained by DIC for a cracked mock-up (ONERA)

In the generic situation, the usual boundary conditions are given on Γ_b. Γ_m is the part where, using DIC acquisition measurements, both displacement and stress vector components (U_m, F_m=0) are available. Lastly, Γ_u is the remaining part of the boundary, where no boundary data is known. For the sake of simplicity, we will assume that Γ_b is empty, but the extension of the method to situations with non-empty Γ_b is straightforward. The boundary Γ_u is generally non-connected and can possibly contain internal surfaces such as cracks or boundaries of cavities and inclusions (see Figure 4).

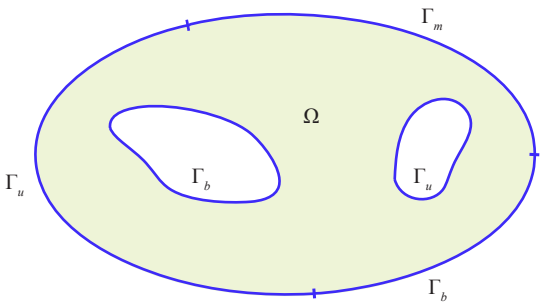


Figure 4 – The geometry of the problem with the partition (Γ_m, Γ_b, Γ_u) of the boundary of the considered body Ω

Within the framework of linearized strain and for isothermal transformation, the constitutive equation for the material constituting the solid is written in an abstract format as a relation between the Cauchy stress tensor σ and the linearized strain tensor ε, and can also involve internal variables α: $f(\boldsymbol{\sigma},\boldsymbol{\varepsilon},\alpha)=0$. This equation is complemented with the evolution equation for the internal variables α: $g(\boldsymbol{\sigma},\boldsymbol{\varepsilon},\dot{\alpha})=0$. All of these equations can also, and generally will, be inclusion equations within the framework of convex analysis, using the notion of sub-differential. The functions f and g entering the constitutive relation can be also functions of the space variable x, given that the solid Ω can be heterogeneous; however, for the sake of clarity, it will be omitted.

The form of the Data Completion Problem (or abusively the Cauchy Problem) that we can select, is to determine the missing Neumann boundary data $\eta=\sigma.n)$ on the part Γ_u of the boundary of the solid such that there exists (u, σ, α) in the domain Ω over the time interval [0, D] fulfilling the following:

$$\begin{cases} div\sigma = 0 & in\ \Omega \\ f(\boldsymbol{\sigma},\boldsymbol{\varepsilon},\alpha)=0, g(\boldsymbol{\sigma},\boldsymbol{\varepsilon},\dot{\alpha})=0 & in\ \Omega \\ \boldsymbol{u}\big|_{t=0}=\boldsymbol{u}_0\ ,\ \boldsymbol{\sigma}\big|_{t=0}=\boldsymbol{\sigma}_0\ \alpha\big|_{t=0}=\alpha_0 & in\ \Omega \\ \boldsymbol{u}\big|_{\Gamma_m}=\boldsymbol{U}_m\ ,\ \boldsymbol{\sigma}.\boldsymbol{n}\big|_{\Gamma_m}=\boldsymbol{F}_m & on\ \Gamma_m \\ \boldsymbol{\sigma}.\boldsymbol{n}\big|_{\Gamma_u}=\boldsymbol{\eta} & on\ \Gamma_u \end{cases} \quad (28)$$

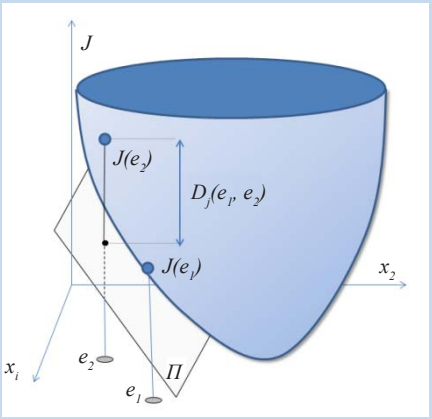
where $\varepsilon(\boldsymbol{u})=\left[\nabla\boldsymbol{u}\right]^{sym}$ is the linearized strain operator. The general variational method, derived firstly in [5] for the Laplace operator (stationary isotropic conduction equation), for solving this problem is based on two steps. First, two auxiliary usual well-posed problems

Box 2 – The Bregman divergence

Introduced within the context of convex optimization ([16], [41]), the Bregman distance, or Bregman divergence, is akin to a metric, although not satisfying the triangle inequality nor symmetry. Let J be a convex function, the generalized Bregman distance between e_1 and e_2 with respect to J is the **non-negative scalar**:

$$D_J(e_1,e_2)=J(e_2)-J(e_1)-\langle p_1,e_2-e_1\rangle \quad \text{for } p_1\in\partial J(e_1)$$

where $\partial J(e)$ stands for the subdiffential of J at e.



Geometrical interpretation of the Bregman divergence

Π is the tangent plane to J at e_1

In IR^n, it is readily seen that the squared Euclidean distance $\|e_1-e_2\|^2$ or Mahalanobis distance $(e_1-e_2)^T Q(e_1-e_2)$ are Bregman divergences, respectively generated by $J(x)=\|x\|^2$ and $J(x)=x^T Qx$. In probability theory, the Kullback–Leibler divergence or relative entropy [43]:

$$D(e,a)=\sum e_i \log\frac{e_i}{a_i}-e_i+a_i$$

is generated by the function

$$J(e)=\sum e_i \log e_i - e_i$$

P_i, $i = 1,2$ are defined, each using only one of the overspecified boundary data on Γ_m and a given normal stress vector field η on Γ_u:

$$\begin{cases} div\sigma_i = 0 & in\ \Omega \\ f(\boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_i, \alpha_i) = 0, g(\boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_i, \dot{\alpha}_i) = 0 & in\ \Omega \\ \boldsymbol{u}_i|_{t=0} = \boldsymbol{u}_0\ ,\ \boldsymbol{\sigma}_i|_{t=0} = \boldsymbol{\sigma}_0\ \ \alpha_i|_{t=0} = \alpha & in\ \Omega \\ \boldsymbol{\sigma}_i.\boldsymbol{n}|_{\Gamma_u} = \boldsymbol{\eta} & on\ \Gamma_u \end{cases} \tag{29}$$

And respectively for $\mathcal{P}_1$ and $\mathcal{P}_2$:

$$(\mathcal{P}_1)\{\boldsymbol{u}_1 = \boldsymbol{U}_m\ \ on\,\Gamma_m\} \qquad (\mathcal{P}_2)\{\boldsymbol{\sigma}_2.\boldsymbol{n} = \boldsymbol{F}_m\ \ on\,\Gamma_m\} \tag{30}$$

With the additional condition (global equilibrium) on $\boldsymbol{\eta}$:

$$\int_{\Gamma_m} \boldsymbol{F}_m dS + \int_{\Gamma_u} \boldsymbol{\eta} dS = 0\ .$$

It is clear that if a surface traction field $\boldsymbol{\eta}_{opt}$ on Γ_u is such that $\boldsymbol{u}_1 = \boldsymbol{u}_2 + \boldsymbol{RBM}$, where $\boldsymbol{RBM}$ is a Rigid Body Motion, the two problems $\mathcal{P}_1$ and $\mathcal{P}_2$ will have the same solution (σ, α). Therefore, the Cauchy problem is solved with $\boldsymbol{u}_1$, and the solution of the Data Completion Problem is $\boldsymbol{\eta}_{opt}$; furthermore, the Dirichlet-Neumann data pair on Γ_u is ($\boldsymbol{u}_1$, $\boldsymbol{\eta}_{opt}$).

The second step will therefore be to minimize a suitable gap functional E between the two states $(\boldsymbol{u}_1, \boldsymbol{\sigma}_1, \alpha_1)$ and $(\boldsymbol{u}_2, \boldsymbol{\sigma}_2, \alpha_2)$ solutions of $\mathcal{P}_1$ and $\mathcal{P}_2$, as a function of the sought stress vector field η on Γ_u, leading then to the variational method:

$$\boldsymbol{\eta}_{opt} = \underset{\boldsymbol{\eta}}{ArgMin}\ G(\boldsymbol{\eta}) \equiv \mathcal{E}\left([\boldsymbol{u}_1, \boldsymbol{\sigma}_1, \alpha_1](\boldsymbol{\eta}), [\boldsymbol{u}_2, \boldsymbol{\sigma}_2, \alpha_2](\boldsymbol{\eta})\right) \tag{31}$$

Clearly, the choice of the gap $\mathcal{E}$ is of primary importance, in view of the overall performance of the method. The gap functional takes into account first that the dimensions of the state variables are various, and second that the states involved are not any collections of the state variables, but rather obey a conservation law and a constitutive relation. What is needed is therefore a kind of metric defined on a geometric variety, rather than a distance in the vector space into which it is plunged.

It is therefore proposed to turn to Bregman divergences, which can be built from any convex function (see Box 2). More precisely, after having selected a convex function J of the state variables suited to each kind of constitutive relation, the gap functional is written as:

$$\mathcal{E}_J(e_1, e_2) = \int_\Omega D_J^s(e_1(x), e_2(x))dV \tag{32}$$

Where $D_J^s(e_1, e_2)$ is the symmetrized Bregman divergence:

$$\begin{aligned} D_J^s(e_1, e_2) &= D_J(e_1, e_2) + D_J(e_2, e_1) \\ &\equiv \langle p_1 - p_2, e_1 - e_2 \rangle \end{aligned} \tag{33}$$

The properties of the Bregman divergence are closely linked to the Legendre-Fenchel inequality in convex analysis ([30], [38]), especially in establishing the following property, which is of upmost importance in the perspective of building a gap functional.

Proposition: *If J is strictly convex, then the following equivalences hold true*

$$D_J^s(e_1, e_2) = 0 \Leftrightarrow e_1 = e_2 \Leftrightarrow \partial J(e_1) = \partial J(e_2) \tag{34}$$

Indeed, coming back to the solution process (13) for solving the data completion problem (10), using a gap functional based on the symmetrized Bregman divergence (33) leads to a functional correctly suited for the minimization:

$$\begin{aligned} &\mathcal{E}_J(e_1, e_2) = \int_\Omega \langle p_1 - p_2, e_1 - e_2 \rangle dV \geq 0 \\ &\mathcal{E}_J(e_1, e_2) = 0 \qquad \Leftrightarrow e_1 = e_2\ \ and\ \ p_1 = p_2 \end{aligned} \tag{35}$$

It is worth noting that the strictly convex function J does not appear anymore in the expression of the symmetrized Bregman divergence, so there is no need to compute it: it is sufficient for such a function to be able to be identified in order for the property (35) to hold true. Returning to the Data Completion Problem for extending mechanical fields inside the solid, the primal state variables are ($\varepsilon(u)$, α), and seeking convex functions J generating an appropriate and effective gap functional $\mathcal{E}_J$, one is naturally led to the convex framework of modeling for the constitutive equation. Actually, the Standard Generalized Materials [35] theory rests on the definition of a convex potential W (the free Helmholtz-Gibbs energy) and a convex pseudo-potential of dissipation Ψ (positively homogeneous of Degree 1), such that the constitutive equations appearing in (28) are written:

$$\begin{aligned} &W = W(\varepsilon - \varepsilon^p, \alpha),\ \psi = \psi(\dot{\varepsilon}^p, \dot{\alpha}) \\ &\sigma = \frac{\partial W}{\partial \varepsilon} = -\frac{\partial W}{\partial \varepsilon^p},\ A = -\frac{\partial W}{\partial \alpha} \\ &\sigma \in \partial_{\dot{\varepsilon}^p}\psi\ ,\ A \in \partial_{\dot{\alpha}}\psi \end{aligned} \tag{36}$$

Here the plastic strain ε^p appears separately in the internal variable list, and the free energy is a function of the elastic strain $\varepsilon^e = \varepsilon - \varepsilon^p$ in order to comply with the usual presentations.

For linear elasticity and non-linear elasticity (in a small strain context) where neither internal variables nor dissipation pseudo-potential are involved, the variational method (31) for solving the Data Completion Problem has been applied to the extension of surfacic displacement fields for various applications, such as identification of contact areas [6], of material parameters in inclusions, of geometry of crack fronts [7], of internal pressure in cavities [9], or determination of linear fracture mechanics parameters [8]. The following gap functional has been used:

$$G(\boldsymbol{\eta}) = \mathcal{E}_W\left(\varepsilon_1([u(\boldsymbol{\eta})]), \varepsilon_2([u(\boldsymbol{\eta})])\right) = \int_\Omega (\sigma_1 - \sigma_2) : (\varepsilon_1 - \varepsilon_2) dV \tag{37}$$

In the case of linear elasticity (quadratic potential W), the symmetrized Bregman divergence turns out to be simply twice the elastic energy of the displacement field $u_1 - u_2$: $(\sigma_1 - \sigma_2) : (\varepsilon_1 - \varepsilon_2) = 2W(\varepsilon(\boldsymbol{u}_1 - \boldsymbol{u}_2))$.

For the elastoplastic case with implicit incremental formulation [4] and taking advantage of the linearity of the Bregman divergence, that is:

$$D_{\lambda J + \mu F}(e_1, e_2) = \lambda D_J(e_1, e_2) + \mu D_F(e_1, e_2) \quad \forall \lambda, \mu \geq 0 \tag{38}$$

the following single-parameter family of gap functions can be built ($0 \leq \chi \leq 1$):

$$\begin{aligned} \mathcal{E}_{\chi W + (1-\chi)\Psi} = &\int_\Omega (\Delta\sigma_1 - \Delta\sigma_2) : \left(\chi(\Delta\varepsilon_1 - \Delta\varepsilon_2) + (1-2\chi)(\Delta\varepsilon_1^p - \Delta\varepsilon_2^p)\right) dV \\ &+ \int_\Omega \langle A_1 - A_2, (1-2\chi)(\Delta\alpha_1 - \Delta\alpha_2) \rangle dV \end{aligned} \tag{39}$$

Box 3 – Solution algorithm for the data completion problem (extension of mechanical fields for surface displacement data)

Given a Dirichlet-Neumann data pair on Γ_m: (U_m, F_m)

i) Pick an initial value η_0 of the sought stress vector on boundary Γ_u

ii) Solve the two auxiliary problems: $\mathcal{P}_1^k$ with (U_m,η_k) as boundary conditions, $\mathcal{P}_2^k$ with (F_m,η_k) as boundary conditions

iii) Compute the Gap between the two solutions of $\mathcal{P}_1$ and, $\mathcal{P}_2$: $Gk=\mathcal{E}(e_1^k,e_2^k)$

iv) If $Gk < tol$ then end, else

v) Compute the gradient of G by solving two adjoint problems

vi) Update η_k by any descent algorithm, go to ii)

The condition of strict convexity of $\chi W+(1-\chi)\Psi$ can be ensured only for $\chi>0$, because the potential Ψ is not strictly convex (as a positively homogeneous function of Degree one: $\Psi(\lambda x)=\lambda\Psi(x)$ for $\lambda\geq 0$). Within this family, the gap function obtained for $\chi=$ 1/2, which balances exactly between the free energy gap and the dissipation gap, has the peculiar feature of involving only the stress and strain tensors and can then be called the Drücker Gap:

$$\mathcal{E}_D=\int_\Omega\left(\Delta\sigma_1-\Delta\sigma_2\right):\left(\Delta\varepsilon_1-\Delta\varepsilon_2\right)dV \qquad (40)$$

The Drücker Gap leads to the only computable function $G(\eta)$ by a boundary integration over the entire external boundary of the solid, as is the case in linear elasticity and hyperelasticity, using the virtual power principle. This feature has been widely used previously to improve the global performance of the solution algorithm for a linear Cauchy problem and to reduce the computational burden [6].

The following figures show the kind of results that can be achieved with respect to the extension of the mechanical field inside a solid from surface displacement measurements. The method and algorithm are summarized in Box 3. Results for linear elastic behavior were obtained with very few iterations (~ 10), whereas for strongly non-linear materials such as elastoplatic materials (elastoplaticity with linear hardening), a hundred minimizing iterations are needed.

It can be seen that either the contact surface for an elastic complex structure (Figure 5), or the residual stress after unloading for an elastoplatic solid (Figure 6) can be recovered with good accuracy.

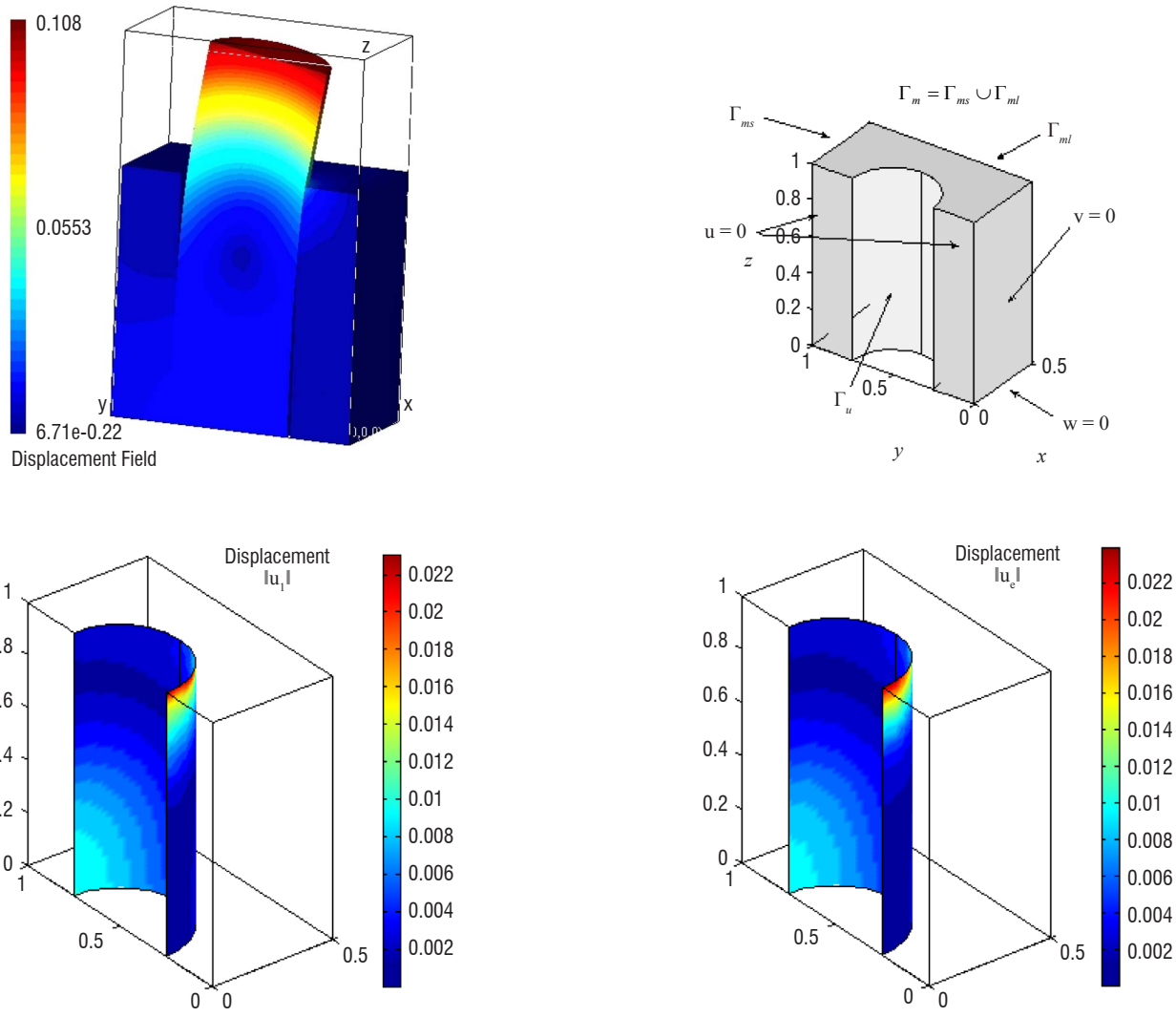


Figure 5 – 3D Identification of the displacement on the interior surface of an elastic mock-up loaded by a flexured rod (from [6])

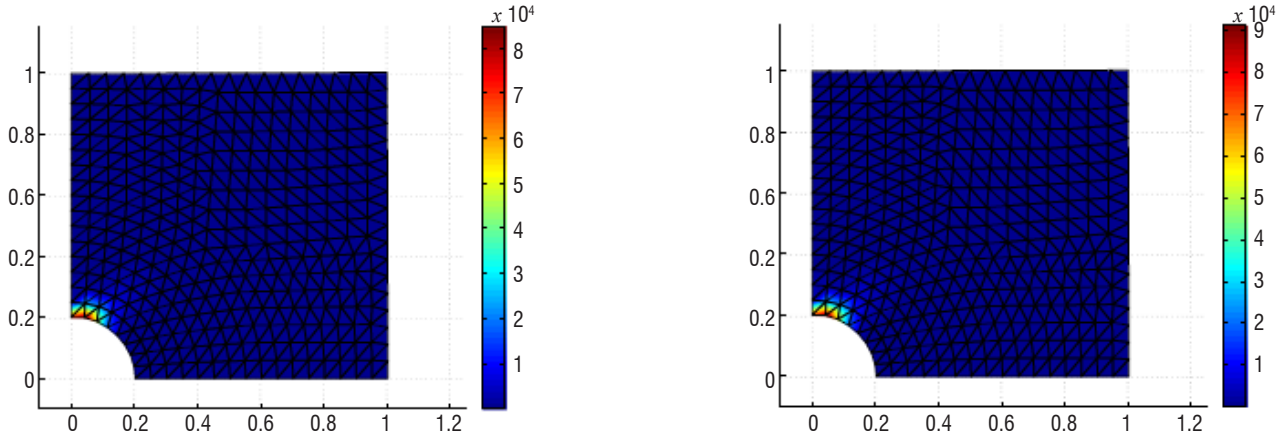


Figure 6 – Identification of residual stress (Von Mises norm) in a holed solid from a (synthetic) displacement given on the stress-free top during a loading and unloading process (from [4])

The proposed approach, using the minimization formulation (31), does not explicitly incorporate any kind of regularization, although the Cauchy problem is very ill-posed. Nevertheless, the use of a conjugate gradient algorithm entails a form of regularization in the numerical applications: the value of the stopping criterion can also plays the role of a regularization parameter [42]. For moderate noise in the data (under 2%), this regularization appears to be sufficient, as shown for the thermal conductivity equation [5].

Conclusions

This paper overviewed some results and techniques of inverse problems that can be used in signal processing on data emanating from various experimental situations arising in the aerospace domain. Two specific topics related to image-based identification problems have been focused on.

With regard to the near future, some leads or trends can be pointed out. Two promising, and already fruitful, ways can be identified in the perspective of combining measurements of distinct data acquisition nature. The first is the coupling between the 3D-PIV techniques and the 3D Background Oriented Schlieren (BOS), which is aimed at the reconstruction of the density field of instantaneous flows via the analysis of the light deflection through a medium with inhomogeneous optical index. The second is the simultaneous exploitation of the DIC displacement field measurements and the thermal camera temperature field measurements through nonlinear thermomechanics modeling. Originated by an interesting change of perspective, the passive approach to wave probing, which uses ambient noise or diffuse fields (nondestructive testing, seismic identification, echography, etc.) or non-cooperative sources of illumination (radar), stimulates a wide field of research. The deep learning applied to the analysis of a large amount of experimental data would also generate high hopes, although the number of really convincing results remains still low in the aerospace domain. From the numerical simulation side, progress can be expected thanks to the development of data assimilation, taking into account uncertainties both in the results and in the methods themselves, and enhancing performances through the use of model reduction techniques. ■

References

[1] J. D'ADAMO, N. PAPADAKIS, E. MEMIN and C. ARTANA - *Variational Assimilation of POD Low-Order Dynamical Systems*. J. Turbul, 8(9), 2007, 1-22.

[2] R. J. ADRIAN - *Particle Imaging Techniques for Experimental Fluid Mechanics.* Ann. Rev. Fluid Mechanics, 23, 1991, 261-304.

[3] S. ANDRIEUX - *Recalage, identification, suivi en service des structures.* Presses des Ponts, Paris, 2016.

[4] S. ANDRIEUX, T. N. BARANGER - *On the Determination of Missing Boundary Data for Solids with Nonlinear Material Behaviors, Using Displacement Fields Measured on a Part of their Boundaries*. J. Mech. Phys. Solids, 2016, http://dx.doi.org/10.1016/j.jmps.2016.02.008i.

[5] S. ANDRIEUX, S., T. N. BARANGER, and A. BEN ABDA - *Solving Cauchy Problems by Minimizing an Energy-Like Functional*. Inverse Problems, 22(1), 2006, 115-133.

[6] S. ANDRIEUX, S., T. N. BARANGER - *An Energy Error-Based Method for the Solution of the Cauchy Problem in 3D Linear Elasticity*. Comput. Methods Appl. Mech. Eng. 197, 2008, 902-920.

[7] S. ANDRIEUX, T. N. BARANGER - *Emerging Crack Front Identification from Tangential Surface Displacements.* C.R. Mécan. 340(8), 2012, 565-574.

[8] S. ANDRIEUX, T. N. BARANGER - *Three-Dimensional Recovery of Stress Intensity Factors and Energy Release Rates from Surface Full-field Displacements*. Int. J. Solids Struct. 50(10), 2013, 1523-1537.

[9] S. ANDRIEUX, T. N. BARANGER - *Solution of Nonlinear Cauchy Problem for Hyperelastic Solids*. Inverse Problems, 31(11), 2015, 115003-115022.

[10] R. ARCANGELI - *Pseudo-solution de l'Equation Ax=y, comptes rendus de l'Académie des Sciences. Paris, Series A*. Vol. 263A, 1966, 282-285.

[11] G. ARTANA, A. CAMMILLERI, J. CARLIER, and E. MÉMIN - *Strong and Weak Constraint Variational Assimilations for Reduced Order Fluid Flow Modeling*. Journal of Computational Physics, Volume 231, Issue 8, 20, 2012, 3264-3288.

[12] S. AVRIL, M. BONNET, A. BRETELLE, M. GRÉDIAC, F. HILD, P. IENNY, F. LATOURTE, D. LEMOSSES, A. PAGANO, E. PAGNACCO, and F. PIERRON - *Overview of Identification Methods of Mechanical Parameters Based on Full-Fields Measurements*. Exp.Mech. 48, 2008, 381-402.

[13] F. BEN BELGACEM - *Why is the Cauchy Problem Severely Ill-Posed ?* Inverse Problems 23(2), 2007, 823-836.

[14] C. BOUMAN and K. SAUER - *A Generalized Gaussian Image Model for Edge-preserving Map Estimation.* IEEE Trans. Image Processing 2, 1993, 296-310.

[15] G. E. P. BOX, J. S. HUNTER, and W. G. HUNTER - Statistics for Experimenters, John Wiley & Sons, 2005.

[16] L. BREGMAN - *The Relaxation Method of Finding the Common Point of Convex Sets and its Application to the Solution of Problems in Convex Programming*. USSR Computational Mathematics and Mathematical Physics 7 (3), 1967, 200-217. [17] J. BRYNJARSDOTTIR, and A. O'HAGAN - *Learning About Physical Parameters: The Importance of Model Discrepancy.* Inverse Problems, 30, 2014.

[18] D. CALVETTI, J. P. KAIPIO, and E. SOMMERSALO - *Inverse Problems in the Bayesian Framework.* Inverse problems, 30, 2014, Bayesien et proba.

[19] E. J. CANDÈS, and M. B. WAKIN - *An Introduction To Compressive Sampling.* IEEE Signal Processing Magazine, 21, 2008.

[20] E. J. CANDÈS, J. ROMBERG - *Sparsity and Incoherence in Compressive Sampling*, Inverse Problems, vol 23, n°3, 2007, 969-985.

[21] A. CHAMBOLLE - *An Algorithm for Total Variation Minimization and Applications.* Journal of Mathematical Imaging and Vision 20, 2004, 89-97.

[22] F. CHAMPAGNAT, A. PLYER, G. LE BESNERAIS, B. LECLAIRE, S. DAVOUST, and Y. LE SAN - *Fast and Accurate PIV Computation Using Highly Parallel Iterative Correlation Maximization*. Experiments in Fluids, 50, 2011, 1169-1182.

[23] G. CHENEGROS, L. M. MUGNIER, F. LACOMBE, and M. GLANC - *3D Deconvolution of Adaptive-Optics Corrected Retinal Images.* Proc. SPIE 6090, Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XIII, 60900P, February 23, 2006, DOI:10.1117/12.645233.

[24] J. M. CONAN, L. MUGNIER, T. FUSCO, V. MICHAU, and G. ROUSSET - *Myopic Deconvolution of Adaptive Optics Images by Use of Object and Point-Spread Function Power Spectra*. Applied Optics, Vol. 37, N°21, 1998.

[25] P. CORNIC, F. CHAMPAGNAT, A. CHEMINET, B. LECLAIRE, and G. LE BESNERAIS - *Fast and Efficient Particle Reconstruction on a 3D Grid Using Sparsity*. Experiments in Fluids, 56(3), 2015, 1-7.

[26] S. DAVOUST, L. JACQUIN, and B. LECLAIRE - *Dynamics of m=0 and m=1 Modes and of Streamwise Vortices in a Turbulent Axisymmetric Mixing Layer.* J. Fluid Mech. Vol. 709, 2012, 408-444.

[27] G. DESODT, L. AOUCHICHE, and O. RABASTE - *Extract Before Detect, N-Signal Complex Approximate Message Passing Applied to Radar Signals*. COSERA (COmpressed SEnsing Applied to RAdar Signals), Bonn, 2013.

[28] D. L. DONOHO - *Compressed Sensing.* IEEE Trans. IT, vol. 52/no. 4, 2005, 1289-1306.

[29] D. L. DOHONO, and X. HUO - *Uncertainity Principles and Ideal Atomic Decomposition.* IEEE Trans. Inform.Theory, Vol. 47, n°7, 2001, 2845-2862.

[30] I. EKELAND, and R. TÉMAM - *Convex Analysis and Variational Problems*, Siam, Philadephia, 1999.

[31] D. P. G. FOURES, N. DOVETTA, D. SIPP, and P. J. SCHMID - *A Data-Assimilation Method for Reynolds-Averaged Navier-Stokes-Driven Mean Flow Reconstruction.* J. Fluid Mech., 759, 2014, 404-431.

[32] A. GRONSKIS, D. HEITZ, and E. MÉMIN - *Inflow and Initial Conditions for Direct Numerical Simulation Based on Adjoint Data Assimilation.* Journal of Computational Physics ,242, 2013, 480-497.

[33] J. HADAMARD - *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. Dover, NewYork, 1923.

[34] K. HADJ-SASSI, and S. ANDRIEUX - *Vers une nouvelle stratégie d'estimation conjointe des paramètres matériaux et de l'état des structures par assimilation de données et recalage.* 18ème Congrès Français de Mécanique, Grenoble, 2007.

[35] B. HALPHEN, and Q. S. NGUYEN - *Sur les matériaux standards généralisés*. J. Mécanique, 14, 1975, 39-63.

[36] P. HEITZ, D. HEITZ, and E. MÉMIN - *Multiscale Regularization Based on Turbulent Kinetic Energy Decay for PIV Estimation with High Spatial Regularization.* 8th Int. symposium on particle image velocimetry-PIV 2009, Melbourne, Victoria, Australia, 2009.

[37] D. HEITZ, E. MÉMIN, and C. SCHNÖRR - *Variational Fluid Flow Measurements from Image Sequences: Synopsis and Perspectives*. Exp. Fluids, 48, 2010, 369-393.

[38] J.-B. HIRIART-URRUTY, and C. LEMARÉCHAL - Convex Analysis and Minimization Algorithms, I and II. Springer Verlag, 1993.

[39] J. KAIPIO, and E. SOMERSALO - *Statistical Inverse Problems: Discretization, Model Reduction and Inverse Crimes.* J. Computational and Applied Mathematics, 198, 2007, 493-504.

[40] R. E. KALMAN - *A New Approach to Linear Filtering and Prediction Problems.* Transactions of the ASME - Journal of Basic Engineering (Series D), 1960, 82: 35-45.

[41] K. C. KIWIEL - *Proximal Minimization Methods with Generalized Bregman Functions*. SIAM J. Control Optim. 35 (4), 1997, 1142-1168.

[42] M. Y. KOKURIN - *Stable Iteratively Regularized Gradient Method for Nonlinear Irregular Equations Under Large Noise.* Inverse Problems 22, 2006, 197-207.

[43] S. KULLBACK, and R. A. LEIBLER - *On Information and Sufficiency*. Annals of Mathematical Statistics 22 (1), 1951, 79-86.

[44] M. M. LAVRENTIEV - Some Improperly Posed Problems of Mathematical Physics. Springer, Berlin, 1967.

[45] F. X. LE DIMET, and O. TALAGRAND - *Variational Algorithms for Analysis and Assimilation of Meteorological Observations: Theoretical Aspects.* Tellus, 38A, 1986, 97-110.

[46] J. L LIONS - Optimal Control of Systems Governed by Partial Differential Equations. Springer Verlag, 1971.

[47] A. MOHAMAD-JAFARI, and M. DUMITRU - *Bayesian Sparse Solution to Linear Inverse Problems with Non Stationary Noise With Student-t Priors.* Digital Signal Processing, Special Issue on Bayesian Signal Processing: in Honour of William J. Fitzgerald, 2015.

[48] J. MORLIER, and D. BETTEBGHOR - *Compressed Sensing Applied to Modeshapes Reconstruction.* in Topics in Modal Analysis I, Volume 5, Part of the series Conference Proceedings of the Society for Experimental Mechanics Series,2012, 1-8.

[50] V. A. MOROZOV - *On the solution of functional equations by the method of regularization*. Soviet Math. Dokl. 7, 1966, 414-417.

[51] L. MUGNIER, T. FUSCO, J.-M. CONAN - *MISTRAL: a Myopic Edge-Preserving Image Restoration Method, with Application to Astronomical Adaptive-Optics-Corrected Long-Exposure Images.* J. Opt. Soc. Am. A, 21, 2004, 1841-1854.

[52] I. M. NAVON - *Practical and Theoretical Aspects of Adjoint Parameter Estimation and Identifiability in Meteorology and Oceanography.* Dynamics of Atmosphere and Oceans, 27, 1997, 55-79.

[53] D. L. PHILLIPS - *A Technique for the Numerical Solution of Certain Integral Equations of the Ffirst Kind.* J Assoc Comput Mach, 9, 1962, 84-97.

[54] L. I. RUDIN, S. OSHER, and E. FATEMI - *Nonlinear Total Variation Based Noise Removal Algorithms*. Physica D 60, 1992, 259-268.
[55] P. RUHNAU, and C. SCHNÖRR - *Optical Stokes Flow Estimation: An Imaging-Based Control Approach*. Exp. Fluids, 2007, 42: 61-78.
[56] P. RUHNAU, A. STAHL, and C. SCHNÖRR - *On-line Variational Estimation of Dynamical Fluid Flows with Physics-Based Spatio-Temporal Regularization.* In Pattern Recognition – 28th DAGM Symposium, volume 4174 of LNCS, Springer, 2006, 444-454.
[57] P. RUHNAU, A. STAHL, and C. SCHNÖRR - *Variational Estimation of Experimental Fluid Flows with Physics-Based Spatio-Temporal Regularization*. Measurement Science and Technology, 2007, 18: 755-763.
[58] F. SARTOR, G. LOSFELD, and R. BUR - *PIV Study on a Shock-Induced Separation in a Sransonic Flow.* Exp. Fluids Vol. 53, 2012, 815-827.
[59] C. SCHNÖRR - *Determining Optical flow for Irregular Domains by Minimizing Quadratic Functionals of a Certain Class.* Int. J. Computer vision, 6:1, 1991, 23-35.
[60] P. SMITH, S. L. DANCE, M. J. BAINES, N. K. NICHOLS, and T. R. SCOTT - *Variational Data Assimilation for Parameter Estimation: Application to a Simple Morphodynamic Model.* Ocean Dynamics, 59 (5), 2009, 697-708.
[61] A. SUTTON, and J. H. S. ORTEU - Image Correlation for Shape, Motion and Deformation Measurements: Basic Concepts, Theory and Applications. Springer, USA, 2009.
[62] R. TIBSHIRANI - *Regression Shrinkage and Selection via the Lasso.* J. of the R. Stat. Soc, Series B, Vol. 58, 1 , 1996.
[63] A. N. TIKHONOV and V. Y. ARSENIN - Solutions of Ill-Posed Problems. Wiley, New York, 1977.
[64] E. ROSENCHER - *Optical Diagnostics of Flows: an Introduction.* AerospaceLab Journal, 1, 2009, 1-2.
[65] G. VASILE, J. P. OVARLEZ, F. PASCAL, and C. TISON - *Coherency Matrix Estimation of Heterogeneous Clutter in High-Resolution Polarimetric SAR Images, Geoscience and Remote Sensing.* IEEE Transactions on 48 (4), 2010, 1809-1826.
[66] A. VLASENKO, and C. SCHNÖRR - *Physically Consistent Variational Denoising of Image Fluid Flow Estimates.* IEEE Trans Image Process,19(3), 2010, 586-952.
[67] M. J. WAINWRIGHT - *Sharp Thresholds for High-dimensional and Noisy Sparsity Recovery Using I1-Constrained Cuadratic Programming (Lasso).* IEEE Trans.IT, v.55, no. 5, 2009.
[68] R. YEGAVIAN, B. LECLAIRE, F. CHAMPAGNAT, and O. MARGUET - *Performance Assessment of PIV Super-Resolution with Adjoint-Based Data Assimilation.* 11th Int. Symposium on Particle image velocimetry – PIV15, Santa Barbara, California, Sep 14-16, 2015.

AUTHOR

Stéphane ANDRIEUX received his Engineering diploma from the *Ecole Nationale des Ponts et Chaussées*, France, in 1980, and later his PhD in 1983. He joined the R&D division of *Electricité de France* in 1985 and led various research groups and mechanics departments there. After being the Scientific Director of the EDF R&D division, he is now the General Scientific Director of ONERA and professor at *Ecole des Ponts Paristech*. His work is focused on damage mechanics and inverse problems. In 2006, he received the Dechelle Prize from the French Academy of Sciences.

E. P. Busso, D. Boivin, D. Lévêque
(ONERA)

E-mail : esteban.busso@onera.fr

DOI : 10.12762/2016.AL12-02

On Recent Advances in Microstructural Characterization and *In-Situ* Testing Techniques to Study Material Behavior

The deformation fields within grains in polycrystalline materials are generally highly heterogeneous and can be the precursors to the nucleation of micro-cracks or cavities. Such behavior is conditioned by microstructural features, such as grain structure, texture, morphology, and size. The understanding of the inter-relationship between the material's microstructural characteristics and complex deformation and damage phenomena is crucial to the formulation of structural integrity assessments of engineering components, since they constitute the physical bases of the required constitutive models of deformation and failure.

In this work, an overview is provided of recent experimental techniques developed at ONERA and elsewhere to characterize the microstructure and to measure the deformation of metallic materials. Some of the most promising characterization, measurement and identification techniques and methods used to develop, calibrate and validate physics-inspired constitutive models will be discussed, and illustrative examples will be given. It will be shown that the individual use of traditional interferometric, characterization and testing techniques is being replaced by the combined use of state-of-the-art techniques based on image correlation (2D and 3D), *in-situ* and *ex-situ* scanning electronic microscopy (SEM), EBSD, amongst others. Discussions about future trends, including the great potential offered by the combination of *in-situ* techniques at various scales with real time computations, are presented.

Introduction

The development of new materials and the optimization of existing ones require an understanding of how their microstructures at all relevant scales affect their mechanical and functional properties. Such relationships can be very complex, since they generally involve physico-chemical processes that act during their entire life cycle. In order to reduce development and characterization costs, accurate multi-physics and multi-scale models based on sound physical principles are needed. In addition, suitable characterization, identification and validation methods are required at different microstructural scales. These may involve measurement techniques that need to be coupled with the numerical integration of complex constitutive material models (e.g., see [19]), sometimes in real time. Handling the resulting large amount of data that need to be stored is becoming a major challenge, since the resolution and accuracy of detailed microstructural measurements is continuously increasing. An additional difficulty is the efficient visualization of the very large amount of data that needs to be handled. For instance, a single 3D image obtained by computed tomography and reconstructed numerically may require up to 16 GBytes of computer disk space [10]. Furthermore, the measurement of full fields by techniques such as image correlation is equally demanding in terms of digital storage capabilities [23]. However, it is the simultaneous use of multiple observation and measurement techniques such as those described above that offers exciting new opportunities for the development and calibration of physics-inspired constitutive models of materials with complex heterogeneous microstructures.

In this work, the challenges associated with some of the most promising characterization, measurement and identification techniques to study the mechanical behavior of materials will be addressed.

However, since this overview is not intended to be exhaustive by any means, special emphasis is placed on 2D and 3D imaging and *in-situ* micromechanical scanning electronic microscopy (SEM) testing techniques, coupled with electron back-scattered diffraction (EBSD), developed at ONERA and elsewhere. Other important experimental techniques, such as transmission electronic microscopy (TEM), energy-dispersive X-ray spectroscopy (EDS), thermography and tomography, should be combined with those to be discussed next, in order to address a broader range of physico-chemical phenomena.

2D and 3D Measurement Methods

The study of material behavior has traditionally relied on either *in-situ* or *ex-situ* surface measurements of deformation fields, as previously discussed, and on post mortem characterizations of the material microstructure. Since the mid 90s, traditional interferometric techniques (e.g., speckle, Moiré, holography) to measure deformation fields have gradually been replaced by a new family of techniques based on image correlation: 2D-Digital Image Correlation (2D-DIC) (*i.e.*, correlations of images of a surface acquired with a single camera) (e.g., [10][13]), and by 3D-DIC (*i.e.*, measurements of 3D surface shapes and displacements using two cameras) (e.g., [12] [15]). However, the pertinence of relying on surface information to interpret bulk phenomena is not sufficient in many instances. In the last decade, the ability to visualize the volume of a material in three dimensions (3D) in a non-destructive way at the scale of a few micro-metres or less has opened up exciting new ways to study and characterize the microstructure and behavior of a broad range of materials [28]. An example is the detailed local microstructural information that can be obtained about, for instance, grain morphology, defect population, and damage, using tomography through either an X-ray source of a synchrotron or a laboratory scale micro-tomograph [5]. The data associated with the digitized material volume obtained from such 3D imaging techniques are then combined with suitable algorithms to identify the morphology of the microstructural features of interest and local displacement fields. The resulting techniques are known as either Digital Volume Correlation (DVC) or Volumetric DIC (V-DIC).

In what follows, a discussion of the main methods used for 2D and 3D digital image correlation at ONERA will be presented and compared with current practice. This will include details about efficient algorithms and parallel computing implementation, in particular the fast option offered by Graphic Processing Units (GPUs). Illustrative examples on the use of 2D and 3D-DIC in aeronautical materials and structures in terms of both computational performance and accuracy will be presented and discussed.

Basic principles and approaches

Full-field measurements in solids using imaging tools follow either classical global [11] and local [3] approaches, and the pixel-wise displacement estimate (referred to as "dense" by the authors) approach proposed recently by ONERA [15], which is a particular type of local approach. Their main differences concern the choice of reference domain used. To illustrate that point, let us recall the basic principle on which DIC is based.

DIC relies on a cross-correlation to measure shifts in datasets by finding the maximum of the correlation between pixel intensities on two or more corresponding images. If a vector $\boldsymbol{x}$ defines the position of a point in the reference image, and $\boldsymbol{u}(\boldsymbol{x})$ the displacement vector associated with the new position of that point in the new configuration, then the function that defines the characteristics of the point (e.g., the intensity of the pixel's grey-scale at that point) in the reference image can be expressed as,

$$f(\boldsymbol{x}) = g\left(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})\right) \tag{1}$$

where g is the corresponding gray-scale value in the image's new deformed configuration. The main objective of all image correlation algorithms is to determine the transformation, $\boldsymbol{u}(\boldsymbol{x})$, in the region of interest, Ω, of the reference image, knowing the values of f and g. The correlation procedure must be carried out in a certain domain and solved by minimizing a given quantity. The most commonly used minimization quantity is that defined by the least square difference between the gray-scale values of the images over the domain of interest, Ω. Here,

$$\eta = \int_{\Omega} \left[f(\boldsymbol{x}) - g\left(\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x})\right) \right]^2 d\boldsymbol{x} \tag{2}$$

However, the problem is ill-posed, since the available information, e.g., the gray-scale values of the pixels, is insufficient to compute the displacement vector, $\boldsymbol{u}(\boldsymbol{x})$, that is, the number of unknowns is greater than the number of equations. To overcome that problem, the correlation procedure must rely on additional hypotheses associated with the definition of the domain to be correlated.

In classical local approaches, the domain Ω is defined by small correlation windows in the region of interest, each window containing more than one pixel, the number determined by both the spatial resolution desired and the degree of overlapping between windows, if any. This issue will be further discussed below. In contrast, the domain Ω in global approaches is defined by the whole region of interest, rather than just by a small window. The discretization of the global region is generally done following a similar scheme to that of the finite element method. Thus, the displacement fields can be interpolated according to a relation of the form,

$$\boldsymbol{u}(\boldsymbol{x}) = \sum_{n \in N} u_n \boldsymbol{\Psi}_n(\boldsymbol{x}) \tag{3}$$

Here, u_n are the degrees of freedom and $\boldsymbol{\Psi}_n$ are the chosen interpolation vector functions. The solution of the correlation problem in the global approaches requires the minimization of Eqn. 2 assuming the functional form of the displacements given by Eqn. 3. This is done by seeking the optimal values of the unknown displacement coefficients, u_n. However, in local approaches, Eqn. 3 is not required to carry out the correlation itself, but it is instead used as a post-processing tool to smooth the estimated displacement fields.

In order to deal with the main challenges resulting from the manipulation of the huge amount of data obtained through 2D and 3D imaging, fast processing algorithms such as those originally developed for particle image velocimetry (PIV) used in experimental fluid mechanics, have recently been developed for 2D-DIC or 3D-DIC [1]. One such

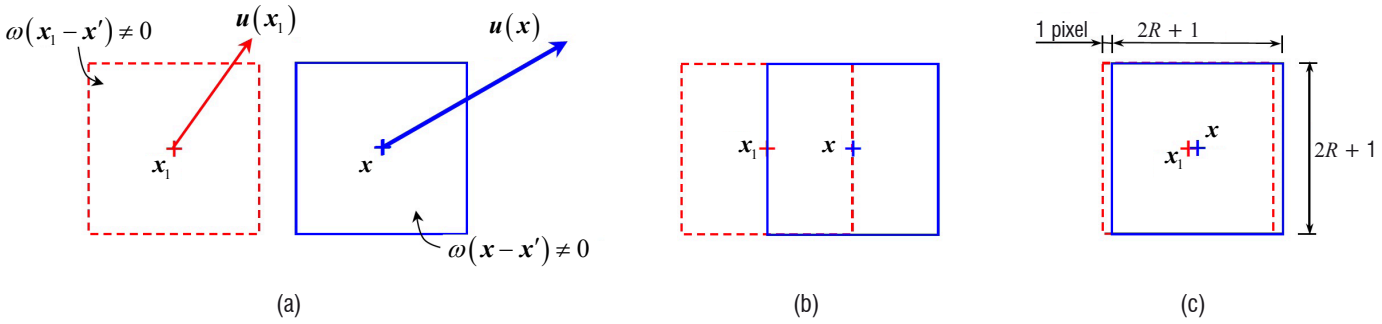


Figure 1 – Illustration of the possible relative configurations between two neighboring correlation windows centered at pixels x and x_1, for which the displacements $u(x)$ and $u(x_1)$, respectively, are calculated: (a) non-overlapping and (b) partially overlapping window configurations typical of classical local DIC approaches, and the (c) maximum possible overlapping window configuration of the FOLKI-D method. Each local window is delimited by its corresponding weight function, ω

contribution is the recent work by Le Besnerais *et al*. [15], which describes ONERA's fast and promising image correlation method for solid mechanics applications, named FOLKI-D (a French acronym for the Iterative Lucas-Kanade Optical Flow). The FOLKI-D method [14] is an extension of Lucas' and Kanade's original algorithm [16] adapted for PIV in [7].

In contrast to the classical global DIC approach, ONERA's pixel-wise displacement estimate approach (FOLKI-D) [15] is a local one, since it relies on the information associated with a window centered on a particular pixel to calculate its displacement. The minimizing quantity is similar to that given in Eqn. 2, but it is performed over a single window at a time using a weight function $\omega(x - x')$ – with ω typically representing a $(2R+1)^2$ square shaped window, and R being the desired scaling parameter. Then,

$$\eta_x(u) = \sum_{x' \in \Omega} \omega(x - x') \left[f(x') - g(x' + u(x)) \right]^2 \qquad (4)$$

The minimization of Eqn. 4 is then carried out using a Gauss-Newton iterative algorithm.

Most classical local approaches consider correlation windows that either do not overlap at all, or that exhibit small partial overlaps. In contrast, FOLKI-D estimates the displacement of each pixel by considering that neighboring correlation windows exhibit the maximum possible overlap between them. Short of these being 100% coincidental, such configuration implies that the spacing between neighboring windows is just a single pixel. This point is illustrated in Figure 1, which shows the possible relative configurations between two neighboring correlation windows centered at pixels x and x_1, with (a) representing non-overlapping and (b) representing partially overlapping window configurations, both typical of classical local DIC approaches, and (c) the maximum possible overlapping window configuration of the FOLKI-D method.

It should be pointed out that, even though the minimization of a relation of the type given by Eqn. 2 is performed independently for each pixel, all minimization operations can, in principle, be performed simultaneously on all pixels of the image. However, the use of such an approach on the maximum possible overlapping window configuration illustrated in Figure 1(c) would result in an excessively large number of redundant computations to update each displacement vector independently, since a pixel shares most of its corresponding window with its neighbors. In order to overcome this problem, the FOLKI-D method relies on a first order Taylor expansion of Eqn. 4 and on the reformulation of the problem in terms of global convolutions and pixel-wide operations. As a result, the optimization only involves regular and local operations over the whole image, leading to a method which is highly efficient and suitable for modern parallel computing architectures, such as GPUs. See [15] for further details.

The pixel-wise displacement estimate DIC approach offers another considerable advantage with respect to the classical global approaches, since there are no a priori assumptions to be made about the nature of the displacement fields, with the corresponding gain in computing efficiency, accuracy and potential for parallelization using, for instance, GPUs. These issues will be discussed later in the text, in one of the case studies (multi-scale deformation behavior of a woven composite).

Another promising application of DIC or DVC is when it is combined with finite element (FE) modeling and remeshing techniques to locate and introduce into the FE model microcracks in either quasi-real time or as a post-processing. Recent work carried out at ONERA [8] has proposed a method whereby the crack growth path can be identified from the image processing by following the peak of a representative local material degradation measure, such as the magnitude of a displacement discontinuity between neighboring pixels.

For additional discussions of current issues and future trends in 2D-DIC, 3D-DIC and DVC, refer to [29].

Representative experimental test cases

In this section, two experimental test cases reported in [15] involving 3D surface displacement field measurements using ONERA's FOLKI-D algorithm are presented.

The first test case consists of a stereographic 3D-DIC measurement during the uniaxial compression of a carbon-epoxy laminate composite panel after an impact test. Such a testing procedure is widely used by the aeronautic industry for certification purposes and to demonstrate repair process worthiness. In this test, two images are recorded at each instant, to compute both the specimen shape and

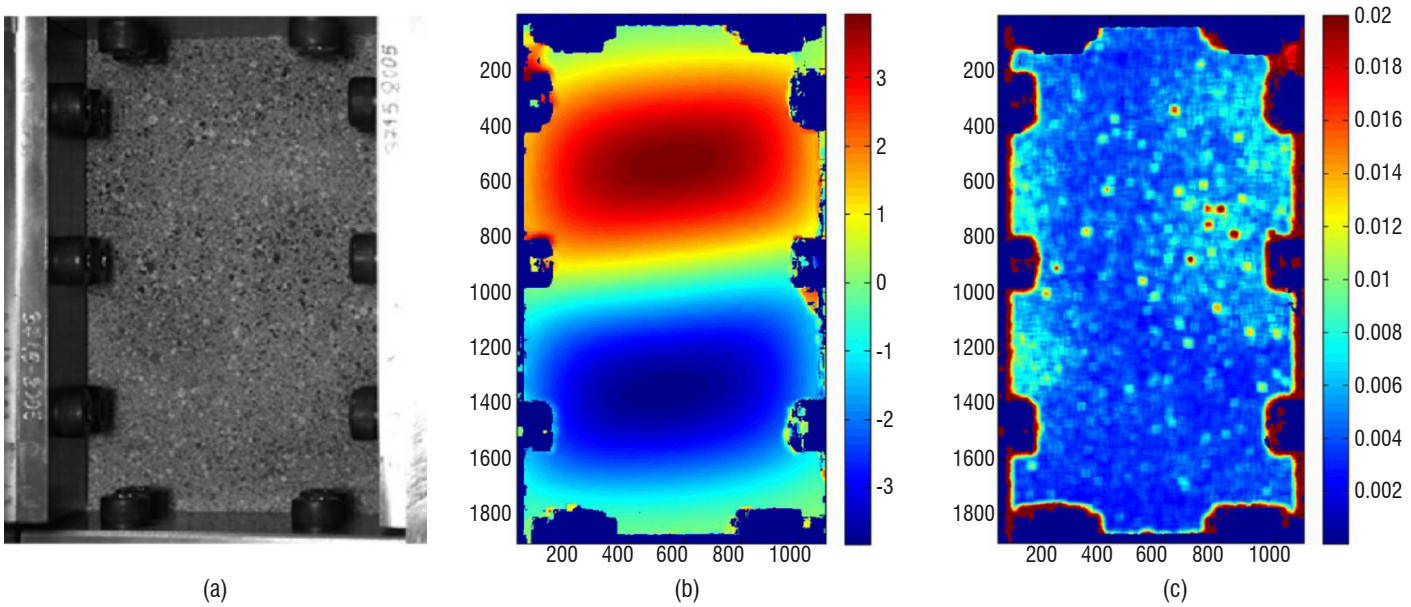


Figure 2 – Uniaxial compression of a carbon-epoxy laminate composite panel after an impact test: (a) specimen, (b) out-of-plain displacement component, and (c) uncertainty estimation obtained with FOLKI-D [15]

the 3D displacement field. Figure 2 shows (a) an image of one of the specimen's sides, (b) a plot of the measured out-of-plain displacement component, and (c) the distribution of the uncertainty estimation obtained with FOLKI-D. The scales in Figures 2(b) and (c) are in pixel units. It can be seen from Figure 2(b) that the measured out-of-plain displacement component distribution implies that specimen buckling has occurred. A comparison between the accuracy of the computed displacements and those obtained using a commercially available code showed that they differed by less than 2.5%. However, the computing time to process the two stereographic 2 MPixel images of FOLKI-D was found to be 24 times faster than the commercial code. Another advantage of FOLKI-D is that it enables information close to the specimen borders to be obtained due to the large number of the resulting displacement vectors (2 million for the window shown in Figure 2's case, against 32 k vectors for the commercial code). Finally, the uncertainty distribution of the computed displacement at each pixel plotted in Figure 2(c) is obtained from a classical local statistical analysis.

The second test case is intended to illustrate the robustness of FOLKI-D when dealing with degraded image textures, as is usually the case when measurements are conducted in high-temperature environments. Here, the growth of a fatigue crack at 900°C in a polycrystalline superalloy plate with a center hole is investigated. This test is representative of local conditions seen in aero-engine combustion chambers. Figure 3 shows the crack configurations after (a) 47000 and (b) 53000 cycles at the maximum cycle load, and (c) and (d) the vertical displacement component fields measured by FOLKI-D

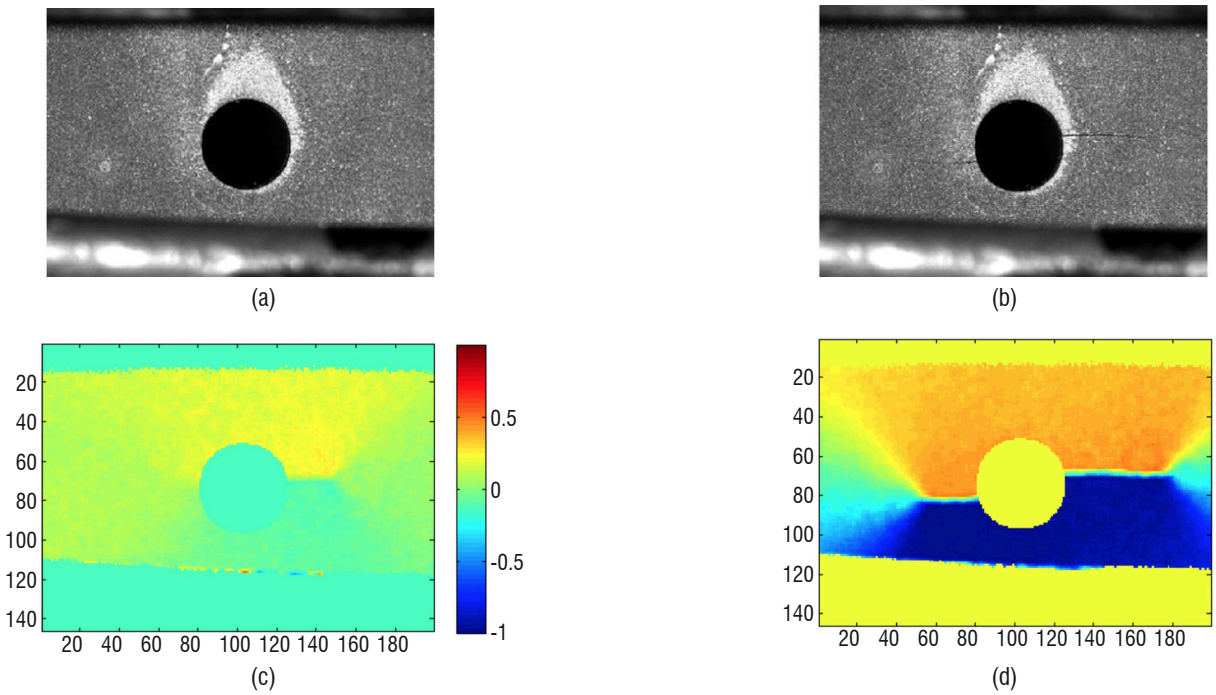


Figure 3 – Growth of a fatigue crack at 900°C in a polycrystalline superalloy plate with a center hole: crack configurations after (a) 47000 and (b) 53000 cycles at the maximum cycle load, and (c) and (d) vertical displacement component fields measured by FOLKI-D corresponding to the configurations in (a) and (b), respectively [15] (the applied load is in the vertical direction)

corresponding to the configurations in (a) and (b), respectively [15]. It should be pointed out that, despite the low contrast on the specimen surface, the FOLKI-D results were capable of providing an accurate estimate of the actual crack lengths and crack tip positions by relying on a local discontinuity criterion. As discussed in [15], high-temperature DIC measurements require that the light emission from the heated surface be minimized by the use of suitable filters [9] and by ensuring that the time between images is sufficiently short to avoid the development of oxidation-related surface texture. Nevertheless, uncertainties in the measured displacements in some regions where the local variance of the calculated residuals is high were found to be up to 0.05 pixel.

Combined Local Microstructural Characterization and Measurement Methods

Scanning electronic microscopes (SEM) are not just sophisticated high-resolution instruments to observe the surface of samples at the sub-micron scale. They can also be used to identify the crystallographic characteristics and chemical composition of the sample when coupled with electron back-scattered diffraction (EBSD) and energy-dispersive X-ray spectroscopy (EDS) or wavelength dispersive spectroscopy (WDS) techniques, respectively.

An analysis by EBSD provides information about the individual grains in the material. A typical inversed pole figure (IPF) produced by EBSD of a typical AISI 316L-type austenitic stainless steel is shown in Figure 4 [17]. Here, Figures 4(a) and (b) are IPFs of the microstructure in the annealed and cold-rolled state, respectively. The close-up image of the microstructure in the cold-rolled state in Figure 4(b) shows also the slip traces within the grains.

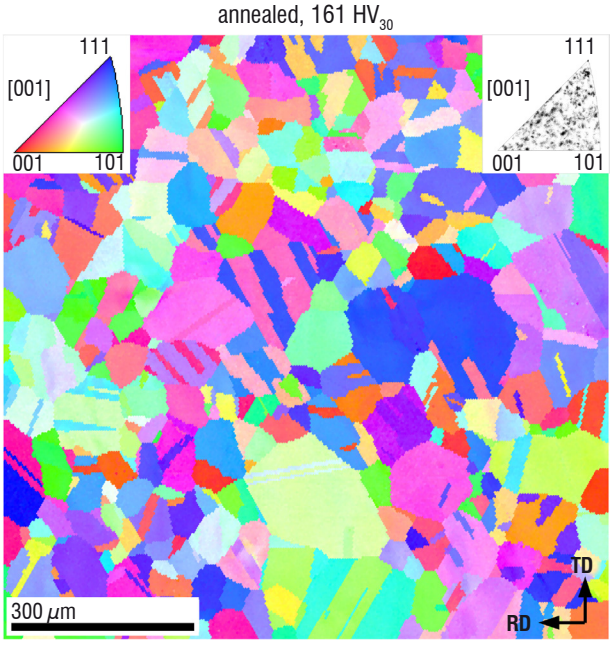
The microscope's electronic beam can also be used as a micro-lithography electronic tool to create surface patterns on the sample to be tested and thus serve as a micro-extensometer to measure its deformation during either *in-situ* or *ex-situ* micromechanical tests. This capability to link several observation and characterization techniques has recently opened up a wide range of opportunities to study the behavior of materials with heterogeneous microstructures. The simultaneous acquisition of crystallographic information and deformation fields enables a qualitative and quantitative understanding of the material behavior, and provides local information to validate local constitutive material and life prediction models.

In what follows, two examples are given of recent original work carried out at ONERA based on a multiple material characterization of high-temperature alloys used in aerospace applications.

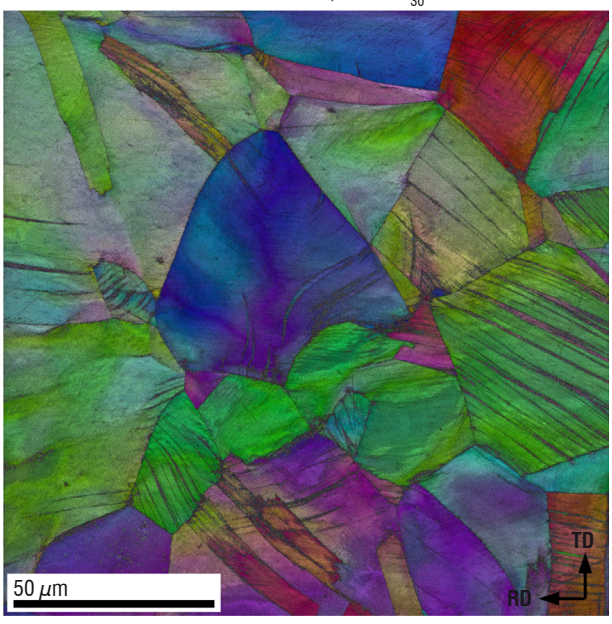
Combined EBSD and *in-situ* micromechanical SEM testing to measure high-temperature grain boundary sliding in a Ni-base superalloy

Intergranular creep is one of the important mechanisms that conditions the behavior and lifetime of high-temperature gas turbine disc and blades. Generally, the amount of grain boundary (GB) sliding which contributes to the macroscopic time-dependent deformation is quite small, of the order of tens of nanometers, thus specialized microscopy techniques are needed to measure such deformation. One of the common techniques used for that purpose is known as "marking", whereby some visible markers are deposited on the material surface to serve as a visual reference and thus enable small relative GB displacements to be measured during deformation. The technique developed at ONERA will be discussed later in the text.

A combination of EBSD, which provides information about the local crystallographic deformation, with marking at the grain scale was undertaken at ONERA to study the intergranular GB sliding due to creep in a Ni base superalloy at 850°C. Even though the method is relatively simple, its development requires expertise in processing, heat treatments, *in-situ* micro-extensometry within the chamber of an SEM, analysis by EBSD, and complex data processing.



(a)



(b)

Figure 4 – Microstructure of an austenitic stainless steel observed by EBSD: (a) annealed state shown in inverse pole figure (IPF) and (b) cold-rolled state, in a close-up view, shown by a superposition of the IPF and image quality maps [17]

The results to be reported here are some of those obtained as part of the work by Thibault [31] [32]. Note that the discussion of these representative results is intended to illustrate the potential and accuracy of advanced characterization and *in-situ* testing techniques, and it does not constitute an exhaustive study of superalloy creep behavior. For more detailed and complete studies conducted in this area at ONERA, it is suggested that the reader refers to [25, 26, 27, 31, 32].

Identification and analyses of the grain boundary characteristics by EBSD

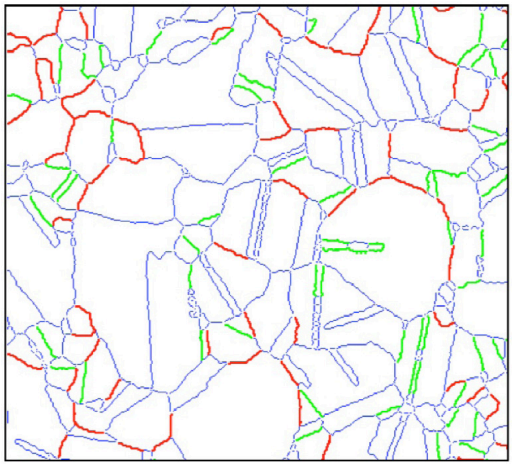
In order to study the characteristics of grain boundaries, the coincidence site lattice (CSL) model is used. The model enables the nature of the GBs, such as highly misoriented ones ($\theta \geq 15°$) to be defined. Generally, GBs are classified by their coincidence index, denoted by Σ. For instance, grains that share a $\Sigma 5$ boundary contain $1/5^{th}$ of atoms located on common lattice sites. Thus, boundaries are said to be generic if $\Sigma > 29$, and special if $\Sigma \leq 29$ [21]. Amongst the special boundaries, it is worth identifying those where the degree of lattice coincident sites is the maximum, namely the $\Sigma 3$ boundaries, often associated with twins and known as twin boundaries. The expectation is that boundary structures with low Σ values will be more regular than a general boundary and hence would have low energies because of their good atomic fit. This consideration is central to the concept of "grain boundary engineering", which relies on maximizing the fraction of CSL boundaries as a way of increasing the material strength.

The processing of EBSD information can take many different forms. For instance, it can be used to identify and visualize a particular type of grain boundary, such as the $\Sigma 3$ boundaries and generic ones, as shown in the example of Figure 5. Here, Figure 5(a) shows the GB network extracted from the EBSD information: red GBs are of the $\Sigma 3$ type, green ones are generic ones, both are of between 20 to 40 μm in length, and blue ones are generic GBs of other sizes. Figure 5(b) shows a distribution of the GB lengths identified in Figure 5(a). Other types of information can be inferred from EBSD measurements, depending on the criteria used. For instance, EBSD software enables the identification of individual GB intersections, and the fraction of GB intersections connected to 0, 1, 2 or 3 boundaries of the $\Sigma 3$ type, for instance. In reality, such intersections play a particular role in intergranular sliding as their resistance to sliding is greater than other types of GBs. Thus, their spatial distribution and the different nature of GB intersections constitutes a useful piece of information to study grain boundary creep.

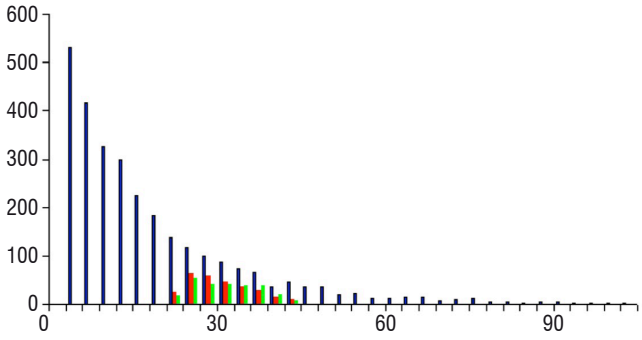
Displacement and strain field measurements at high temperatures by micro-extensometry

Micro-extensometry is a technique to measure local displacement and strains, well suited to study intergranular phenomena. In a SEM, it is based on image correlation techniques of the observed sample's polished surface before and after deformation. As previously discussed, if the surface does not contain natural markers that can be identified on the SEM image, then it is necessary to introduce artificial features on the surface, for instance by engraving [22][4], internal marking [2], or film deposition [18][20][24]. In the work done at ONERA, the contrast was achieved by depositing a surface grid by electronic micro-lithography. The SEM images of the surface produced in this way were obtained from before and after *ex-situ* creep tests. The processing is done by a suitable image correlation algorithm that is capable of tracking the position of the micro-lithography markers, in order to calculate the displacement fields. From the latter, the strain fields and the amplitude of the GB slidings are extracted.

The creep specimen marking by electronic micro-lithography was performed within the SEM chamber and involved the deposition of a $\approx$ 10 nm thin film of a different material on the polished sample surface following a particular pattern. In this study, it consisted of a 0.3 x 5 μm grid over a 320 x 320 μm area. The different stages involved in the micro-lithography process are shown schematically in Figure 6. It is worth noting that the grid material is deposited by either evaporation or spraying and must satisfy several criteria. For instance, its atomic number must be significantly different from that of the sample material in order to provide a good contrast in the SEM images obtained by Back Scattered Electron (BSE) signals. In addition, the microgrids were exposed to temperatures of 700°C during the creep test. Even though creep tests were performed in a vacuum, there is always a small partial oxygen pressure that may compromise the microgrid stability. Thus, the latter is another important requirement for the micro-grid material. Finally, the microgrid must be capable of



(a)



(b)

Figure 5 – (a) Grain boundary network with boundaries of the $\Sigma 3$ type in red and generic ones in green, both of between 20 to 40 μm in length, and with the blue GBs being generic ones of other sizes, and (b) distribution of the GB lengths identified in (a)

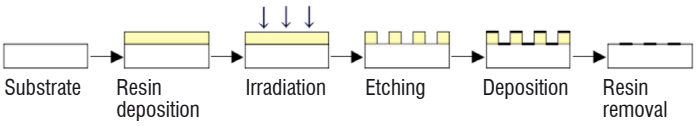


Figure 6 – Principal steps of the SEM micro-lithography deposition process

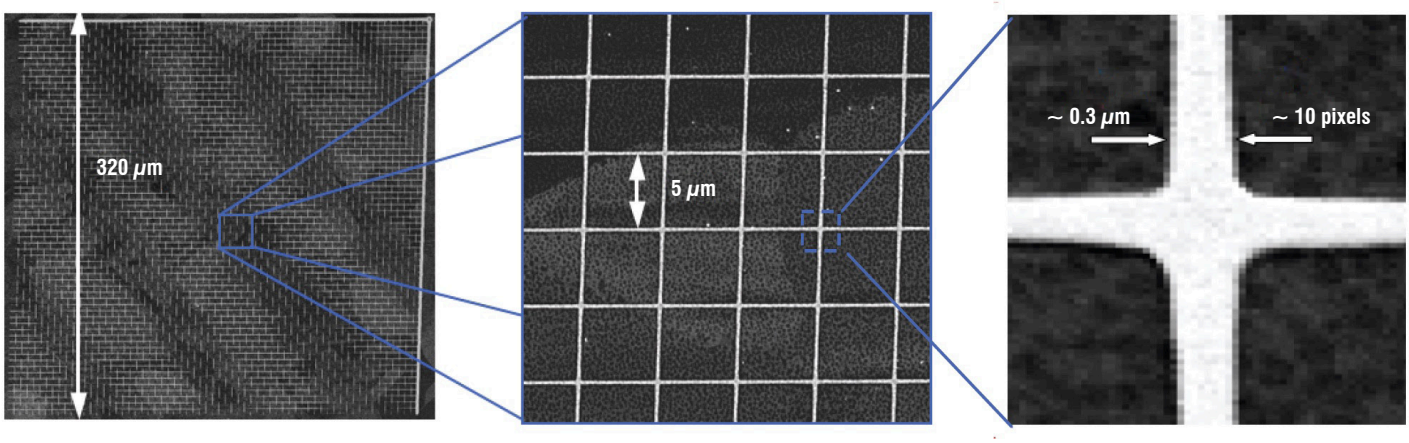
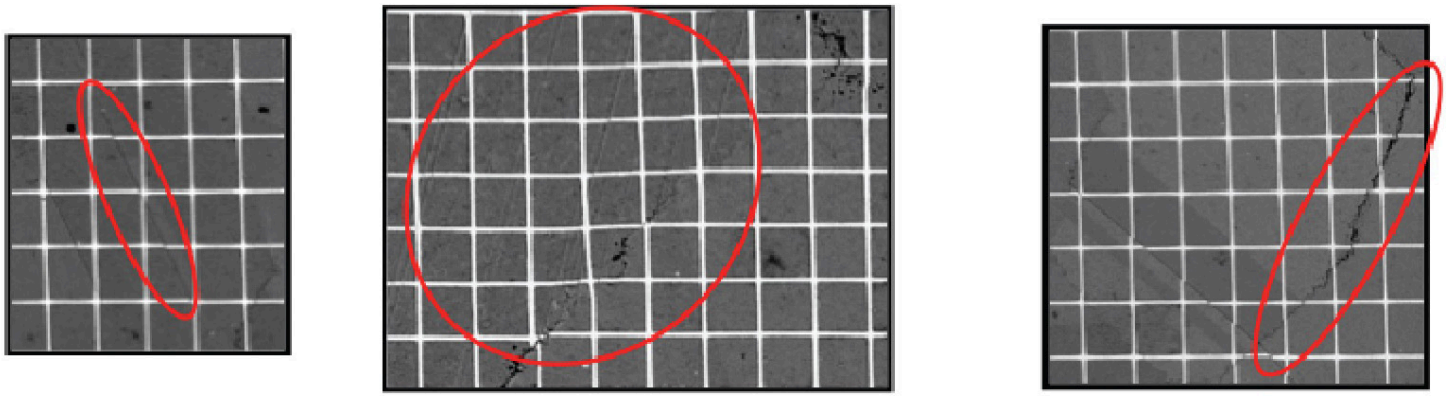


Figure 7 – SEM image of the microgrid produced by micro-lithography deposition in backscattered contrast



(a) (b) (c)

Figure 8 – Deformation phenomena revealed by the deformed microgrid: (a) intragranular shear, (b) localized plastic deformation or slip near a GB, and (c) intergranular shear

accommodating the substrate deformation without compromising its adhesion and mechanical integrity. Amongst the various options considered here, the material that best fulfilled the above requirements was HfO_2, deposited as a thin film of $\approx$ 10 nm by cathodic spraying.

The image correlation (IC) method used to obtain the displacement and strain fields requires high resolution images with a good image-to-noise ratio. For that purpose, the electron beam current and acquisition time must be optimized. It should also be noted that, since the microgrid nodes constitute the markers that the IC will follow, they should be represented by a sufficient number of pixels. For instance, for the 0.3 μm wide microgrid lines shown in Figure 7, 7 to 8 pixels are typically required across the line width. Then, the pixel size that would be required in this case is approximately 0.04 μm in size. Other issues that condition the pixel size choice are the image size and acquisition time desired. For the 320 x 320 μm image size shown in Figure 7, a minimum of 8000 x 8000 pixels will be required. The acquisition time for such an image size using the SEM system described here is of the order of 30 min.

The creep deformation of the superalloy at high temperatures reveals itself in different forms, through the microgrid deformation as well as the shearing and curvature of its originally straight segments. In Figure 8, it is possible to distinguish (a) intragranular shear, (b) localized plastic deformation or slip within the grain or near GBs, and (c) intergranular shear.

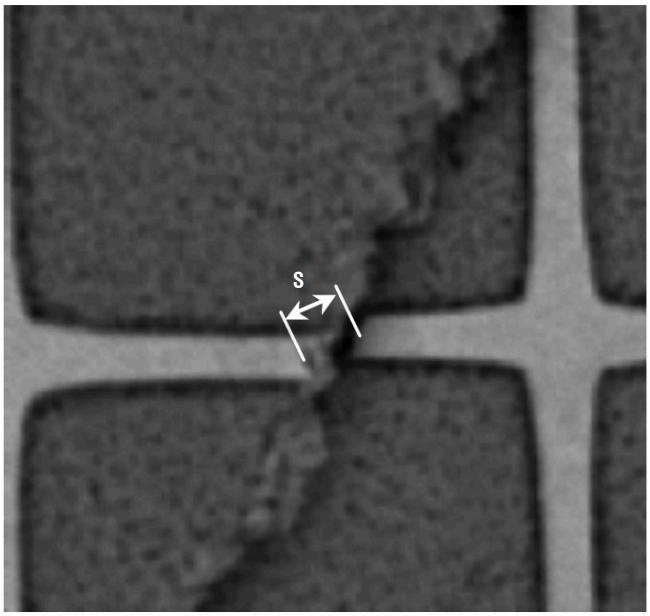
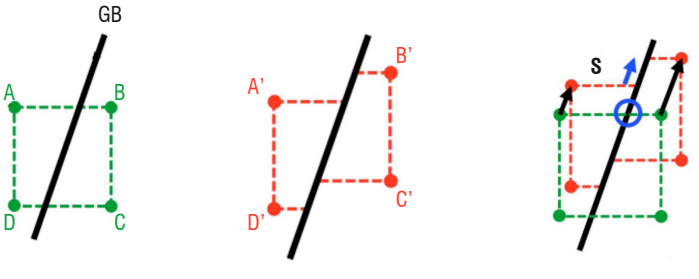


Figure 9 – Typical sheared grain boundary and attached microgrid with an associated shear displacement vector



(a) Undeformed (b) Deformed (c) Shear vector, **s**

Figure 10 – Determination of the shear displacement vector at a GB-microgrid intersection

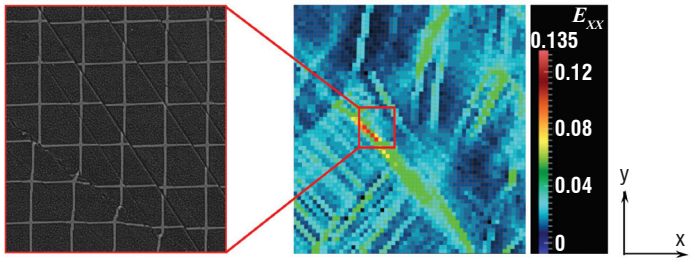


Figure 11 – Example of the local distribution of the E_{xx} strain component calculated from the microgrid after deformation. The zoomed image shows the deformed microgrid corresponding to the framed square in the right hand side image

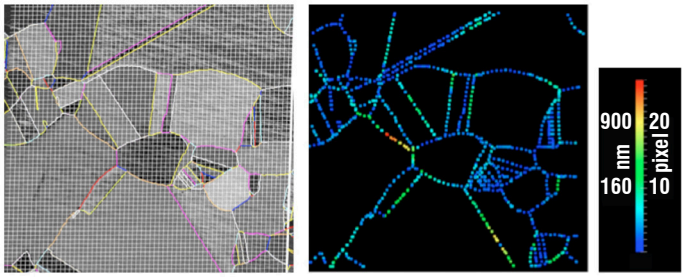
The calculation of the strain associated with a sheared grain boundary is based on the displacement vector which characterizes the direction and amplitude of the localized shear deformation of the straight microgrid segment. This is given by the vector **s** in the deformed grid shown in Figure 9. The determination of the shear displacement vector **s** in the discrete microgrid node space relies on a comparison between the undeformed and deformed microgrids, as illustrated in Figure 10. It is then possible to represent the shear displacement amplitude across each microgrid-grain boundary intersection according to a color code associated with a reference scale.

Figure 11 shows a typical example of the local distribution of the strain component along the x-axis, calculated from the deformed microgrid. A part of the latter can be seen in Figure 11's zoomed region.

Combination of deformation and crystallographic data

The combination of the local crystallographic characteristics with the measured deformation fields contained in the EBSD and SEM images, respectively, require that they be available in the same reference system provided by the microgrid. However, this is not something that can be easily done as the operational conditions are quite different. For instance, in EBSD the sample is inclined at an angle of 70° with respect to the electron beam, whereas the latter is normal to the sample during the SEM imaging. It is also impossible to obtain both types of images of exactly the same physical region. Nevertheless, the most important difference is the size of the digital sampling; in this work, they were typically 2000 x 2000 pixels for the EBSD, and 8000 x 8000 pixels for the SEM images.

In order to make the EBSD and SEM images compatible, they need to be mapped into le same numerical space. The mapping of the



(a) (b)

Figure 12 – (a) Superposition of the SEM and GB network images, and (b) corresponding plot of the GB shear displacement amplitudes

EBSD images into the resolution of the SEM images is done by a homography method based on the selection of point correspondences in the associated image pair. Here, four singular points that have corresponding ones in the two image domains are defined, and the subsequent superposition of the two images enables the validation of the procedure by data correlation. This can be seen in Figure 12(a), which shows an example of the superposition of the SEM and GB network images after the mapping procedure.

Once the mapping is done, the process of measuring the GB shearing can begin. It is convenient to assign a color code to the magnitudes of the GB shear displacements, in order to obtain quantitative plots of the microstructural region of interest. Figure 12(b) shows such a plot of GB shear displacement amplitudes for the image identified in Figure 12(a). In view of the fact that individual grains can now be identified by their own morphological and crystallographic characteristics, a range of additional information can be extracted. For instance, the mean shear amplitude of different types of grain boundaries, such as the proportion of $\Sigma 3$ and other generic type of GBs. An example is given in Figure 13, which shows the effect of the type of grain boundary and creep temperature on the GB mean shear displacements measured after creep tests at 350 MPa. The results reveal that the $\Sigma 3$ boundaries in Ni-base superalloys are more creep-resistant than those with greater coincidence indexes, with the GB shear displacement decreasing with temperature. One reason for such behavior is that recovery processes at high temperatures, such as that arising from dislocation annihilation, is much slower in CSL boundaries than in general boundaries due to the associated lower point defect diffusion kinetics. An additional mechanism reported by [30] is that dislocations accumulate at CSL boundaries, giving rise to internal stresses, which inhibits the dislocation motion.

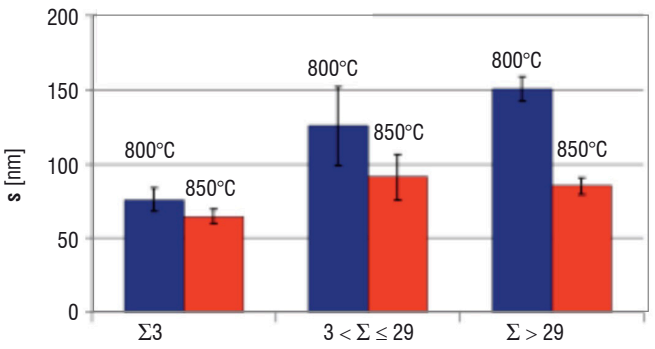


Figure 13 – Effect of the type of grain boundary and temperature on the GB mean shear displacement after a creep test at 350 MPa

Additional creep tests at 700°C and 800°C under an applied stress of 700 MPa were also carried out to investigate the effect of temperature on the deformation of the superalloy. Figures 14(a) and (b) show a comparison between the measured mean GB shear displacement vectors at these two temperatures. It can be seen that GB shear displacement decreases with temperature, from a maximum value of 113 nm at 700°C to 60 nm at 800°C. The contribution of GB sliding to the overall creep deformation along the loading direction was found to decrease from 33 % to 19 % for the same increase in temperature. This somehow counter-intuitive result is due to the known decrease of γ'-precipitate shearing activity by the γ-phase dislocations with temperature. Such shearing promotes deformation localization within narrow and planar slip bands, and leads to grain boundary shearing due to the high levels of stress concentration that develop at slip band- grain boundary intersections [27]. As temperature increases, the deformation becomes more homogeneous as matrix dislocations are now able to circumvent the precipitates through thermally activated mechanisms (e.g., climb, cross-slip), and the amplitude of grain boundary sliding decreases.

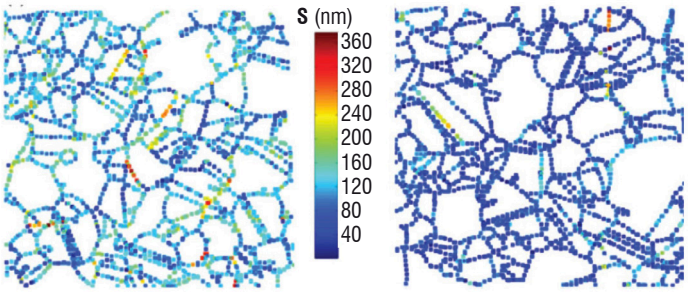


Figure 14 – Comparison between measured amplitudes of GB shearing displacements at two different temperatures: (a) 700°C and (b) 800°C

It is worth noting that superalloy microstructures can also exhibit very different GB morphologies depending on their metallurgical history. An interface or GB exhibiting a rather smooth morphology is likely to behave differently at high temperatures where GB sliding becomes an important deformation mechanism than one containing a rough or tortuous morphology. The results of an investigation of such GB morphology effects are presented in Figure 15, which shows a comparison between the measured amplitudes of GB shearing displacements in superalloy microstructures with the two different GB morphologies shown in the insets: (a) smooth and (b) tortuous. It can be seen that the inhibition of GB sliding due to the introduction of tortuous GBs leads to a significant decrease in the amplitude of the GB sliding. Such beneficial effect of GB morphology on the high-temperature behavior of superalloys is one of the GB engineering parameters being considered at ONERA for the next generation of polycrystalline superalloys.

In this section, the application of the combined SEM and EBSD techniques to study intergranular sliding in a Ni base superalloy under high temperature creep conditions has revealed the potential of such techniques to study and quantify complex deformation phenomena at the grain level. However, its development and setup require a great deal of experimental effort and know-how, be it in the marking of

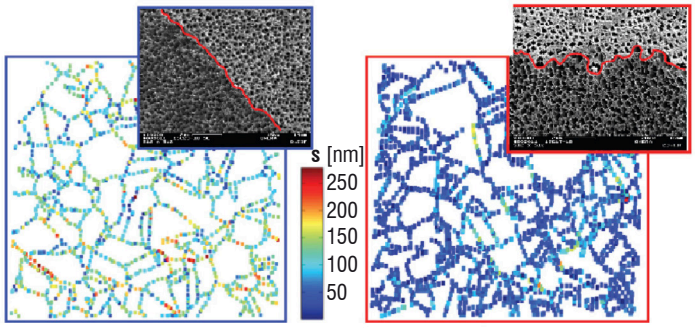


Figure 15 – Comparison between measured amplitudes of GB shearing displacements in superalloy microstructures with the two different GB morphologies shown in the insets: (a) smooth and (b) tortuous

samples by micro-electronic lithography, the data acquisition for the EBSD and SEM, or the development of the required software and post-processing tools.

Combined DIC with *in-situ* SEM and optical techniques to study multi-scale deformation phenomena

An illustration of how microstructural observation and mechanical characterization techniques can be combined inside the chamber of a SEM to study the deformation behavior of materials at two or more different microstructural scales is provided next. The need to simultaneously observe the evolution of the material microstructure at more than one scale can be found in a broad range of materials. A representative example is that of Nickel-base superalloys, widely used in aerospace and power generation applications, where microstructural heterogeneities can be found at different microstructural scales. As can be seen in Figure 16(a), mesoscale heterogeneities are introduced by the presence of 10-20 μm casting pores, usually found in the interdendritic regions, while those at the microscale (Figure 16(b)) result from the 0.5-1 μm γ' precipitates embedded in a nickel solid solution (γ) matrix. Typical micro-cracks which emanate from the casting pore, known to be due to localized deformation in the pore's vicinity (e.g., see [6]), can also be seen in Figure 16(a).

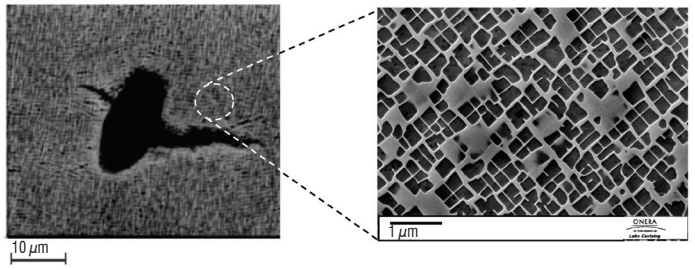
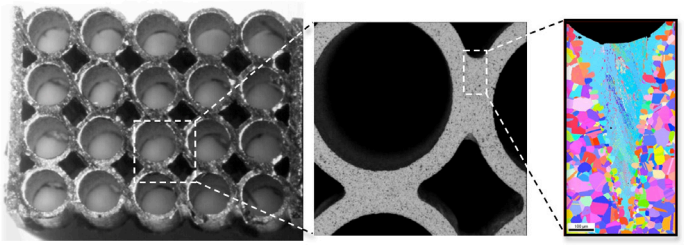


Figure 16 – Typical length scales associated with (a) a cracked casting defect [6] and (b) the superalloy $\gamma - \gamma'$ microstructure

Another class of materials that exhibits distinct observation scales are architectured materials, where not only the intrinsic scales of the microstructure (e.g., the grain size) are relevant, but also that of its architecture arising from processing and/or manufacturing.



(a) (b)

Figure 17 – Example of a multiscale architectured material consisting of brazed Inconel tubes: (a) optical image at the scale of a group of tubes, and SEM images of (b) the local welded tubes and (c) the weld microstructure

A typical case is that of the multi-scale architectured material shown in Figure 17, consisting of brazed Inconel tubes. Here, Figure 17(a) is an optical image at the scale of a group of tubes, and (b) and (c) are SEM images of the local brazed tubes and of the weld microstructure, respectively.

At ONERA, microscopists, metallurgists and engineers have recently developed a novel experimental *in situ* SEM setup to enable the simultaneous microstructural observation and measurement of the material deformation behavior at two different scales: a macroscopic one at the scale of the *in situ* specimen gauge length, and at the micrometer and nanometer scales associated with the actual material microstructure. By combining an *in situ* mechanical testing device, an optical microscope with a long lens and an SEM, it is possible to measure surface deformation fields over several millimeters and quantify the deformation at the scale of either individual grains in metals or fibers in composites. This original experimental set up, which has now been developed commercially and which has been named MEBIRIS, is illustrated in Figure 18. Here, the optical microscope

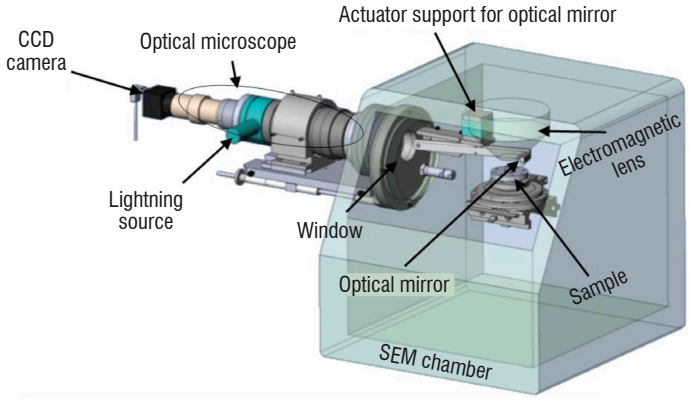


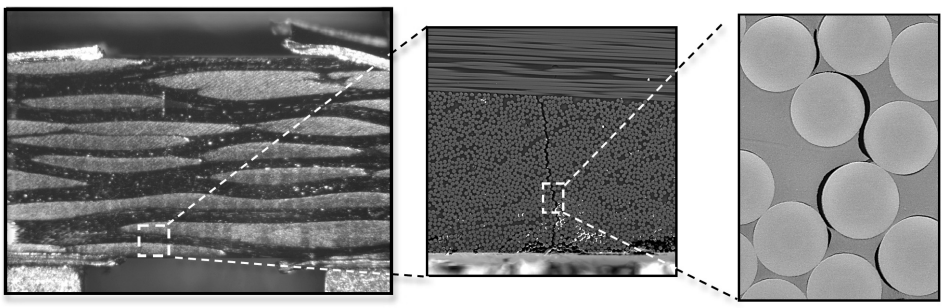
Figure 18 – Schematic set-up of the *in situ* SEM prototype named MEBIRIS

is externally mounted in one of the SEM ports. The resulting optical images are obtained with the help of a mirror positioned directly above the electronic beam column and inclined at an angle of 45° towards the microscope. The electronic image requires that the mirror be retractable. This was made possible by attaching it to an articulated arm driven by an externally controlled micro-motor.

For the development of MEBIRIS, it was necessary to respect a certain number of constraints, such as the optical performance defined by a variable field size (ranging from mm^2 to cm^2), a focal plane capable of being positioned within a range of up to ~20 mm to follow a moving specimen (e.g., such as during a bending test), and the capability to obtain the acquired images with an error smaller than a pixel after the mirror retraction and insertion. There is also the need to precisely align the images obtained by optical means and by the SEM and for their centers to coincide. Additional challenges are associated with the fact that the SEM chamber works in a vacuum using a testing device mounted within the SEM chamber itself, and that an integrated lighting system is required for the optical microscope. Furthermore, no perturbations or noise can be tolerated by the SEM imaging.

An example of the capabilities of the prototype MEBIRIS will be shown next on a material with a multi-scale microstructure, namely the woven C-fiber polymer matrix composite shown in Figure 19. The main objective of this study was to simultaneously investigate during an *in-situ* tension test the development of mesoscopic cracks in the strand, and of interfacial micro-cracks resulting from the decohesion of the individual fibers from the matrix. For that purpose, optical and electronic images were captured alternatively during the test. Figure 19(a) shows an optical image at the scale of a group of strands of carbon fibers, and (b) and (c) are SEM images of a crack inside an individual strand, and of the decohesion between the fibers and the matrix, respectively. In this particular example, the scales of the fibers (10 μm) and the strands (500 μm) are only accessible via the SEM, but not the collection of strands shown in Figure 19(a), since the region of interest here is too large to be observed by SEM and needs instead to be obtained using the optical microscope.

The optical image shown in Figure 19(a) covers the sample width in the gauge length region, 7.4 × 7.4 mm in size, with a 5 μm resolution per pixel. The acquisition time per image was of ≈ 200 ms. Figure 19(b)'s electronic image, on the other hand, enabled the observation of local damage in a 450 × 450 μm region of an individual fiber strand with a resolution of 0.2 μm per pixel. However, the acquisition time per image was 64 s. Here, the stress applied at the moment when the images were taken was 240 MPa.



(a) (b) (c)

Figure 19 – Multiscale microstructure of a woven C fiber polymer matrix composite: (a) optical image at the scale of a strand of carbon fibers, and SEM images of (b) a crack inside an individual strand, and (c) of the decohesion between the fibers and the matrix

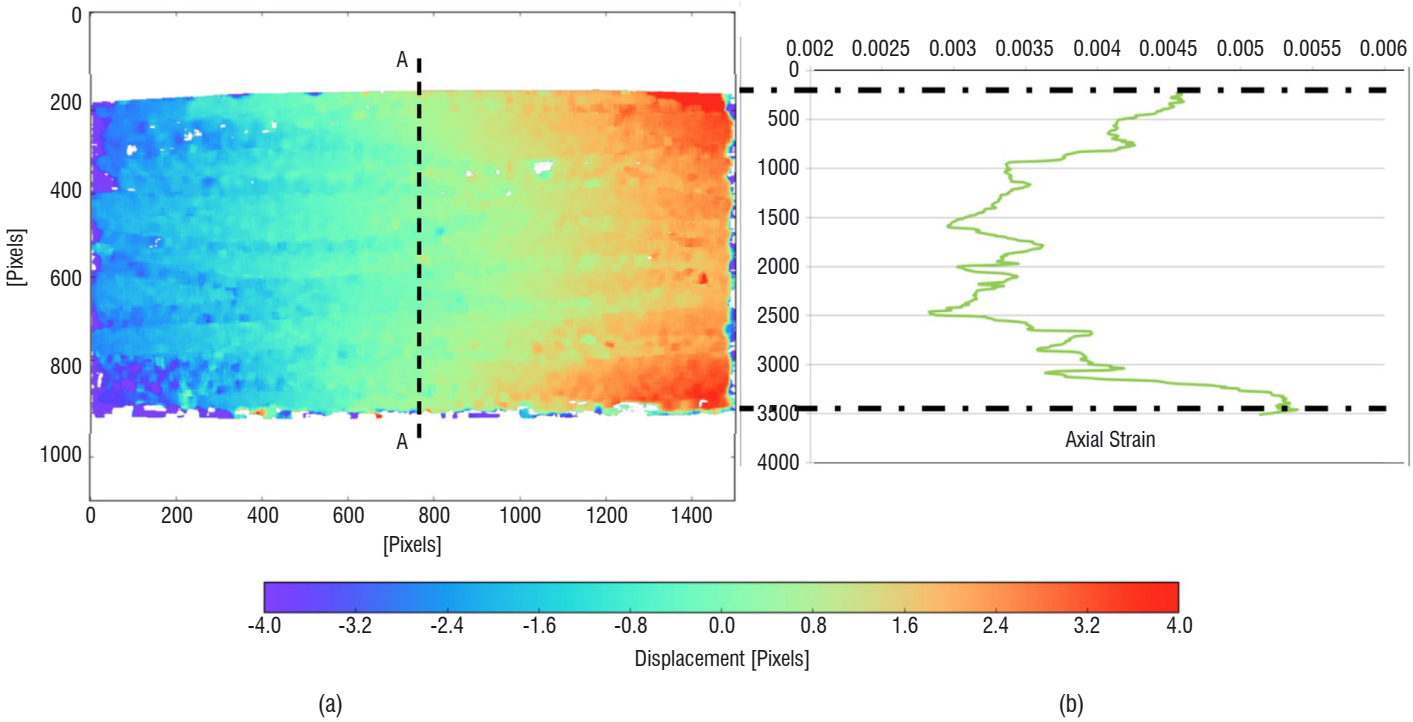


Figure 20 – (a) Measured axial displacements corresponding to Figure 19(a)'s optical image, and (b) axial strain profile along the line AA in (a) for an applied axial stress of 240 MPa

Figure 20(a) shows a color plot of the axial strains obtained by correlating the optical images such as that of Figure 19(a), between the undeformed and the deformed states, using the ONERA's FOLKI-D DIC method previously described. It can be seen that the quality of the optical images is sufficient for the image correlation. It is also possible to distinguish from the plot of the axial strain distribution the different deformation behavior exhibited by the axial and transverse strands. Figure 20(b) presents the measured axial strain profile across the composite specimen thickness, varying from 0.45 % at the surface, to about 0.30 % at its center. These values are consistent with those measured from macroscopic uniaxial tensile tests on the composite. From these type of measurements, a mean value of axial strain can be extracted for each of the loading steps to obtain an estimate of the average linear elastic stiffness of the composite. In this work, a value of $E = 77.3$ GPa was found, which compares well with the Young modulus of 72 ± 2 GPa measured macroscopically.

Concluding remarks

In this overview, we have discussed different techniques to describe local 2D and 3D material microstructures and to measure their deformation behavior. Several examples of *in-situ* or *ex-situ* morphological and structural imaging of heterogeneous materials have been given. A comparative discussion between the classical global and local DIC approaches with the pixel-wise resolution ONERA approach, FOLKI-D, revealed considerable improvements of the latter in both resolution, due to its one-displacement vector-per-pixel architecture, and computational time arising from its suitability for parallel computing implementation using GPUs.

A contrast was provided between EBSD and optical and electron microscopy techniques, highlighting their complementarity. It is anticipated that the combination of different characterization and measurement techniques will significantly accelerate in the near future, to address outstanding issues related to multi-scale deformation phenomena spanning a few hundred of nanometers (defect) to a few micrometers (grain). This will likely involve the combination of techniques such as SEM, TEM, EBSD, EDS, tomography, DIC, thermography and improved software to integrate the resulting combined data sets for model calibration and validation. It is also expected that these advances will lead to the use of *in-situ* techniques combining experimental and numerical methods, including imaging over a broader range of length scales, and real-time computations to study the behavior of novel materials with complex heterogeneous microstructures ■

Acknowledgements

The authors are very grateful to Jean-Didier Garaud and Guy Le Besnerais from ONERA, for the discussions on the numerical aspects of the DIC work, and to A. Soula and K. Thibault for their doctoral research, which provided the foundations for some of the results of the case studies discussed here.

Bibliography

[1] E. BAR-KOCHBA, J. TOYJANOVA, E. ANDREWS, K.-S. KIM, C. FRANCK - *A Fast Iterative Digital Volume Correlation Algorithm for Large Deformations*. Exp. Mechanics, V. 55(1), 2015, 261-274.

[2] R.L. BELL, C. GRAEME-BARBER - *Surface and Interior Measurements of Grain Boundary Sliding During Creep*. Journal of Materials Science, 5, 1970, 933-944.

[3] M. BORNERT, F. BREMAND, P. DOUMALIN, *et al.* - *Assessment of Digital Image Correlation Measurement Errors: Methodology and Results*. Exp. Mech 49, 2009, 353-370.

[4] H. BRUNNER, N.J. GRANT - *Deformation Resulting from Grain Boundary Sliding*. Transactions of the Metallurgical Society of AIME, 215, 1959, 48-56.

[5] J-Y BUFFIERE, E. MAIRE, J. ADRIEN, J-P. MASSE, and E. BOLLER, *In Situ Experiments with X Ray Tomography: an Attractive Tool for Experimental Mechanics*. Exp. Mech. V. 50, 2010, 289-305.

[6] S. DUMOULIN, E. P. BUSSO, N. P. O'DOWD - *A Multiscale Approach for Coupled Phenomena in FCC Materials at High Temperatures*. Philosophical Magazine, V. 83 (31-34), 2003, 3895-3916.

[7] F. CHAMPAGNAT, A. PLYER, G. LE BESNERAIS, B. LECLAIRE, S. DAVOUST, and Y. LE SANT - *Fast and Accurate PIV Computation Using Highly Parallel Iterative Correlation Maximization*. Exp. Fluids 50, 2011, 1169-1182.

[8] S. FELD-PAYET, V. CHIARUTTINI, J. BESSON and F. FEYEL - *A New Marching Ridge Algorithm for Crack Path Tracking in Regularized media*. Int. J; Solids and Struct., V. 71, 2015, 57-69.

[9] B. GRANT, H. STONE, P. WITHERS, and M. PREUSS - *High Temperature Strain Field Measurement Using Digital Image Correlation*. J. Strain. Anal. Eng. 44, 2009, 263-271.

[10] F. HILD and H. D. ESPINOSA - *Full Field Measurements and Identification in Solid Mechanics*. Procedia IUTAM , 4, 2012, 1-6.

[11] F. HILD and S. ROUX - *Comparison of Local and Global Approaches to Digital Image Correlation*. Experimental Mechanics V. 52, 2012, 1503-1519.

[12] J.T. HAMMER, T.J. LIUTKUS, J.D. SEIDT, A. GILAT - *Using Digital Image Correlation (DIC) in Dynamic Punch Tests*. Exp. Mechanics, V. 55(1), 2015, 201-210.

[13] F. LAGATTU, F. BRIDIER, P. VILLECHAISE and J. BRILLAUD - *In-Plane Measurements on a Microscopic Scale by Coupling Digital Image Correlation and In Situ SEM Technique*. Materials Characterization, V 56, Issue 1, 2006, 10-18.

[14] G. LE BESNERAIS and F. CHAMPAGNAT - *Dense Optical Flow by Iterative Local Window Registration*. proc. IEEE International Conference on Image Processing 2005, IEEE, Vol. 1, 1-137.

[15] G. LE BESNERAIS, Y. LE SANT and D. L'EVÊQUE - *Fast and Dense 2D and 3D Displacement Field Estimation by a Highly Parallel Image Correlation Algorithm*. Strain, V. 52, 2016, 286-306.

[16] B.D. LUCAS and T. KANADE - *An Iterative Image Registration Technique with an Application to Stereo Vision in*. IJCAI 81, 1981, 674-679.

[17] H. POMMIER, E.P. BUSSO, T.F. MORGENEYER and A. PINEAU - *Intergranular Damage During Stress Relaxation in AISI 316L-Type Austenitic Stainless Steels: Effect of Carbon, Nitrogen and Phosphorus Contents*. Acta Mat. 103, 2016, 893-908.

[18] J.D. PARKER, B. WILSHIRE - *A Surface Measurement Study of Grain Boundary Sliding During Creep of a Two Phase, Copper-Cobalt Alloy*. Materials Science and Engineering, 29, 1977, 219-225.

[19] J.-C. PASSIEUX, F. BUGARIN, C. DAVID, J.-N. PÉRIÉ, L. ROBERT - *Multiscale Displacement Field Measurement Using Digital Image Correlation: Application to the Identification of Elastic Properties*. Exp. Mechanics, V. 55(1), 2015, 121-137.

[20] R.C. POND, D.A. SMITH, P.W.J. SOUTHERDEN - *On the Role of Grain Boundary Dislocations in High Temperature Creep*. Philosophical Mag- azine, 37, 1978, 27-40.

[21] L. PRIESTER - *Les joints de grains. De la théorie à l'ingénierie.* EDP Sciences, Les Ulis, 2000.

[22] W.A. RACHINGER - *Relative Grain Translations in the Plastic Flow of Aluminium*. Journal of the Institute of Metals, 81, 1952, 33-41.

[23] P. L. REU, W. SWEATT, T. MILLER, D. FLEMING - *Camera System Resolution and its Influence on Digital Image Correlation*. Exp. Mechanics, V. 55(1), 2015, 9-25.

[24] C. REY, P. VIARIS DE LESEGNO, R. CHIRON - *Analysis of Shear Localization in Iron Crystals by Local Strain Field and Lattice Rotation Field Measurements*. Local Strain and Temperature Measurement, 1999, 30-39.

[25] A. SOULA - *Étude de la déformation intergranulaire au cours du fluage à haute température d'un superalliage à base de nickel polycristallin*. Doctoral Thesis, INPG, Grenoble, 2008.

[26] A. SOULA, Y. RENOLLET, D. BOIVIN, J.-L. POUCHOU, D. LOCQ, P. CARON, Y. BRECHET - *Analysis of High-Temperature Creep Deformation in a Poly-Crystalline Nickel-Base Superalloy*. Materials Science and Engineering A 06, 2009, 510:301-306.

[27] A. SOULA, D. LOCQ, D. BOIVIN, Y. RENOLLET, P. CARON, Y. BRECHET - *Quantitative Evaluation of High Temperature Deformation Mechanisms: A Specific Microgrid Extensometry Technique Coupled with EBSD Analysis*. Journal of Materials Science 10, 2010, 45.

[28] S.R. STOCK - *Recent Advances in X-Ray Microtomography Applied to Materials*. Int. Mater. Rev. 53(3), 2008, 129-181.

[29] M. A. SUTTON, F. HILD - *Recent Advances and Perspectives in Digital Image Correlation*. Exp. Mechanics, V. 55(1), 2015, 1-8.

[30] V. THAVEEPRINGSRIPORN and G.S. WAS - *The Role of CSL Boundaries in Creep of Ni-16Cr-9Fe at 360°C*. Metall. Trans. 28A, 1997, 2101.

[31] K. THIBAULT - *Influence de la microstructure sur le glissement intergranulaire lors du fluage d'un superalliage pour disques.* Doctoral Thesis, INPG, Grenoble, 2012.

[32] K. THIBAULT, DIDIER LOCQ, PIERRE CARON, DENIS BOIVIN, YVES RENOLLET, YVES BRECHET - *Influence of Microstructure on Local Intra- and Intergranular Deformations During Creep of a Nickel-Based Superalloy At 700°C.* Mat. Sci. and Eng. A 12, 2013, 14-21.

AUTHORS

Esteban Busso is currently the Scientific Director of ONERA's Materials and Structures Division. He was formerly Professor of Mechanics of Materials at the *Ecole des Mines* de Paris and director of the *Ecole*'s *Centre des Matériaux* and, from 1994 till 2005, Professor at Imperial College's Department of Mechanical Engineering in London, UK. He obtained his MSc and PhD degrees from the Massachusetts Institute of Technology (MIT) in Cambridge, USA, in 1987 and 1990, respectively. He also worked in industry in the UK, Japan, South Africa and Argentina. His research involves micromechanics studies of deformation and fracture of materials and interfaces, with an emphasis on the development of multiscale and multiphysics concepts in mechanistic models to predict deformation and fracture processes.

Denis Boivin is a senior research scientist with 30 years of experience in scanning electron microscopy with associated microanalysis and crystallographic EBSD techniques. He obtained his degree in instrumental physics from the *Conservatoire National des Arts et métiers* in Paris, France, in 1991. Taking advantage of the acquired experience along side high level specialists at ONERA, he now leads the Scanning Electron Microscopy laboratory, where improvements and control of the existing characterization techniques are continuously maintained to the benefit of materials research.

David Lévêque is a senior research scientist with 20 years of experience in mechanics of materials for aerospace applications. He obtained his PhD degree from the *Ecole Normale Supérieure* in Cachan, France, in 1998. Since his hiring at ONERA, he has been involved in several multi-partner national and international projects on composite and metallic materials (French-Japanese supersonic cooperation, CleanSky, ATTLASII, ITP2, etc.). He now leads the Mechanical Characterization Laboratory, which is dedicated to the metallic material characterization and the development of various high temperature instrumented tests to better understand and identify the mechanical behavior of high temperature materials.

A. Hurmane, A. Mavel,
P. Paulmier, F. Laurin
(ONERA)

E-mail: Antoine.Hurmane@onera.fr

DOI: 10.12762/2016.AL12-03

Combined Experimental and Modeling Approaches for Strength Analysis of 3D Woven Composites: From Elementary Coupons to Complex Aeronautical Structures

New generations of 3D woven composite materials have been recently developed to be used in aeronautics as an alternative to the classical laminated composite materials, for structures exposed to impact. Therefore, it has been necessary to determine precisely the damage and failure scenarios for such materials subjected to different kinds of loadings through a large experimental testing campaign performed at ONERA on unnotched coupons. These tests have been multi-instrumented to understand the different damage and failure mechanisms encountered in 3D woven composite materials. Based on these observations, a model, named ONERA Damage Model for Polymer Matrix Composites (ODM-PMC), has been developed specifically for such materials. This non-linear material approach takes into account the different observed sources of non-linearity (viscoelasticity of the matrix, in–plane matrix damage, inter-yarn debondings and fiber yarn failures) and has been validated through comparisons with available tests on unnotched specimens. Moreover, the predicted failure loads, obtained with the ODM-PMC model, on plates containing different kinds of geometrical singularities, such as a hole or a milled groove, have been compared successfully to multi-instrumented test results also performed at ONERA. Finally, the ODM-PMC model has been applied to large 3D woven composite structures, quite representative of real industrial components. The predicted damage and failure scenarios seem to be relevant as compared to data available in the literature. Moreover, the obtained computational times are compatible with usage in an industrial environment. Therefore, it has been demonstrated that this approach, implemented in a commercial finite element code, could be used in design offices in aeronautical industries.

Introduction

Some critical components of aircraft structures, such as center wing box, wings or fuselage, are nowadays manufactured from laminated composite materials, owing to their high specific mechanical properties. However, these materials are particularly sensitive to low-velocity/energy impact events, such as dropped tools, which induce a strong decrease in the residual strengths after impact. Therefore, due to the poor impact resistance of classical laminated composite materials, 3D woven composites have been recently developed to be used in industrial applications [1] exposed to impact. It has been demonstrated experimentally [2-6] for different kinds of loadings that the damage and failure mechanisms are strongly linked to the local architecture of these composite materials and are very different from those observed in classical laminated composites. Due to very attractive mechanical properties of the new generations of 3D woven composites, such materials are taken into consideration in order to manufacture some large components in aeronautics, possibly exposed to impact. These composite structures present a complex geometry, including some geometrical singularities which induce locally severe stress gradients. Moreover, these composite structures could be subjected to a wide range of multiaxial loadings, such as combined bending/tensile/compressive loadings.

The massive use of finite element simulations is an absolute necessity in order to efficiently design innovative 3D woven composite structures. This article deals with the development of a physically-based damage and failure approach developed specifically for 3D woven composite materials. Therefore, this material approach allows an accurate description of the different damage and failure mechanisms observed in elementary coupons subjected to different loadings; moreover, the resulting model has to be transferable and easy to use in design offices to predict the strength of large composite components, representative of industrial applications.

In order to improve the understanding of the damage and failure mechanisms encountered in new generations of 3D woven composites [7-9], a large testing campaign has been performed at ONERA. Simple coupons have been subjected to off-axis tensile, compressive and bending loadings, complementary to those already presented elsewhere [2-6]. These tests have been multi-instrumented, in order to establish the damage and failure scenarios for the different loading conditions under consideration: tension, compression, and bending. The ONERA Damage Model for Polymer Matrix Composite (ODM-PMC), specifically developed for 3D woven composite structures under static loadings, is briefly presented. Only the main ideas of the model are reminded. Some comparisons with the available experimental data on unnotched specimens are presented and the predictive capabilities of the model are discussed. Then, one of the objectives of this study is to evaluate the predictive capabilities of this material model on academic composite structures, containing geometrical singularities inducing stress gradients, such as holes or notches. Then, the multi-instrumented tests performed at ONERA on open-hole plates or notched specimens subjected to compressive loading are presented and the results are compared with the strength predictions of the ODM-PMC model, implemented in a commercial finite element code. Finally, this model is applied to a potential industrial component, to demonstrate the capabilities of this approach to be used in a design office in aeronautical industries.

Box 1 - Studied material

The material under investigation is a highly unbalanced 3D woven composite material consisting of carbon fiber yarns (48K) embedded in an epoxy matrix. Based on the generic architecture reported in Figure 1 [10], the studied 3D woven composite material has been optimized in order to prevent large delamination after impact and thus to obtain a good impact resistance (the exact architecture of this material is, however, confidential as requested by the manufacturer). The thickness of the tested material is about 9.5 mm, which is rather thick compared to classically studied laminates or other 3D woven composites. It should be noted that the Representative Elementary Volume of such a material is rather large (a few centimeters) as compared to other composite materials and thus prohibits the use of existing testing standards, and necessitates alternative designs of the testing samples even for the elementary coupons.

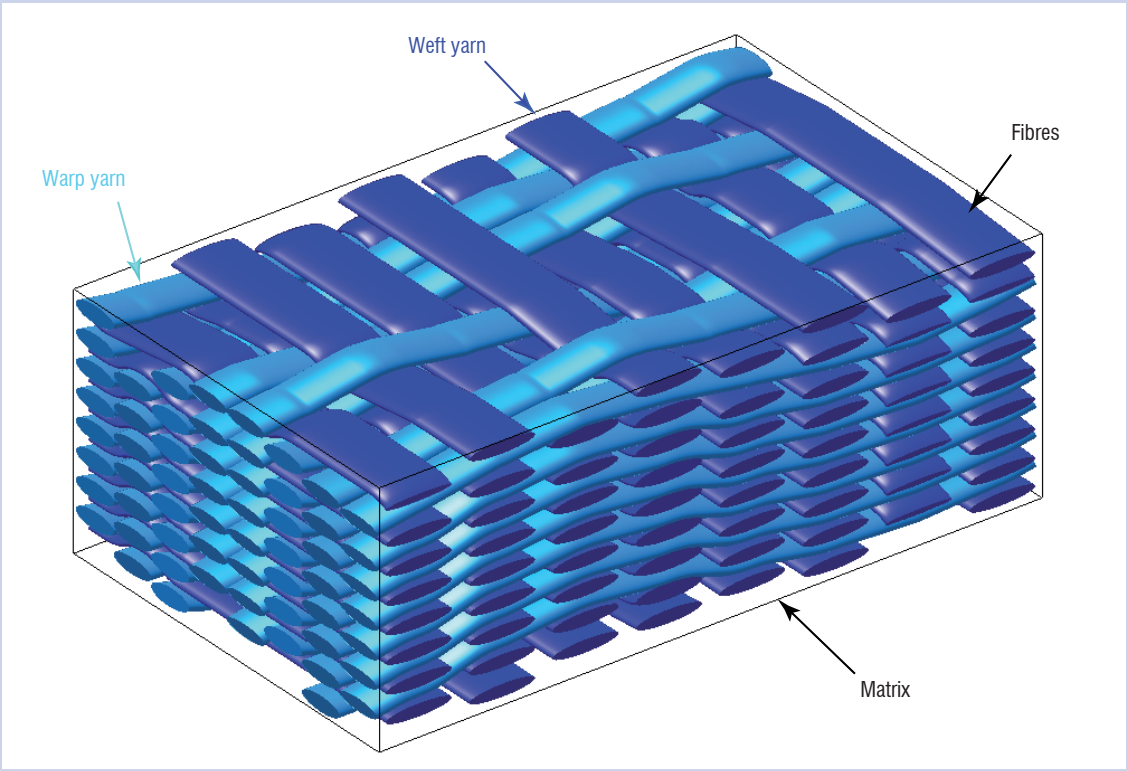


Figure B1-01 – Generic architecture of an unbalanced 3D woven composite material.

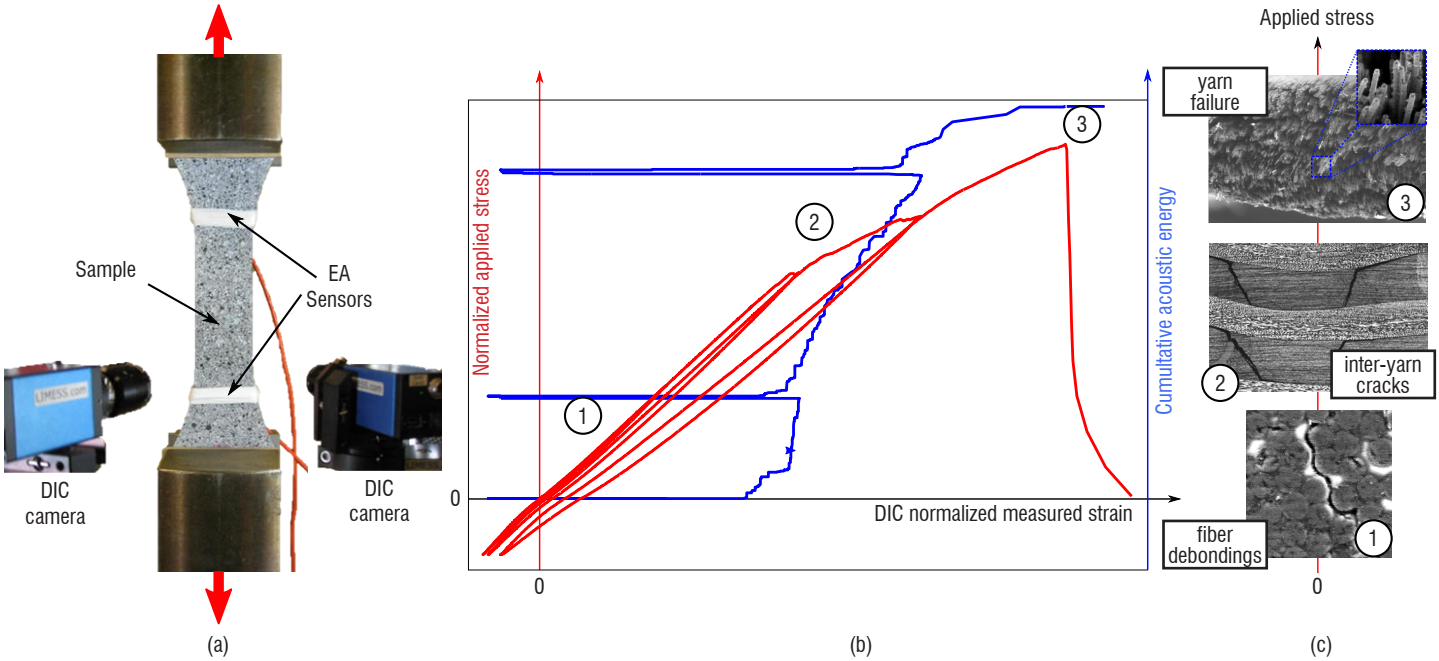


Figure 1 – Damage and failure scenario in tension in the weft direction. (a) Experimental set-up and associated multi-instrumentation, (b) macroscopic behavior, evolution of the applied stress with respect to strain and cumulative energy/strain evolution, (c) micrographs of the different damage and failure mechanisms.

Experimental study of the behavior of a 3D woven composite

Experimental testing campaign

In order to study the behavior of a new generation of 3D woven composite materials (see the box entitled studied material), an experimental testing campaign has been performed at ONERA. An electro-mechanical Schenck machine (150 kN maximum capacity) is used. The material is tested under tensile (Figure 1a) and compressive (Figure 2a) loadings. Because of the specific unbalanced architecture of the material, the behavior is investigated in different material axes (*i.e.*, tests are performed at 0°, corresponding to the warp direction, 45° and 90°, corresponding to the weft direction). Each test is repeated in order to estimate the potential result scattering (which remains rather low for this material). Tension tests are load-controlled, whereas compression tests are displacement-controlled. This testing campaign is aimed at understanding non-linear behavior sources and at observing the damage and rupture mechanisms under different loadings applied in different directions. Hence, multi-instrumentation is used for each mechanical test. Digital Images Correlation (DIC) is performed with the commercial code Vic3D© to measure displacement fields on the specimen surface, to estimate the strain fields [7,8,11] evolution during the tests, or to detect surface cracking [12,13]. Optical microscopy [11,14] enables the observation of the different kinds of damage on one edge of the sample, establishing the damage scenarios precisely. Acoustic Emission (AE) records acoustic events within the sample (volume information) and provides useful information on the evolution of the damage [10,15]. Moreover, X-Ray micro-tomography is performed on some samples after interrupted tests, in order to complete and validate, through the volume of the material, the scenarios established using surface microscope observations. Thanks to the information collected during these tests, it is possible to (i) cross-validate the different measurements [13] and improve the confidence in the measurements, and (ii) to establish the damage and failure scenarios for different loadings, specific for this material. Damage and rupture scenarios have been established in tension and compression and are rather different. These scenarios are similar in the warp and weft directions and detailed in the following sections.

Damage and failure scenario in tension

Under tensile loading, the non-linear response of the material can be decomposed into three main phases, as reported in Figure 1b: (i) a viscoelastic non-linear behavior, (ii) the onset and evolution of different matrix damage mechanisms (mostly in-plane transverse cracking and inter-yarn debondings), and (iii), finally, the rupture of the fiber yarns, which induces the failure of the elementary tested specimens. For low stress levels, the observed non-linearity is rather moderate and is mainly due to the viscous behavior [16] of the composite material (due to the polymer matrix). This point has been enhanced through creep tests, especially for off-axis tension at 45°. The second phase of the macroscopic behavior is clearly non-linear, as reported in Figure 1b for a tensile test in the weft direction, and is due to mesoscopic damage. This damage mostly consists in in-plane inter-yarn cracks, as illustrated in Figure 1c. The mesoscopic damage is clearly governed by the architecture of the material and through the analysis of X-Ray tomography only diffuse damage in the material is observed. It can be noted that the mesoscopic damage observed in classical 2D woven composites is very different, and is located mostly within the fiber yarns [17-19]. Moreover, at inter-yarn crack tips, inter-yarn debondings are generated. Again, no large delamination crack is observed, the different matrix cracks being confined in the microstructure of the material. The last phase of the scenario corresponds to the failure of the coupon when fiber yarns break. A Scanning Electron Microscope (SEM) examination of a yarn after failure (Figure 1c) enables failure patterns (with failed fiber clusters [18,20]) to be observed, which are classically observed in 2D woven and laminated composites with unidirectional plies.

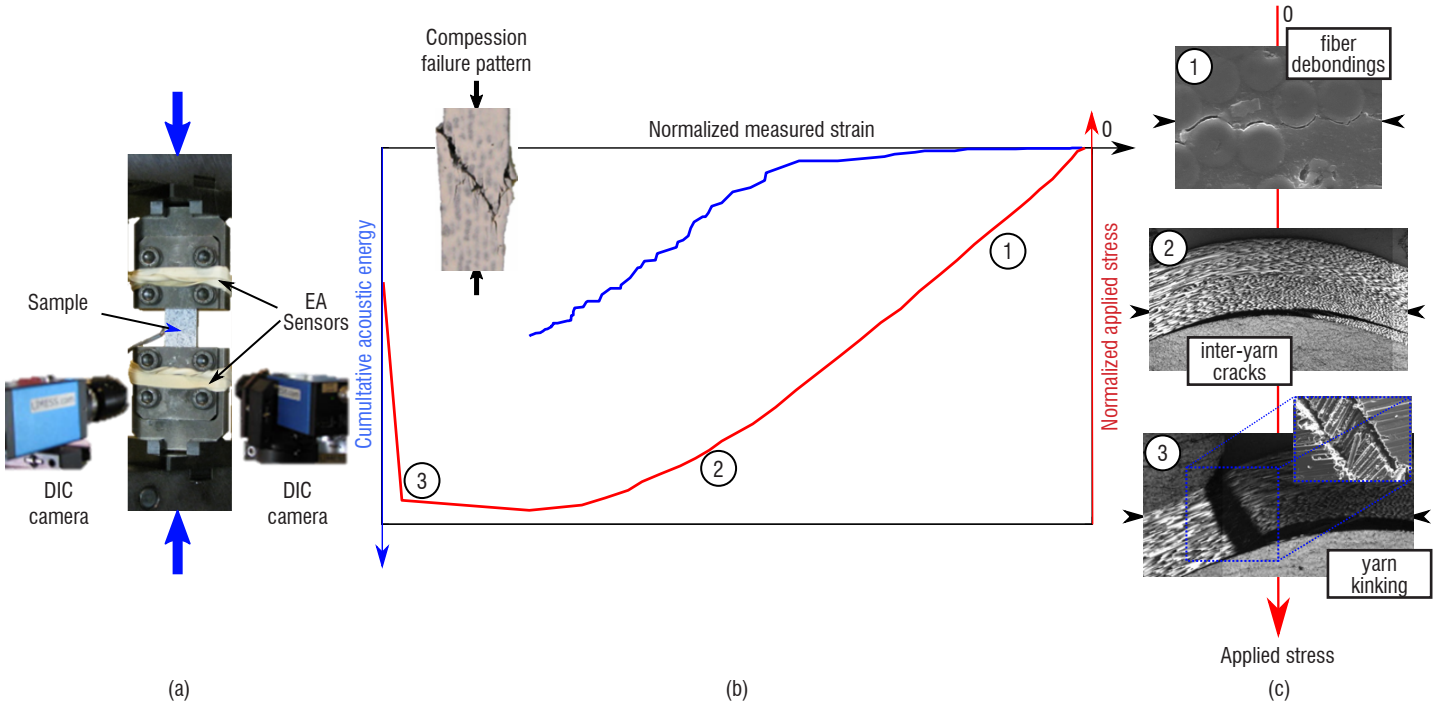


Figure 2 – Damage scenario in compression in the weft direction. (a) Experimental set-up and associated multi-instrumentation, (b) global behavior, evolution of the applied stress with respect to strain and cumulative energy/strain evolution and (c) micrographs of the different damage and failure mechanisms.

Damage and failure scenario in compression

Under compressive loading, the macroscopic behavior can also be decomposed into three main phases, as reported in Figure 2b: (i) a viscoelastic behavior, (ii) a slightly non-linear part, mainly due to inter-yarn debondings, and (iii), finally, the rupture of the fiber yarns leading to failure of the tested elementary specimen. The first phase is again due to the matrix viscosity, which has been further evidenced through the analysis of 45°-off-axis compressive creep tests. The non-linear behaviors in tension and in compression are different [21,22]. Indeed, in tensile creep tests, fiber/matrix debondings within the fiber yarns (or matrix micro-cracks), normal to the applied loading (see Figure 1c), are observed, thus increasing the apparent non-linearity. In compression creep tests, fiber/matrix debondings are generated through the Poisson effect and are thus parallel to the loading direction, as shown in Figure 2c. Therefore, this micro-damage has a negligible influence on the viscosity in the loading direction. The second phase, which is slightly non-linear, is due to inter-yarn debondings resulting from the compressive loading of the fiber yarns presenting the larger initial waviness, as illustrated in Figure 2c. Then, inter-yarn debonding promotes the failure of the fiber yarns due to kinking of the latter, also linked to the initial waviness of the yarns. Kinking of yarns always occurs in the vicinity of inter-yarn debondings, as observed in Figure 2c. It can be noted that the result scattering is rather low compared to that of classical unidirectional plies because the initial waviness of the fiber yarns is controlled during the manufacturing process.

The ONERA Damage Model for 3D woven Polymer Matrix Composites (ODM-PMC)

As mentioned previously, only diffuse damage (in-plane inter-yarn matrix cracking and inter-yarn debondings) is observed, contrary to damage studied in laminates constituted of unidirectional plies. The ONERA Damage Model for Composites with a Polymer Matrix (ODM-PMC) is based on a continuum damage approach [23] and is defined at the macroscopic scale to be used for the strength predictions of academic composite structures, but also for large composite structures representative of industrial components. For new generations of 3D woven composites, multiple sources of non-linearity have been determined experimentally through the analysis of quasi-static tests.

The ODM-PMC model has been proposed, in order to describe the non-linear behavior, damage and failure of 3D woven composites subjected to quasi-static loadings. Only the main ideas of this model are presented here; more details can be found in [8,9,11,24-26]. This approach is thermodynamically consistent and the macroscopic behavior, expressed in Eq. 1, derives directly from the Helmholtz free energy.

$$\underline{\sigma} = \underline{\underline{C}}^{eff} : (\underline{\varepsilon} - \underline{\varepsilon}^{ve} - \underline{\varepsilon}^{th} - \underline{\varepsilon}^{0}) - \underline{\underline{C}}^{0} : (\underline{\varepsilon}^{s} + \underline{\varepsilon}^{r} - \underline{\varepsilon}^{0}) \qquad (1)$$

where $\underline{\sigma}$ is the stress tensor, $\underline{\varepsilon}$ is the total strain tensor, and $\underline{\varepsilon}^{ve}$ is the viscoelastic strain tensor. Taking into account the viscosity of the polymer matrix is essential to accurately describe the macroscopic behavior of specimens subjected to off-axis loadings, as presented previously. The viscoelastic approach, taking into account the influence of micro-damage (noted δ_1, δ_2, δ_3) on the viscosity, used in this study is thus not detailed here, but more information can be found in [27,28].

In Eq.1, $\underline{\underline{C}}^{0}$ corresponds to the initial elastic stiffness tensor ($\underline{\underline{S}}^{0}$ is the initial elastic compliance) and $\underline{\underline{C}}^{eff}$ is the effective elastic stiffness tensor taking into account the effects of the different damage and failure mechanisms, as expressed in Eq. 2.

$$\underline{\underline{C}}^{eff} = \left(\underline{\underline{S}}^{0} + \sum_{i} d_i \, \underline{\underline{H}}_{i}^{m} + \sum_{j} \left(D_j^t \, \underline{\underline{H}}_{j}^{ft} + D_j^c \, \underline{\underline{H}}_{j}^{fc} \right) + D_3 \, \underline{\underline{H}}_{3}^{f} \right)^{-1} \quad (2)$$

with $(i, j) = \{1, 2\}$

For such a material, the high contrast between the mechanical properties of the constituents (matrix and fiber yarns) leads to crack orientations induced by the microstructure. Therefore, each damage or failure mechanism is described in the model by a scalar variable. In this approach, these different damage variables are classified by their effects on the macroscopic behavior. Indeed, two types of variables are considered: (i) mesoscopic damage variables induce a non-negligible non-linearity in the macroscopic behavior and (ii) macroscopic rupture variables lead to a violent decrease in the apparent rigidity.

Mesoscopic damage variables (d_1 in the warp direction, d_2 in the weft direction) are related to in-plane matrix cracking and induce a notable non-linear effect on the macroscopic behavior through the term ($\sum_{i=1}^{2} d_i \, \underline{\underline{H}}_{i}^{m}$) in Eq. 2.

The ($\underline{\varepsilon}^{0}$, $\underline{\varepsilon}^{s}$) strain tensors allow the unilateral aspect of damage to be taken into account in this model, while ensuring the continuity of the macroscopic behavior for any kind of complex non-linear loading. The unilateral aspect of damage means that the model takes into account the fact that the cracks, opened under tensile loading, will close under compressive loading. Therefore, the effects of damage mechanisms in a material subjected to tension are very different from those under compressive loading, due to crack closure. This point is essential to perform robust finite element simulations taking into account the unilateral aspect of damage. Moreover, the residual strain ($\underline{\varepsilon}^{r}$) describes the remaining strain after unloading at null stress, which is assumed to be strongly connected to the in-plane mesoscopic damage generated during the loading phase, as proposed by [29,30]. It has been demonstrated [31] that the introduction of this residual strain is essential to predict the permanent indentation after impact.

The out-of-plane macroscopic damage (D_3), which corresponds to the inter-yarn debondings, can play a major role in the macroscopic behavior, for instance during an impact test or a bearing test, and its effects are taken into account through the term ($D_3 \, \underline{\underline{H}}_{3}^{f}$) in Eq. 2.

As explained, for unnotched specimens the rupture in the material axis is due to the failure of fiber yarns. The failure mechanisms are different in tension and in compression, and are thus distinguished in the proposed modeling. The macroscopic variables (D_1^t, D_2^t) describe the effects of yarn failures in tension, in the warp and weft directions respectively, and the variables (D_1^c, D_2^c) describe the effects of yarn failures in compression. These yarn failures induce a violent and softening macroscopic behavior through the term $\sum_{j=1}^{2} \left(D_j^t \, \underline{\underline{H}}_{j}^{ft} + D_j^c \, \underline{\underline{H}}_{j}^{fc} \right)$ in Eq. 2. It can be noted that the influence of the hydrostatic pressure on the apparent strength of fiber yarns in compression is taken into account based on previous studies performed on laminated composites [32]. The description of the softening behavior due to yarn failures is necessary, in order to accurately predict the final failure of composite structures containing geometrical singularities, such as open-hole plates or notched specimens.

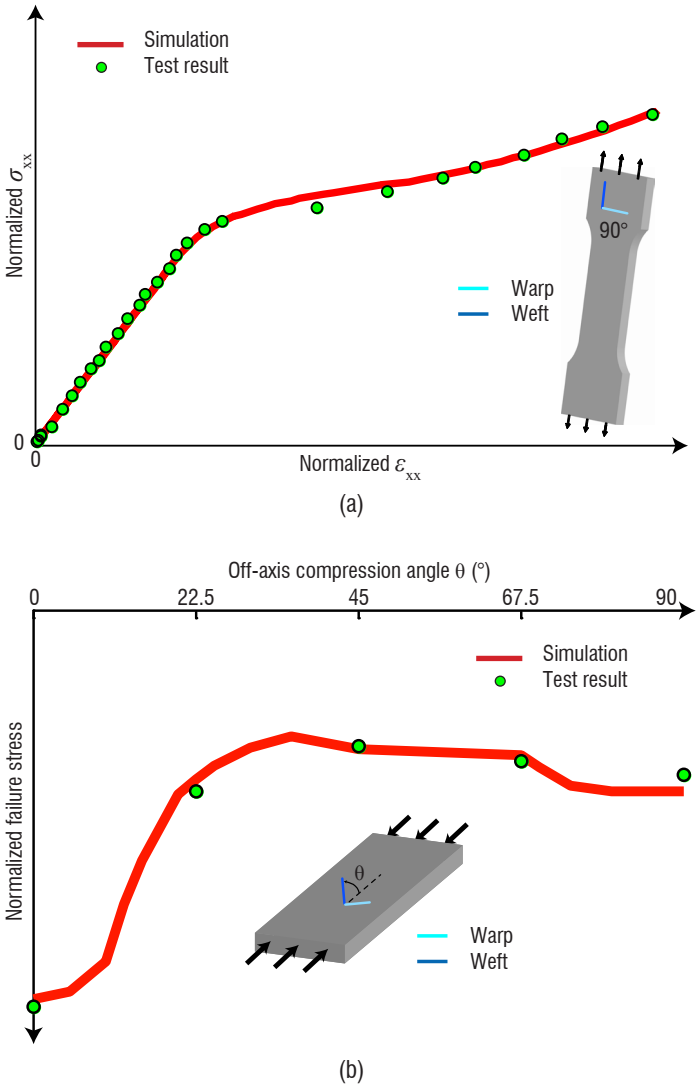


Figure 3 – a) Comparison between the experimental strain/stress curves and the prediction obtained with the ODM-PMC model for a tensile test in the weft direction and b) comparison between the measured and predicted stresses at failure for different off-axis compressive tests on unnotched samples.

The identification process for viscoelasticity, mesoscopic damage, residual strains, and the onset of inter-yarn debondings and yarn failure is well established through the analysis of unnotched coupons subjected to off-axis tensile and compressive loadings. The comparison between the measured and predicted macroscopic behavior for tensile tests in the weft direction is reported in Figure 3a, and the proposed model is able to accurately describe the various phases of the macroscopic behavior. Moreover, in Figure 3b, the predicted macroscopic stress at failure for various off-axis compressive tests are compared successfully with the available experimental results on unnotched samples tested at ONERA. Nevertheless, the parameters linked to the evolution law of the fiber yarns cannot be determined on unnotched samples because the first yarn failure induces the final failure of the tested unnotched specimens. The degradation law associated with the fiber yarn failure can only be determined through the analysis of tests performed on structures containing geometrical singularities, as presented in the following section.

Strength prediction of composite structures with geometrical singularities

The definition of the final rupture of high stress gradient parts of composite components, especially for open-hole plates, cannot be reduced to the first fiber yarn failure, due to the high stress gradient around the hole. Such a definition of final failure will lead to underestimating the failure strength by a factor of 2 or 3. It is therefore an absolute necessity to describe the progressive degradation of the mechanical properties due to fiber yarn failure, especially for compressive loading. In order to evaluate the predictive capacities of the ONERA Damage Model for Polymer Matrix Composites (ODM-PMC), a large experimental campaign on academic composite structures with different geometrical singularities subjected to compressive loading is performed. Two different kinds of geometrical singularities are thus considered: (i) classical circular open-hole plates with different diameters, and (ii) plates containing a milled groove or milled double notches, which generate more severe local stress gradients.

Strength prediction of open-hole plates subjected to compressive loading

For the open-hole plates subjected to compressive loading, four configurations were tested with different hole diameters, from 6 mm to 40 mm with a constant w/d ratio (plate width to hole diameter) equal to 6, as illustrated in Figure 4a. These tests, performed in the warp and weft directions, are instrumented with 2 CCD cameras for Digital Images Correlation (DIC) and Acoustic Emission (AE) sensors, in order to monitor the evolution of the various damage and failure mechanisms. A hydraulic Maser machine with a 500 kN maximum capacity was used, as reported in Figure 4b. Two different load cells have been used as a function of the hole diameter and the associated width. The tests were performed in the machine-controlled displacement mode and a constant (0.1 mm/min) displacement rate is imposed. The dimensions of the samples were chosen through a preliminary FE analysis, in order to avoid premature buckling of the specimens before kinking of the fiber yarns. Each test was repeated three times, in order to estimate scattering. The failure patterns, shown in Figure 4c, are rather similar for the various configurations tested in the warp and weft directions, and failure is due to kinking of the fiber yarns. The observed macroscopic crack is normal to the in-plane loading direction and presents a through-the-thickness angle, strongly linked to the architecture of the material and to the yarn kinking propagation through the thickness, as observed in smooth specimens (Figure 2b). The progressive propagation of the fiber yarn cracks through kinking is clearly evidenced through the analysis of the displacement field measured by DIC, as reported in Figure 4d.

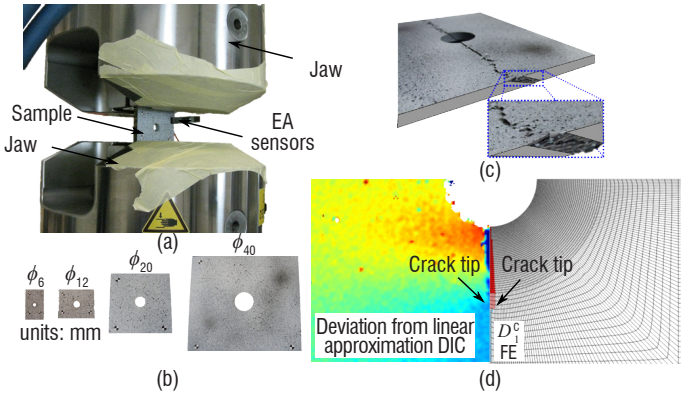


Figure 4 – a) Experimental set-up and associated multi-instrumentation to test open-hole plates subjected to uniaxial compressive loading, b) presentation of the tested open-hole plates with scale effects (ϕ_6 means $\phi = 6$ mm), c) observed failure pattern for open-hole plates in the warp and weft directions and d) comparison of the measured (with the deviation from the linear approximation of the displacement [12]) and predicted macroscopic crack lengths at the peak load for the open-hole plate with a 40 mm in diameter hole

In order to predict the strength of 3D woven composite structures, the ODM-PMC model has been implemented in the Abaqus/standard commercial finite element code. It can be noted that, because macroscopic failure variables (i.e. yarn failures and inter-yarn debondings) induce a softening behavior, it is necessary to introduce a regularization technique in order to avoid mesh dependence and localization of the solution. The delay effect method [33] is associated with yarn failure to avoid these problems, for the sake of simplicity.

To simulate these tests, Finite Element (FE) calculation conditions (geometry and mesh) have been established. The mesh size in the vicinity of the geometric singularity is the same for all of the tested configurations, because of the chosen regularization technique. The configuration, with a 6 mm in diameter hole, is used to identify the evolution law of fiber yarns under compression in the warp and weft directions. For the other configurations, the predicted failure loads are obtained with the same set of material parameters. In order to evaluate the benefit of using such a complex material approach, as compared to the classical semi-empirical models widely used in design offices in aeronautics, the Point Stress Method [34,35] is also applied to open-hole plates subjected to compression as a post-treatment of a linear finite element simulation and is described in the box entitled Presentation of the Point Stress Method. The characteristic distance d_0 is also identified in the configuration with a hole diameter equal to 6 mm in the warp and weft directions.

Box 2 - Presentation of the Point Stress Method

The Point Stress Method [34,35] is a semi-empirical method widely used in design offices to predict the strength of open-hole plates because of its simplicity. The main idea, illustrated in Figure B2-01, consists in evaluating, as a post-treatment of a linear elastic Finite Element simulation, a failure criterion at a given distance d_0 from the geometrical singularity, such as a hole, which induces the stress gradient. The specimen is considered as broken when this criterion is fulfilled for the first time at one point within the structure.

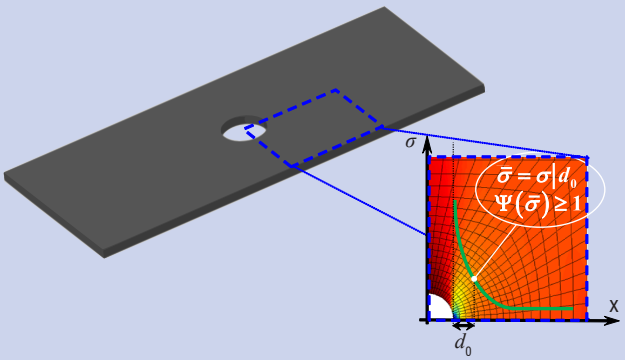
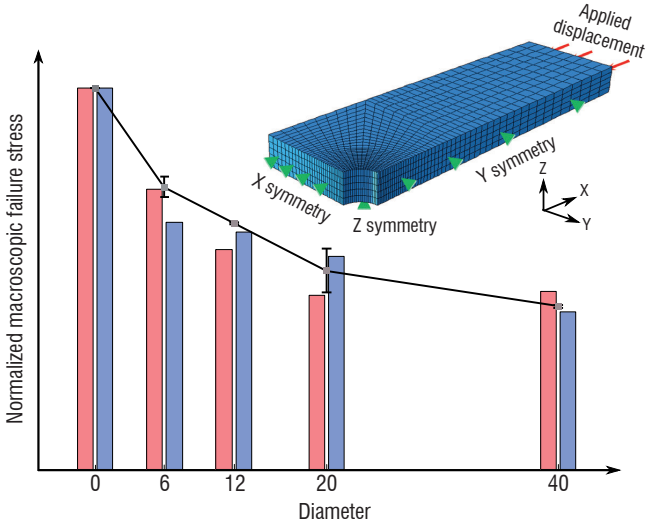


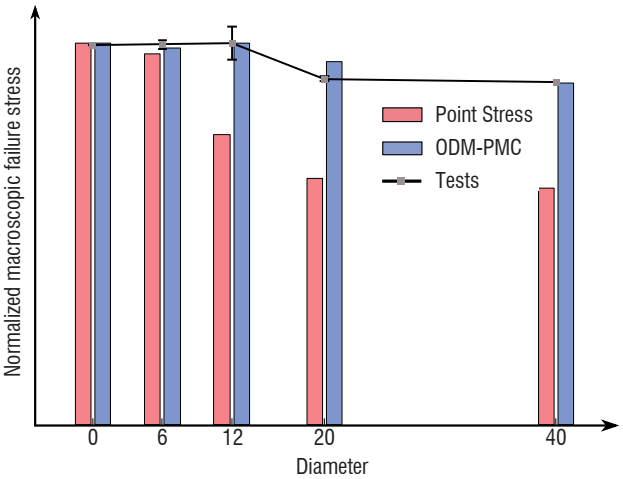
Figure B2-01 – Principle of the Point Stress Method applied to an open-hole plate.

In the present approach, owing to the introduction of the progressive degradation of the mechanical properties due to the fiber yarn rupture, the predicted failure macroscopic stresses at failure, defined as the failure load (peak force) divided by the real section of the tested specimen, are in good agreement with the experimental measurements for open-hole plates with different diameters subjected to compression in the warp direction, as illustrated in Figure 5a. Moreover, the predicted length of the fiber yarn cracks at the experimental peak load is compared successfully with that measured by DIC, as reported in Figure 4d. The predictions obtained with the Point Stress Method, reported in Figure 5a, are also in good agreement with experimental data. These results are not surprising because the macroscopic behavior is almost linear up to failure due to the high yarn ratio in this material direction.

Figure 5b shows the comparison between the experimental macroscopic stresses at failure in the weft direction and those predicted with the ODM-PMC model and the Point Stress Method. Failure strengths have been normalized to the failure strength of the un-notched specimen, corresponding to the 0 mm hole diameter. Firstly, the scale effect on the macroscopic stress at failure is not the same in the warp and weft directions. While the evolution with respect to the hole diameter in the warp direction is similar to that observed in laminated composites with unidirectional plies [36], the influence of the hole diameter is markedly different in the weft direction. Indeed, for a hole diameter smaller than 12 mm its influence on the stress at failure is negligible, but for larger hole diameters the apparent strength decreases. In the weft direction, the Point Stress Method predicts an evolution of the macroscopic stress at failure similar to that obtained in the warp direction, and is thus not in good agreement with available experimental data. Nevertheless, the ODM-PMC model enables the prediction of the scale effect on an open-hole plate, both in the warp and weft directions, very accurately. The difference between the two different directions is due to the amount of mesoscopic damage (in-plane matrix cracking) and inter-yarn debondings, which is higher in the weft direction in the present case of a highly unbalanced 3D woven composite. Moreover, it also explains the fact that the Point Stress Method, used as a post-treatment for linear elastic FE simulations, leads to results that are in poor agreement with experimental data in the weft direction.



(a)

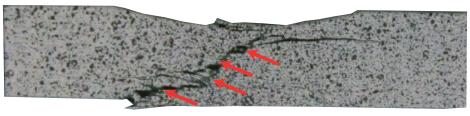


(b)

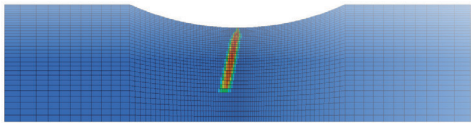
Figure 5 – Evolution of the normalized macroscopic stresses at failure measured and predicted with the ODM-PMC model and the Point Stress Method for open-hole plates subjected to compression in the a) warp and b) weft directions.

Strength prediction of machined plates subjected to compressive loading

In order to validate the predictive capabilities of the ODM-PMC model, additional machined composite structures have been subjected to uniaxial compressive loading. Three plates containing a milled groove have been machined, as illustrated in Figure 6. The radius of the groove is 90 mm and its depth in the plate is equal to 4 mm. Moreover, three additional plates containing milled double notches (*i.e.*, notches located on two opposite edges) have been machined and are illustrated in Figure 6b. The radius of the milled notches is 45 mm and their depths (along the width and through the thickness)

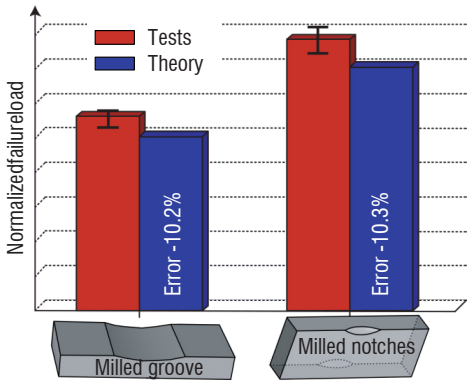


Observed failure pattern in a milled grooved specimen



Predicted failure pattern in a milled grooved specimen

(a)



(b)

Figure 6 – a) Observed and predicted failure patterns in a milled grooved 3D woven composite plate, b) comparison of the measured and predicted failure loads for different machined plates subjected to compression in the warp direction.

are equal to 5 mm. The geometry of these specimens has been controlled through stereo-digital image correlation.
These tests on machined specimens are performed only in the warp direction. The testing machine and the associated multi-instrumentation are similar to those already presented. The failure patterns for these two test configurations are very similar to those already observed for unnotched specimens and open-hole plates.

The finite element simulations are performed with the ODM-PMC model and the material parameters identified previously. Again, the size of the element of the mesh close to the geometrical singularities is similar to those used for open-hole plates, which is non-trivial for the complex geometries under consideration.

The damage and failure scenario for such complex structures is described correctly by the proposed approach. Indeed, inter-yarn debondings are first predicted close to the singularities, due to the compressive loading and, finally, fiber yarn kinking is predicted in the vicinity of the singularities and then propagates through the thickness with a given angle, which is rather consistent with the observation, as reported in Figure 6a. Moreover, Figure 6b presents the comparison between the predicted and measured failure loads for the two configurations of machined structures, inducing different local stress gradients. The predictions are conservative as compared to the experimental data and are in good agreement with experimental data (the error is around -10%).

To conclude, this non-linear material approach ODM-PMC enables the prediction of the final failure load of open-hole plates with scale effects, as well as the strengths of composite structures containing more complex singularities, such as a milled groove or milled double notches. Moreover, the damage and failure scenarios for all of the composite structures considered are accurately described by the proposed modeling. Finally, the benefit of using an advanced damage and failure approach as compared to semi-empirical approaches, such as the Point Stress Method, is demonstrated for some configurations inducing an important amount of matrix damage.

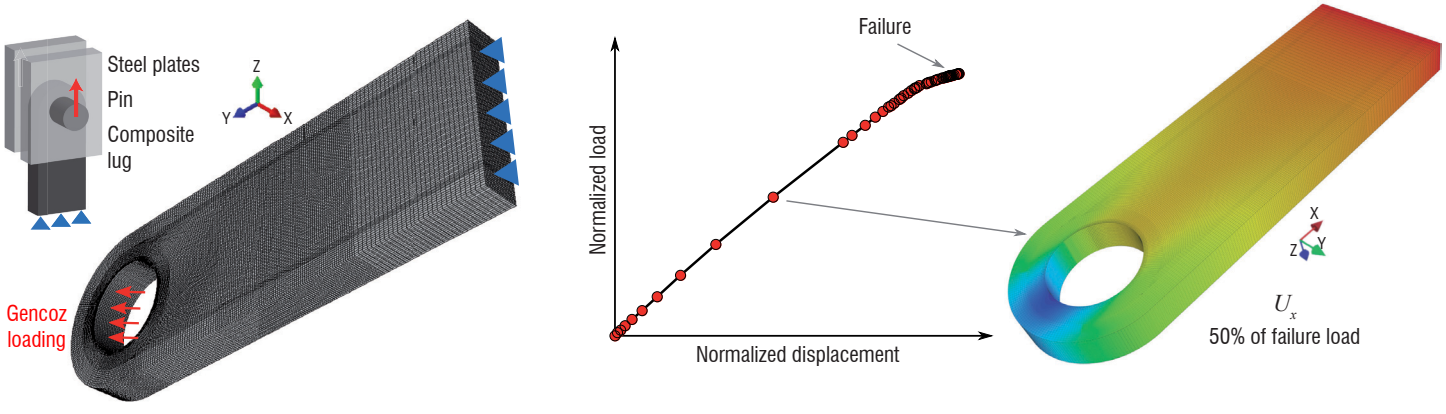
Application to a composite component

Since the ODM-PMC model has been validated on simple smooth specimens and composite structures containing geometrical singularities, such as holes, grooves or notches, this approach can be applied to 3D woven composite components, such as those tested in the higher levels of a validation pyramid.

The considered test case consists in a classical lug [37,38] subjected to tensile loading through an iron pin, as illustrated in Figure 7a. The geometry has been slightly simplified for the considered study. Based on [39], an empirical formula is used to describe the distribution of bearing pressure for pin and lug contact, in order to avoid managing contact and thus to reduce the computational time. The finite element mesh of such a structure contains 1.1 million degrees of freedom and necessitates 35 GB of RAM for non-linear calculation with the ODM-PMC model. It can be noted that special attention has been paid to the computation of the consistent tangent matrix, to ensure the convergence of such a complex material model. Moreover, the implementation in Abaqus/standard is consistent with multithread requirements, in order to reduce the time of computation. Therefore, the computational time for this test case is around seven hours with 30 CPUs, which could match the requirements of a design office in aeronautical industries, whereas the predicted global response, reported in Figure 7b, is markedly non-linear from 80% of the failure load. The predicted damage and failure scenario seems to be rather consistent with those reported in the literature for bolted joint problems applied to 3D woven composite materials [40,41], and with those observed in woven composite lugs tested in industry [38]. Further works should consist in performing, at ONERA, multi-instrumented tests on 3D woven composite lugs subjected to tensile or compressive loading, in order to validate this approach, especially through comparisons between the predicted displacement fields, as reported in Figure 7c, and those measured through DIC during the tests.

Conclusions / Perspectives

Due to the poor impact resistance of classical laminated composite materials, 3D woven composites have been recently developed to be potentially used in industrial applications exposed to impact. These composites having been recently developed, it has been necessary to determine precisely the damage and failure scenarios for such materials subjected to different kinds of loadings. Therefore, a large experimental testing campaign has been conducted at ONERA on smooth elementary coupons subjected to tensile, compressive or bending loadings. These tests have been multi-instrumented,



(a) (b) (c)

Figure 7 – Large structural test case: a) Boundary conditions and mesh of a lug subjected to tensile loading b) normalized predicted macroscopic displacement / load curve and c) predicted displacement field Ux at 50% of the failure load.

with Digital Image Correlation, acoustic emission, and X-Ray tomography, in order to understand the different damage and failure mechanisms encountered in 3D woven composite materials. It has been demonstrated that the matrix damage (in-plane matrix cracking and inter-yarn debondings) is diffuse within the material, confined to the local architecture, which constitutes a major difference compared with laminated composites.

Therefore, a model, named ONERA Damage Model for Polymer Matrix Composites (ODM-PMC) and based on the continuum damage theory, has been developed specifically for such a material. This non-linear material approach takes into account the different sources of non-linearity observed in smooth samples, such as the viscoelasticity of the matrix, in–plane matrix damage, inter-yarn debondings, and fiber yarn failures. The predictions of this approach, in terms of non-linear behavior, matrix damage evolution, and failure predictions, have been validated through comparisons with available tests on simple unnotched specimens.

Moreover, in order to evaluate the capability of the ODM-PMC model to predict the strength of 3D woven composite structures, it has been implemented in the Abaqus/standard commercial finite element code. Then, the model has been applied to plates containing different kinds of geometrical singularities, such as an open-hole, a milled groove, or milled double notches, subjected to compressive loading. These different singularities generate different local stress gradients, enabling the validation of the failure load predicted by the model over a wide range of structural configurations. Moreover, since these structural tests have been multi-instrumented, it has been demonstrated that the damage and failure scenario is also accurately described by the proposed approach.

Finally, the ODM-PMC model has been applied to a large 3D woven composite structure, quite representative of industrial components. The considered structural test consists in a lug subjected to bearing loading. The predicted damage and failure scenario is very promising compared to those reported in the literature for 3D woven composite materials. The corresponding tests should be performed soon at ONERA, in order to validate this approach. Moreover, given that special attention has been paid to the quality of the implementation in Abaqus/standard, the obtained computational times match the industrial requirements. Therefore, this approach, implemented in a commercial finite element code, could be transferred to design offices in aeronautical industries.

Further work would consist in estimating the residual strengths and predicting the fatigue lifetime of large industrial structures with a unified approach based on the ODM-PMC model, initially developed to predict the strength of composite structures subjected to quasi-static loadings. These developments are currently being performed in collaboration with other French research laboratories [9,10,42,43] ■

Acknowledgments

The authors would like to express their sincere gratitude to Dr. R. Valle, for valuable and helpful discussions.

References

[1] A.P. MOURITZ, M.K. BANNISTER, P.J. FALZON, K.H.LEONG - *Review of Applications for Advanced Three-Dimensional Fibre Textile Composites.* Composites Part A: Applied Science and Manufacturing 30, p. 1445-1461, 1999.

[2] P.J. CALLUS, A.P. MOURITZ, M.K. BANNISTER, K.H. LEONG - *Tensile Properties and Failure Mechanisms of 3D Woven GRP Composites.* Composites Part A: Applied Science and Manufacturing 30, 1277-1287, 1999.

[3] B.N. COX, M.S. DADKHAH, W.L. MORRIS - *On the Tensile Failure of 3D Woven Composites.* Composites Part A: Applied Science and Manufacturing 27, 447-458, 1996.

[4] B.N. COX, M.S. DADKHAH, R.V. INMAN, W.L. MORRIS, J. ZUPON - *Mechanisms of Compressive Failure in 3D Composites*. Acta Metallurgica et Materialia 40, 3285-3298, 1992.

[5] B.N. COX, M.S. DADKHAH, W.L. MORRIS, J.G. FLINTOFF - *Failure Mechanisms of 3D Woven Composites in Tension, Compression, and Bending.* Acta Metallurgica et Materialia 42, 3967-3984, 1994.

[6] G.A. BIBO, P.J. HOGG - *The Role of Reinforcement Architecture on Impact Damage Mechanisms and Post-Impact Compression Behaviour.* Journal of Materials Science 31, 1115-1137, 1996.

[7] J. SCHNEIDER, *Mécanismes d'Endommagement dans les Composites Multicouches à Renforts Interlock* - Doctorate thesis, Université de Technologie de Compiègne, France, 2011.

[8] C. RAKOTOARISOA - *Prévision de la Durée de Vie en Fatigue des Composites à Matrice Organique Tissés Interlock*. Doctorate thesis, Université de Technologie de Compiègne, France, 2013.

[8] C. RAKOTOARISOA, F. LAURIN, M. HIRSEKORN, J.-F. MAIRE, L. OLIVIER - *Development of a Fatigue Model for 3D Woven Polymer Matrix Composites Based on a Damage Model.* ECCM15 - 15th European Conference on Composite Materials. Venice, Italy, 1-8, 24-28 June 2012.

[10] J. HENRY, Z. ABOURA, K. KHELLIL, S. OTIN - *Suivi de l'Endommagement en Fatigue d'un Composite à Renfort Interlock Carbone/Epoxy par Emission Acoustique*. JNC17- 17ème Journées Nationales des Composites. Poitiers, France, 1-11, 15-17 June 2011.

[11] A. HURMANE, *Analyse par un Dialogue Essais/Calculs de la Tenue en Compression de Structures Composites Tissées 3D* - Doctorate thesis, Université de Technologie de Compiègne, France, 2015.

[12] P. FEISSEL, J. SCHNEIDER, Z. ABOURA, P. VILLON - *Use of Diffuse Approximation on DIC for Early Damage Detection in 3D Carbon/Epoxy Composites*. Composites Science and Technology 88, 16-25, 2013.

[13] F. LAURIN, J.-S. CHARRIER, D. LEVÊQUE, J.-F. MAIRE, A. MAVEL, P. NUÑEZ - *Determination of the Properties of Composite Materials Thanks to Digital Image Correlation Measurements.* Procedia IUTAM 4, 106-115, 2012.

[14] C. HUCHETTE - *Analyse Multiéchelle des Interactions entre Fissurations Intralaminaire et Interlaminaire dans les Matériaux Composites Stratifiés.* Doctorate thesis, Université de Paris VI, France, 2005.

[15] N. GODIN, S. HUGUET, R. GAERTNER, L. SALMON - *Clustering of Acoustic Emission Signals Collected During Tensile Tests on Unidirectional Glass/ Polyester Composite Using Supervised and Unsupervised Classifiers.* Non Destructive Testing and Evaluation International 37, 253-264, 2004.

[16] A. SCHIEFFER - *Modélisation multiéchelle du comportement mécanique des composites à matrice organique et effets du vieillissement thermique.* Doctorate thesis, Université de Franche-Comté, France, 2003.

[17] C. HOCHARD, S. MIOT, N. LAHELLEC, F. MAZEROLLE, M. HERMAN, J.-P. CHARLES - *Behaviour up to Rupture of Woven Ply Laminate Structures Under Static Loading Conditions.* Composites Part A: Applied Science and Manufacturing 40, 1017-1023, 2009.

[18] S. PIMENTA, S.T. PINHO - *Hierarchical Scaling Law for the Strength of Composite Fibre Bundles.* Journal of the Mechanics and Physics of Solids 61, 1337-1356, 2013.

[19] A. DOITRAND, C. FAGIANO, V. CHIARUTTINI, F. LEROY, A. MAVEL, M. HIRSEKORN - *Experimental Characterization and Numerical Modeling of Damage at the Mesoscopic Scale of Woven Polymer Matrix Composites Under Quasi-static Tensile Loading.* Composites Science and Technology 119, 1-11, 2015.

[20] S. BLASSIAU, A. THIONNET, A.R. BUNSELL - *Micromechanisms of Load Transfer in a Unidirectional Carbon Fibre-Reinforced Epoxy Composite Due to Fibre Failures*. Part I. Composite Structures 74, 303-318, 2006.

[21] A. PUCK, H. SCHURMANN - *Failure Analysis of FRP Laminates by Means of Physically Based Phenomenological Models.* Composites Science and Technology 62, 1633-1662, 2002.

[22] A. PUCK, M. MANNIGEL - *Physically Based Non-Linear Stress/Strain Relations for the Inter-fibre Fracture Analysis of FRP Laminates.* Composites Science and Technology 67, 1955-1964, 2007.

[23] J.F. MAIRE, J.L. CHABOCHE - *A New Formulation of Continuum Damage Mechanics (CDM) for Composite Materials.* Aerospace Science and Technology 1, 247-257, 1997.

[24] L. MARCIN - *Modélisation du Comportement, de l'Endommagement, et de la Rupture des Matériaux Composites à RenfortsTissés pour le Dimensionnement Robuste de Structures*. Doctorate thesis, Université de Bordeaux I, France, 2010.

[25] A. HURMANE, F.-X. IRISARRI, F. LAURIN, S LECLERCQ, M.L. BENZEGGAGH - *Strength Analysis of Woven Interlock Composites Subjected to Compressive Loading: Experiments and Simulations.* ECCM16 - 16th European Conference on Composite Materials. Sevilla, Spain. 1-8, 22-26, June 2014.

[26] L. MARCIN, J.F. MAIRE, N. CARRÈRE, E. MARTIN - *Development of a Macroscopic Damage Model for Woven Ceramic Matrix Composites.* International Journal of Damage Mechanics 20, 939-957, 2011.

[27] F. LAURIN, N. CARRERE, J.F. MAIRE - *A Multiscale Progressive Failure Approach for Composite Laminates Based on Thermodynamical Viscoelastic and Damage Models.* Composites Part A: Applied Science and Manufacturing 38, 198-209, 2007.

[28] F. LAURIN, N. CARRERE, C. HUCHETTE, J.-F. MAIRE - *A Multiscale Hybrid Damage and Failure Approach for Strength Predictions of Composite Structures.* Journal of Composite Materials, special issue for the WWFE-III Part A, 47 (20-21), 2713-2747.

[29] C. BOUVET, S. RIVALLANT, J.J.,BARRAU - *Low Velocity Impact Modeling in Composite Laminates Capturing Permanent Indentation.* Composites Science and Technology 72, 1977-1988, 2012.

[30] N. HONGKARNJANAKUL, C.BOUVET, S. RIVALLANT V*alidation of Low Velocity Impact Modelling on Different Stacking Sequences of CFRP Laminates and Ifluence of Fibre Failure.* Composite Structures 106, 549-559, 2013.

[31] A. ELIAS - *Nocivité des Défauts Induits Par Impact Pour Les Structures Composites Tissées 3D À Matrice Organique.* Doctorate thesis, Ecole Centrale Nantes, France, 2015.

[32] N. CARRERE, F. LAURIN, J.-F. MAIRE - *Micromechanical Based Hybrid Mesoscopic 3D Approach for non-Linear Progressive Failure Analysis of Composite Structures* - Part B : Comparison with experimental data. Journal of Composite Materials, special issue for the WWFE-III Part B, 47 (6-7), 733-741, 2013.

[33] G. DUVAUT, J.-L. LIONS - *Inequalities in Mechanics and Physics, Springer-Verlag, Berlin 1976.* French edition: Dunod, Paris, 1972.

[34] J.M. WHITNEY, R.J. NUISMER,. *Stress Fracture Criteria for Laminated Composites Containing Stress Concentrations.* Journal of Composite Materials 8, 253-265, 1974.

[35] R.J. NUISMER, J.D. LABOR - *Applications of the Average Stress Failure Criterion: Part II - Compression. Journal of Composite Materials 13, 49-60*, 1979.

[36] J. LEE, C. SOUTIS - *Measuring the Notched Compressive Strength of Composite Laminates: Specimen Size Effects.* Composites Science and Technology 68, 2359-2366, 1979.

[37] W. WILSON - *Predicting In- and Out-of-Plane Damage Evolution in Woven Fibre-Reinforced Composites.* JEC Composites Magazine 100, 64-67, 2015.

[38] M. HOFFMAN, V. OTTO, T. HAVAR, E. AHCI - *Numerical and experimental evaluation of fatigue performance of bearing laminates.* ECCM17 - 17th European Conference on Composite Materials. Munich, Germany. 1-10, 26-30 June 2016.

[39] O. GENCOZ, U.G. GORANSON, R.R. MERRILL - *Application of Finite Element Analysis Techniquesvfor Predicting Crack Propagation in Lugs.* International Journal of Fatigue 2, 121-129, 1980.

[40] K.C. WARREN, R.A. LOPEZ-ANIDO, J. GOERING - *Behavior of Three-Dimensional Woven Carbon Composites in Single-Bolt Bearing.* Composite Structures 127, 175-184, 2015. .

[41] K.C. WARREN, R.A. LOPEZ-ANIDO, S.S. VEL, H.H. BAYRAKTAR - *Progressive Failure Analysis of Three-Dimensional Woven Carbon Composites in Single-Bolt, Double-Shear Bearing.* Composites Part B: Engineering 84, 266-276, 2016.

[42] R. DESMORAT, L. ANGRAND, P. GABORIT, M. KAMINSKI, C. RAKOTOARISOA - *On the Introduction of a Mean Stress in Kinetic Damage Evolution Laws for Fatigue.* International Journal of Fatigue 77, 141-153, 2015.

[43] A. ELIAS, F. LAURIN, M. KAMINSKI, L. GORNET - *Experimental and Numerical Investigations of Low Energy/velocity Impact Damage Generated in 3D woven composite with polymer matrix.* Composite Structures 159, 228-239, 2017.

Acronyms

3D	(3 Dimensional)
2D	(2 Dimensional)
ODM-PMC	(ONERA Damage Model for Polymer Matrix Composites)
DIC	(Digital Image Correlation)
AE	(Acoustic Emission)
SEM	(Scanning Electron Microscope)
FE	(Finite Elements)
GB	(Giga Byte)
RAM	(Random Access Memory)
CPU	(Central Process Unit)

AUTHORS

Antoine Hurmane graduated (Engineering Diploma) from the French Institute for Advanced Mechanics (IFMA), and received a Master Degree from Blaise Pascal University (Clermont-Ferrand France) in 2011. He received his PhD in mechanics from the University of Technology of Compiègne in 2015. Since he works at ONERA – The French Aerospace Lab as a research scientist in the Composite Materials and Structures Department. His research includes the development of damage and failure approaches for composite materials, and experimental characterization of damage and failure of composite materials.

Frédéric Laurin graduated (Engineering Diploma) from *École Centrale Marseille* (ECM) in 2002 and received his PhD in mechanics from the University of Franche-Comté in 2005. He joined ONERA in 2005 and works as a research scientist in the Composite Materials and Structures Department. His research interest includes the development of damage and failure approaches for composite materials, strength predictions of composite structures through finite element modeling and experimental investigations on the damage and failure mechanisms encountered in such materials. He is the head of the modeling and characterization unit in the composite department at ONERA since 2016.

Anne Mavel has joined ONERA in 1990, and is in charge of mechanical testing of composite materials in the Composite Materials and Structures Department. Her work consists in the characterization of composite materials through mechanical tests from elementary coupons to components. Since 2015, she also performs thermal characterization on composites with ceramic matrix in order to measure their thermo-mechanical performances under severe environmental conditions.

Pascal Paulmier graduated from a DUT in Mechanics in 1981, and specialized in "aerospace" since 1982. He joined ONERA in 1983. Since 2005, he is the experimental engineer in charge of experimental studies from small samples to large structures and of carrying out mechanical tests (static and fatigue loadings) under different environmental conditions in the metallic materials department and then in the composite materials department.

V. Bonnand, D. Pacou
(ONERA)

E-mail: vincent.bonnand@onera.fr

DOI: 10.12762/2016.AL12-04

Complex Thermo-Mechanical Approaches to Study the Behavior of High-Temperature Alloys

The work discusses the complex thermo-mechanical approaches developed at ONERA within the context of gas turbine component fatigue tests. Emphasis is placed on the general experimental methods to test nickel-based single crystal superalloys used in gas turbine blades under severe thermal and mechanical conditions. First, the successive steps required in combined thermal and mechanical fatigue testing are briefly discussed. Particular focus is placed on the techniques developed to reliably control and measure temperature fields generated by induction heating. The main experimental techniques used for thermo-mechanical testing are then discussed. Here, an axial-torsional thermo-mechanical fatigue (TMF) testing rig used to study the coupling effects of combined thermal and mechanical loads is described. How a thermal gradient is introduced in the TMF specimens to reproduce the internal air cooling technology of modern gas turbine blades is shown. In order to increase the representativeness of in-service conditions, the problem of thermo-mechanical fatigue coupled with thermal gradients is investigated using smooth, multi-perforated, and thermal barrier coated samples. It is shown that such testing capabilities can reproduce realistic thermomechanical loading conditions typical of service and constitute a powerful means to develop and validate fatigue life prediction methods. Furthermore, the effect of an overheating event that can occur during a one-engine-inoperative event during in-service operation is also investigated. This type of engine malfunction is known to lead to a degradation of the microstructure, which can modify the material constitutive behavior and affect the remaining lifetime. Finally, the additional complexities associated with the creep-fatigue interaction, the required constitutive models and the computational challenges are discussed.

Introduction

Most industrial components operating at high temperatures, such as those in gas turbines, present a complex thermal loading including high temperature, rapid temperature transient and temperature heterogeneity. The case of single-crystal turbine blades is particularly representative of the level of complexity reached in the rotating components of an aero-engine. These turbine blades are subjected to combined thermal and mechanical loading (Figure 1). They are subjected to a centrifugal force directly correlated to the rotation speed. The resulting mechanical stresses are cycled according to the operational speed of the engine. The aerodynamical loading inherent to rotative machines generates low stress amplitude at a high frequency. In parallel, these blades are the hottest components in the gas turbine. They reach heterogeneous temperatures above 1100°C. Moreover, in order to avoid failure resulting from excessive operating

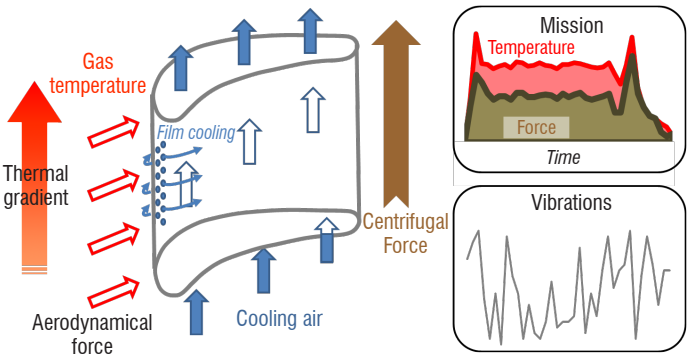


Figure 1 – Schematic description of the complex thermo-mechanical loading of turbine blades

temperatures, turbine blades often use many cooling methods, such as internal channels, film cooling and thermal barrier coatings. These cooling technologies favor temperature heterogeneity and generate thermal stresses to be superimposed on the mechanical loading.

A combined thermal and mechanical loading may develop into various damage mechanisms, such as fatigue, creep, oxidation and also microstructural evolutions. However, the significant interactions between these degradation mechanisms are quite complex, and the prediction of thermo-mechanical damage from basic properties remains difficult. There is an industrial and scientific need to generate material test data under non-isothermal service conditions. In order to be of a good quality, the tests must be conducted in the well-controlled environment of a laboratory.

Over the last six decades, non-isothermal fatigue testing has become an increasingly important method for design, material performance and reliability assessment. In the early 60's, many of the experiments involved thermal cycling. Thermal Fatigue became the terminology adopted for tests with no external mechanical loading. First, thermal fatigue test devices are essentially the same as that designed originally by Coffin [14] and later modified by Carden [10]. Here, a repeated temperature fluctuation in a restrained tube produces cyclic thermal stresses and strains, which lead to fatigue cracks. Another approach to thermal fatigue was also developed on component-like specimens, such as Glenny's tapered discs or wedge shaped specimens tested in fluidized beds [18], or rigs based on heating by the flame of a burner and cooling by compressed air [44]. In these cases, the stresses and strains obtained in the specimen are self imposed, due to the differential thermal expansion brought about by temperature gradients. An example of thermal analysis of the wedge specimens is given in [28]. At ONERA, Chaboche [11] investigated in 1972 the viscoplastic behavior of high-purity Aluminum under thermal cycling. The exploratory test specimen is presented in Figure 2. The hottest zone can reach 200°C with two infrared lamps. The cycle was based on two hold times; the longest one was at the maximal temperature of 200°C, whereas the second one was performed at 90°C.

In parallel to the thermal fatigue improvements, the necessity to perform thermomechanical experiments with externally-imposed loading appears necessary to support constitutive model developments. The

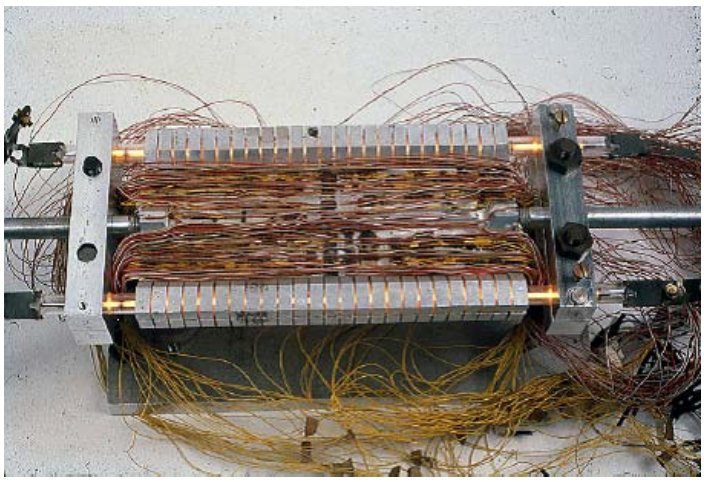


Figure 2 – Pioneering specimen equipped with 60 thermocouples and 40 strain gauges for thermal gradient fatigue (1972). The specimen is heated by two infra-red lamps inside the two external columns and cooled in the middle by a water inlet [11]

generation of precise thermomechanical response data has become possible with the development of closed-loop servohydraulic testing machines [20]. The term "thermomechanical fatigue" became dedicated to strain-controlled non-isothermal tests. Today, the literature references highly thermo-mechanical fatigue. For an overview of more recent activities in this field, the reader should refer to Remy [41], McGaw [36], and Hähner [19] in a special issue of International Journal of Fatigue. Currently, one of the challenges for the topic of thermal and mechanical fatigue is to increase the representativeness of the thermo-mechanical loading compared to the in-service conditions, as well as the representativeness of the specimen compared to real structures.

The overall objectives of the thermo-mechanical approaches discussed in this experimental research are to:

- test calculation methods to improve the life prediction,
- understand the mechanisms that govern non-isothermal damages and appear significantly different to those in the isothermal cases.

This paper reveals why non-isothermal fatigue is a more complex test technique that requires higher skills than traditional mechanical testing. The experimental approaches developed here clearly show the proximity of the turbo-engine industry. Four kinds of non-isothermal fatigue are presented, according to an industrial thermal feature of turbine blades representing the hottest parts of aircraft or helicopter engines. The first step is the characterization in pure thermo-mechanical fatigue when temperature homogeneity is considered. The second step includes a thermal gradient in the specimen wall, as encountered in cooled turbine blades due to internal cooling channels. The third topic is dedicated to severe overheating, which can be encountered in a helicopter engine. In this case, only one exotic cycle appears to be non-isothermal and plays a key role on the remaining life duration of the piece. Finally, the Thermal Barrier Coating within the context of thermal gradient mechanical fatigue will be discussed for the original thermal calibration approach developed.

Thermal aspects in thermal and mechanical testing

This section details the heating system and the temperature measurement and control methods used in the various non-isothermal testing approaches.

Heating methods

High-temperature mechanical testing requires a specific heating system suitable for a servo-hydraulic or electromechanical device. Currently used techniques include resistance heating, radiant heating and (direct or indirect) induction heating. Nevertheless, under non-isothermal conditions, the test cycle is basically based on engine flight, often combining take-off, cruising and landing in the most simplified case. In order to achieve a reasonable test duration in a laboratory, the thermal cycle is accelerated so the temperature ramp rates may exceed 10°C/s and cannot be achieved by conventional resistance furnaces.

Although radiant lamp furnaces or induction coil systems offer a high heating rate capability, another technical possibility comes from burner heating. This type of expensive device consists of a real combustion chamber. It enables the reproduction of experimental conditions close to those encountered in turbo-engines. For this purpose, a burner rig appears interesting to obtain a very high temperature testing device with high heating and cooling rates. The cooling rates can reach 60°C/s due

to the ceasing of the burning and may subject the material to quenching. The MAATRE burner developed at ENSMA – Institut Pprime illustrates the associated metrology and the performance of this test bench [38].

For complex thermo-mechanical approaches, induction heating was preferred to the radiant lamp furnace. It offers a far cheaper and more reliable solution and especially presents the high reactivity required to simulate fast cycles. Another advantage is the possibility of adapting the induction coil to a specific specimen or for measurement access. It could also be a significant drawback, because this operation is time-consuming and cannot easily offer temperature homogeneity as would be the case in a conventional radiant furnace. The different induction coils presented in this paper are always suitable for the constraints and needs of the experiments. Furthermore, the induction coil may provide the ability to perform, as far as possible, in-situ measurement techniques, such as visible image acquisition for Digital Image Correlation, or Infra-Red images for thermography.

In order to properly use induction heating, it is very important to understand the basics. The important feature of the induction heating process is that the heat is generated inside the specimen itself, instead of by an external heat source. Indeed, an inductor coil generates a variable magnetic field. If an electrical conductive part is placed inside this magnetic field, it carries the electrical current induced, also called Foucault currents. In accordance with the Joule effect, the movement of the electrons creating these currents dissipates the heat in the material. Thus, two basic parameters of this heating method must be taken into account:

- The penetration length that characterizes the distribution of the electrical current.
- The dissipated power in the specimen that characterizes the efficiency of the electrical phenomena.

For a semi-infinite plate, the penetration length p is given by the distance from the surface to a density equal to J_0 / e. It is expressed by:

$$p = \sqrt{\frac{1}{\pi f \sigma \mu}} \qquad (1)$$

where σ is the electrical conductivity of the material, μ the magnetic permeability, and f is the frequency of the alternative magnetic field.

The electrical current distribution in the specimen depends on the frequency of the induction unit. Our set-ups are based on a middle-frequency unit at 100 kHz. Few data are available on the electrical resistivity of a Ni-based superalloy. By considering the superalloy CMSX-4 [40], the penetration length is estimated to be 0.72 mm at 1000°C, compared to our specimen thicknesses (1 or 2 mm).

Temperature measurement and control

In high-temperature mechanical testing, the various temperature measurement methods, which can be used either separately or in combination, are:

- thermocouples (N or S type),
- pyrometer,
- infrared camera,
- thermal paintings.

From a global point of view, spot-welded thermocouples are preferred for temperature measurement. Pyrometers and IR cameras depend on the emissivity. Without a stable surface emissivity, results can prove to be inaccurate(temperature drift), as the specimen oxidizes and its emissivity is modified accordingly. Moreover, oxide layer cracking during a test can also significantly alter the measurement. With regard to the thermal painting, only irreversible ones are available at high temperature. This means that they cannot be used under thermal cycling within the context of a gas turbine.

Nevertheless, the presence of a spot-welded thermocouple is not acceptable in the gauge length during a fatigue test because it could cause an alteration in the material and the welded zone might be a preferential site for crack initiation. For this reason, spot-welded thermocouples are used only on specimens for calibration of the nominal temperature, according to the TMF standards [1, 5 and 25]. However, a thermocouple control can be spot-welded outside the gauge length. Indeed, all of the experiments presented in this research have additional water cooling at the specimen grips. This allows the gauge length to remain the hottest part of the specimen and it ensures a dynamically stable temperature field at the specimen ends. Another advantage of this approach is that the cooled ends may offer the possibility of testing new advanced materials that exhibit a very high temperature resistance. In ONERA's ongoing material development activities, sometimes the construction process does not allow for large specimens. This implies that the required gripping system must be machined from a lower quality alloy. Thus, a significant permanent deformation of the grips can be expected, which could also lead to their premature rupture. In return, the thermocouple is less sensitive to local temperature variations, which must be avoided by the control system.

Thermal calibration: general aspects

Calibration specimen

For non-isothermal tests, the general thermal calibration process is based on the TMF-standard [1, 25]. Before each campaign, one of the specimens is dedicated to a thermal calibration process. This step on dummy specimen is aimed at:

- optimizing the inductor coil, in terms of geometry and positioning,
- mapping the temperature field on the heated specimen.

As the basic element of a non-isothermal test, the heating system set-up (power and inductor shape) must be optimized on the calibration specimen with respect to a temperature field, which must be as uniform as possible, and to the thermal cycle. This means that thermal homogeneity must be guaranteed at each instant of the cycle. Classically, the desired zone represents at least the gauge section of the specimen. Nevertheless, it is sometimes not possible under complex thermal situations, as presented in Figure 3. Thus, another strategy

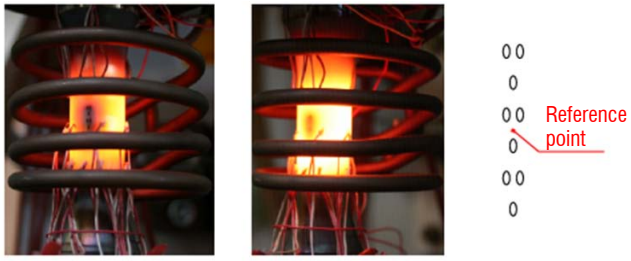


Figure 3 – Impact of the positioning of the induction coil on the temperature distribution in the case of thermal Gradient Mechanical Fatigue on a perforated specimen. (a) Initial cold area for a symmetric position (b) Minimized cold area (c) Reference point for temperature measurement by optical pyrometry

consists in measuring and controlling the nominal temperature at a reference point on the specimen. In this case, we do not seek to avoid temperature heterogeneity, but rather to measure precisely the thermal field of the outer surface. The map must be integrated analytically or numerically using a finite element approach with conventional data interpolation methods.

Whatever the strategy, based on the temperature uniformity sought or on the temperature field subjected to, thermal profiling is required based on thermal imaging or mapping by several thermocouples fixed along and around the specimen. The points of interest given by the thermocouples must be sufficiently numerous for the measured temperatures to accurately represent the thermal gradients over the external surface of the heated specimen.

Temperature control methods

In most of anisothermal fatigue problems, the temperature cycle will remain constant throughout the duration of the test. This assumes that:

- the thermal equilibrium of the load train has been reached,
- there is a dedicated temperature measurement and an associated control system,
- the repeatability and reproducibility of the thermal cycling has been checked.

The thermal equilibrium of the load train is a preliminary step before dynamic calibration. This point is very important because our devices are always equipped with water-cooled specimen grips. It should be noted that no compressed air cooling systems are added to control temperature rates. After this, one thermocouple or pyrometer can serve to perform the temperature control on the dummy specimen. With regard to the temperature control, for simple cases, the experiment could be summarized as the application of a command signal. Nevertheless, in order to ensure the temperature stability during cycling, a control-loop is preferred. Indeed, several temperature deviation sources can be encountered, such as unintended sporadic air flows, a variation of exposure to direct sunlight, or evolution of the thermal response due to an alteration of electrical properties of the material. During the temperature control step, it is necessary to check that the temperature indicated by other non-control sensors, like the thermocouple network of infrared cameras, does not vary.

A classical control-loop system can be used for temperature cycling. All of our testing setups for non-isothermal fatigue are based on a servo-hydraulic fatigue device with the MTS Test suite software. The main advantage of this global controlling system is the possibility of generating and acquiring all of the thermal and mechanical signals synchronously. Additionally, other systems, such as image acquisition or the potential drop method can also be triggered on the command or response.

Thermo-mechanical approaches for single crystal turbine blades

This section is dedicated to thermo-mechanical cycling of uncoated specimens. A schematic description of the experimental loadings is given in Figure 4. The first described setup concerns a strain-controlled thermo-mechanical fatigue test (TMF). The following topics concern the study of Thermal Gradient Mechanical Fatigue conditions (TGMF). The experimental studies have been performed on the Ni-based single crystal superalloy called AM1.

Strain-controlled Thermo-Mechanical Fatigue (TMF)

The first equipment was designed in order to simultaneously simulate thermal and mechanical loads under strain control. The main objective is to assess the accuracy of the constitutive equations under non-isothermal conditions, as well as the accuracy of life time prediction models.

The main issue of strain-controlled TMF experiments relates to the fact that the mechanical strain cycle can only be performed by controlling the total strain applied to the specimen. This means that the thermal strain in the absence of external forces has to be compensated for in an appropriate way, in order to achieve a given mechanical strain:

$$\varepsilon_{mechanical} = \varepsilon_{total} - \varepsilon_{thermal} \quad (2)$$

Hence, special attention must be paid to the dynamic temperature measurement and control because a temperature deviation modifies the imposed mechanical strain, obviously changing the material behavior and leading to an incorrect life duration. In order to avoid

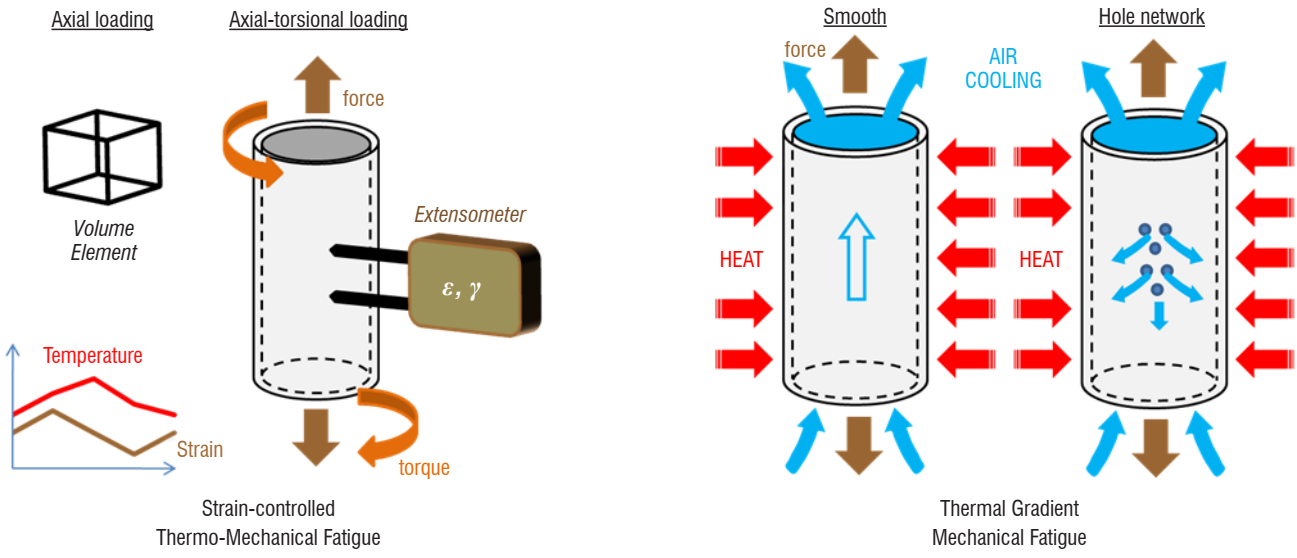


Figure 4 – Schematic description of TMF and TGMF testing developed in this research

and check it, the TMF device presented was developed according to the dedicated experimental procedure of the standard. The testing principles are detailed here ([25] and references therein) and can be divided into three main parts:

- thermal cycling under force control to record thermal strain,
- null force verification in thermal strain control,
- application of additional mechanical strain under extensometer control.

In a time-based control mode, a temperature error automatically leads to an error in the mechanical strain, according to Equation 2. This implies that special attention must be paid to avoiding temperature deviations. The second step of TMF procedure enables the stability and the reproducibility of the thermal loading to be validated. If the nominal temperature deviates or the longitudinal temperature gradient evolves, the resulting force of a thermal strain control will vary. A typical example of the thermal compensation obtained with a temperature range of 650°C to 1100°C is presented in Figure 5, also illustrating the testing rig.

Furthermore, although uniaxial thermo-mechanical fatigue is highly

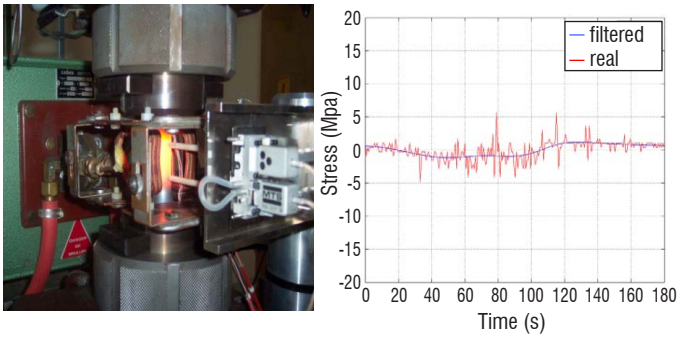


Figure 5 – Presentation of the axial-torsional TMF test device and of an axial thermal compensation cycle [7]

referenced in the literature and recently dedicated standards have appeared [1, 25], its extension to multiaxial loading remains infrequent. Axial-torsional configurations [26, 47, 7 and 9] can be discussed and, more recently, biaxial planar configurations [30, 42]. In our case, an axial-torsional loading was initially developed in [7]. Torsional loading presents specificities linked to single crystal anisotropy. On the one hand, no thermal compensation is necessary due to cubic symmetry of the thermal expansion coefficient and, on the other hand, crystalline viscoplasticity generates high strain bands in the <011> secondary directions. This point implies that care must be taken with the secondary direction of the crystal, in which one extensometer measures and imposes the shear strain used in the servo-loop.

It is worth noting that a comparison between tests and constitutive models could be performed based on the mechanical strain, by subtracting the thermal strain recorded during the first step from the total measured strain. An intermediate step prior to the TMF test can be introduced, in order to assess the Young and the shear moduli. The best practice consists in superimposing a force loading in the elastic regime (or, alternatively, a torque) to the same temperature cycle, as applied in the subsequent TMF test. The elastic moduli are determined as a function of temperature and then the data analysis can be done with the viscoplastic part of the strain (i.e., without the thermal and elastic strains). This research was performed using several strain-temperature paths. The experimental results have been compared with Finite Element simulation based on crystal viscoplasticity identified for AM1. Very good agreement is observed with a temperature range between 650°C and 1100°C, as presented in Figure 6. The selected formalism including octahedral and cubic slip systems has been validated under anisothermal and multiaxial conditions [7].

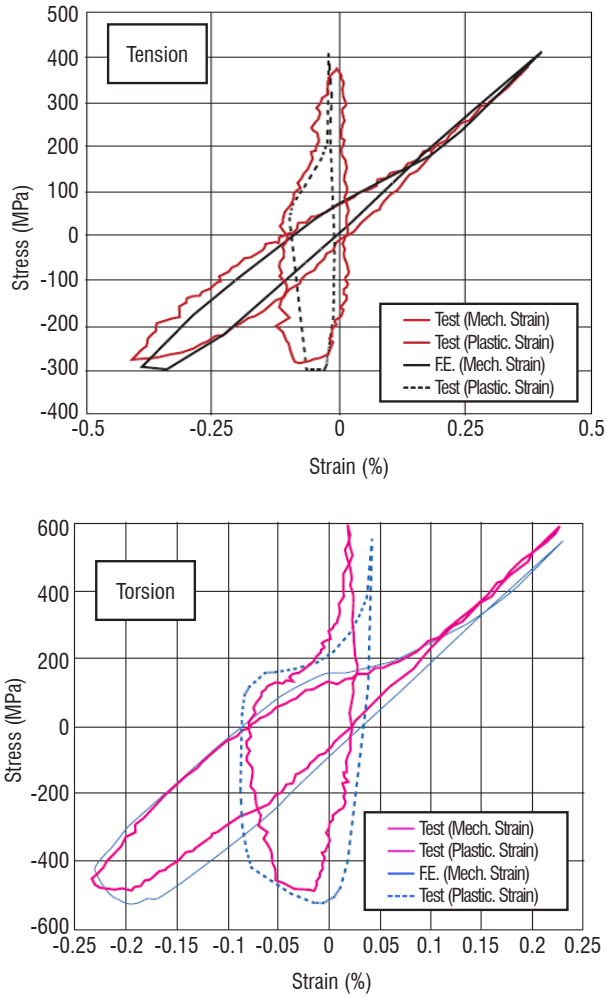


Figure 6 – Typical results for axial and torsional responses of the material [7]

Thermal Gradient Mechanical Fatigue (TGMF)

Another research topic for design engineering relates to the impact of an internal cooling system on the turbine blade life. Although anisothermal experiments are generally performed on volume elements simultaneously subjected to a controlled load and temperature, as previously presented, these tests may not be representative of cooled turbine blades, where a thermal gradient is generated over a very thin wall. The thermal strains and stresses induced by the spatial gradient during heating and cooling may be a significant cause of failure in current single-crystal blades of aero gas turbines.

In order to study the influence of the thermal gradient on fatigue life, experimental initiatives have been promoted for turbine blade analysis. At ONERA, a first experimental setup was developed at the end of the 90's [27], enabling a specific thermal gradient in the region just outside the holes of the thick specimen to be reproduced in the laboratory. Recently, this device was considerably modified and improved by [16] for thin specimens. It can be noticed that similar TGMF testing

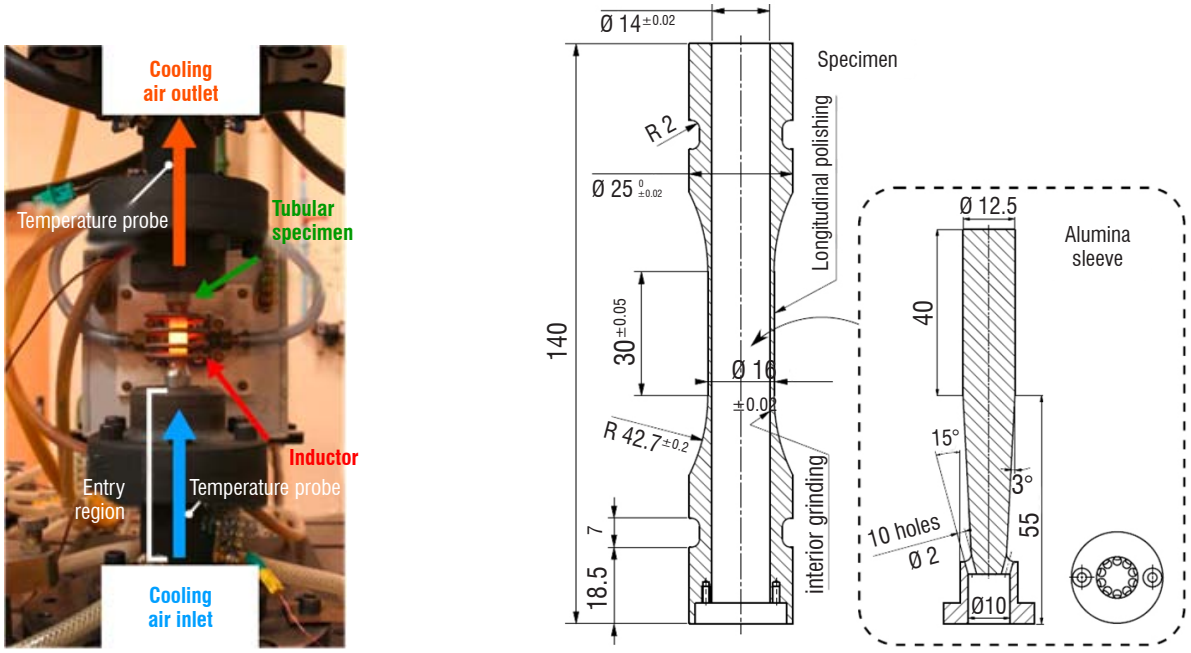


Figure 7 – Photo of the TGMF device and schema of the tubular hollow specimen with alumina sleeve

devices are still rare in the literature [13, 4, 3, 21 and 22]. Although the TGMF facility was firstly concerned with smooth specimens, their use in the presence of multi-perforations, such as on turbine blades, appears to be useful for the gas turbine industry.

TGMF on a smooth specimen

In order to simulate the temperature gradient acting on the blade, the experimental facility developed enables the cooling of the wall of the hollow cylindrical fatigue specimen, during the heating of the external wall. The difficulty lies in obtaining the highest possible thermal gradient, but with an acceptable gauge section temperature distribution. Figure 7 shows the experimental setup and the geometries of the specimen and the sleeve required to increase the heat exchange. The principle is the following one: the tubular specimen is heated by induction and a significant cold air flow is injected inside, in order to obtain a sufficient wall thermal gradient. In order to increase the air flow, and to thereby increase the thermal gradient, an alumina sleeve can be inserted into the hollow specimen.

The TGMF is characterized by (i) force-control fatigue testing, (ii) a temperature control at the external surface of the tubular specimen similar to those encountered in TMF testing, (iii) cooling of the air flow control and (iv) temperature assessment at the internal surface, characterizing the thermal gradient.

Cooling is provided by a compressed air circuit with a thermal mass flow meter at the inlet of the specimen for measuring and controlling mass flow. The mass flow meter has a maximal capacity of 53.88 g/s and a Proportional-Integral-Derivative (PID) controller. The air temperature at the inlet is 15°C and the maximal air flow reached with this installation and a mounted specimen is 43 g/s at 5 bars (28 g/s with the sleeve). The air pressure downstream from the specimen is atmospheric pressure. It is important to mention that small variations of less than 0.5 g/s around the set point are observed.

The main difficulty of this thermo-mechanical test is that the internal surface temperature cannot be measured by common methods like thermocouples or pyrometers. In order to compensate for this lack, special attention was paid to numerically assessing the temperature gradient. The thermal problem is solved by a weak coupling between fluid and solid solvers, based on two software applications developed at ONERA, respectively Cedre and ZéBuLoN. Figure 8 presents the calculated thermal field of the specimen with the sleeve and the exchange coefficient. It exhibits a peak near the injector because the air flow is projected on to the specimen surface. This effect decreases along the specimen z axis, but an injector effect still remains up to the sleeve end. After the peak due to the injector, the exchange coefficient decreases along the specimen axis. Along the gauge length the average value is around 1100 $W.m^{-2}.K^{-1}$, which gives an average thermal gradient of 42°C.

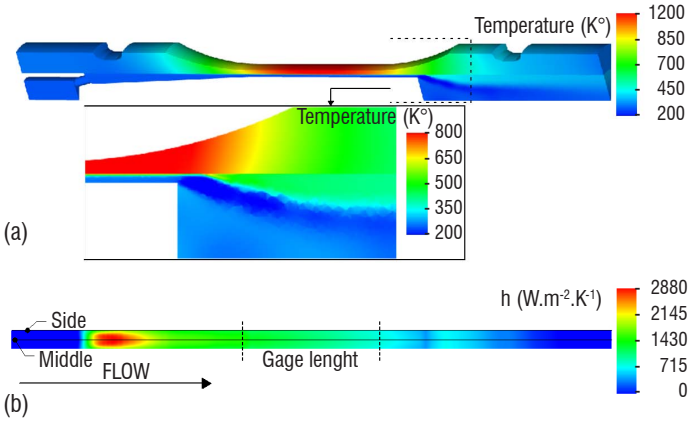
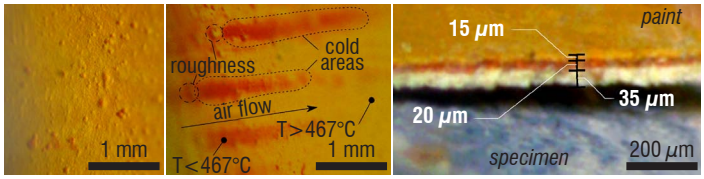


Figure 8 – Coupled calculation results (a) Specimen and air temperature fields (b) Exchange coefficient h at the interface

In order to be able to trust this internal temperature assessment numerical approach, the global heat exchange can be compared with calculations. Its evaluation is characterized by two temperature probes set upstream and downstream from the specimen. Details of the turbulent model and the resulting quite good correlation are presented in [16]. Nevertheless, improvements should be made to several calculation hypotheses, such as the adiabatic sleeve or the internal temperature distribution in the anchors.



(a) (b) (c)

Figure 9 – Thermal paint onthe internal surface after a TGMF test

An opportunity to increase the thermal gradient involves the application of thermal paint on the inner surface of the specimen. Initially added to measure the internal temperature without a thermal cycle, the thermal paint exhibited significant roughness, considerably increasing the turbulence of the cooling air flow. Figure 9 shows the roughness and gives an example of the colour gradient obtained for a thermal gradient up to 120°C/mm in the superalloy. Temperature heterogeneity is observed only close to the air injection, and the temperature gradient in the paint thickness can be also analyzed.

Afterwards, the effect of induction heating on the thermal gradient was also investigated. Induction heating allows a very high heating power and also a great responsiveness for temperature cycles. However, it has the disadvantage of heating the material itself by Joule effect, which affects wall temperature distributions. This could be calculated by considering the tube wall as a semi-infinite plate in which an analytical solution exists:

$$\dot{q}_{induction}(r) = \frac{J_0^2}{2\sigma_e}\exp\left(-\frac{2(r-r_e)}{p}\right) \qquad (3)$$

where $\dot{q}_{induction}$ is the heat source generation power, J_0 is the surface current density, σ_e is the electrical conductivity, r_e is the external specimen radius and p is the penetration depth. The temperature distribution with induction heating or with surface heating (representative of radiant heating for example) can be calculated using the thermal conduction FEM. A convection coefficient h is imposed on the internal surface, associated with an air temperature T_{air}=15°C, as the surface temperature is T_{surf}. The thermal flux leaving the specimen is given by $q_{cool} = h\cdot(T_{surf} - T_{air})$. The surface heating is imposed through a thermal flux on the external temperature and the induction heating by thermal sources in the wall thickness. A schematic comparison of the two cases is shown in Figure 10. The temperature distribution is almost linear for surface heating, whereas it is parabolic for induction heating. The reason is that the heating power is maximal on the external surface and decreases as the distance to the internal surface decreases. It also means that for a same internal convection coefficient, the thermal gradient obtained with an induction heating system is lower than that obtained with a surface heating. The ratio ΔT between the surface heating to induction heating is virtually constant at around 1.27. This illustrates the main disadvantage of the induction heating technology.

Although the presence of a thermal gradient during pure mechanical fatigue can be investigated, a simplified engine-like cycle (Figure 11) was also defined to validate constitutive damage models in the case of realistic and complex thermo-mechanical loadings. The imposed stress cycle is in-phase with the temperature cycle. The thermo-mechanical cycle was obviously shortened to limit the test duration and to present high heat and cooling rates (65°C/s). These rates can be reached due to the high internal cooling flow. Likewise, a high power induction system is required (16 kW). During the thermal calibration, its reproducibility was investigated according to the fluctuations of the controlled air flow. Standard deviations at the extreme values of each sub-cycle are acceptable, as illustrated in Figure 11.

Any experimental and numerical works allowed us to validate the lifetime prediction methods of the single crystal AM1. The lifetime prediction chain is obtained after solving the thermal problem. The stresses and strains are computed using constitutive equations. This is followed in the stabilized cycle by the calculation of the lifetime prediction itself, which is the determination of the number of cycles for crack initiation. Fully reversed cycles at 950°C and complex thermo-mechanical cycles have been investigated. The life prediction chain was confronted with experimental measurements in terms of mechanical response, damage localisation and crack initiation lifetime. The results are detailed in [16]. It is worth noting that necking

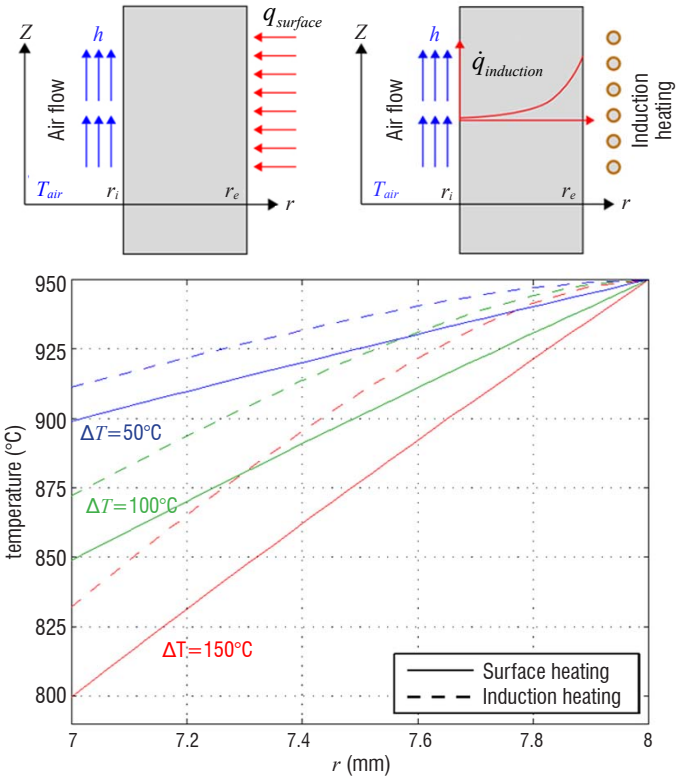


Figure 10 – Comparison of the calculated thermal gradient between the induction and the surface heating for various exchange coefficients and an external surface temperature of 950°C

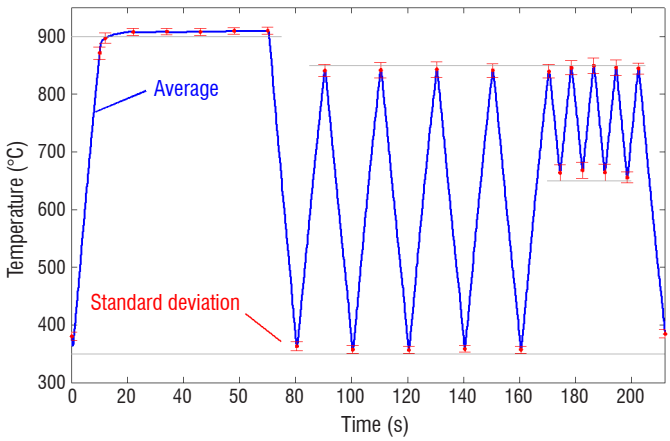


Figure 11 – Engine-like external surface temperature measured on a specimen over 30 cycles

phenomena occurred for high stress levels leading to a lower number of cycles for crack initiation than those simulated. For future works, this leads us to pursue TGMF tests under the strain control mode.

TGMF on a perforated specimen

In civil or military applications, the geometry of air-cooled blades becomes very complex, with areas involving a cooling hole-pattern at the leading edge of turbine blades. These geometrical singularities where cracks can initiate induce significant mechanical gradients additionally to the thermal one. The experimental device previously presented allowed us to evaluate the influence of these multi-perforations on the local behavior and life time.

The drilling pattern representative of areal blade was adapted for a tubular specimen with 8 mm radius; the channels are tilted with a 400 μm diameter.

In this research [16], a constant thermal gradient with fatigue cycle and engine-like cycle are also performed. Once again, a numerical analysis was performed in order to adequately compute the thermal fields inside the cooling channels. The same approach was carried out with coupled calculations in Thermal Solid/Fluid mechanics [29]. Here, the temperature control at the outer surface was performed by a pyrometer pointing at the reference point discussed in Figure 3. In order to compensate for the heat sinks represented by the holes, the inductor coil was off-centre, the injected power being dependant on the distance between the surface and the coil. Nevertheless, a strong necking is always localized near a two-hole row up to the opposite side of the specimen, as observed in Figure 12. The measured circumferential thermal gradient (100°C) seems to generate a high level of creep in the hottest area. For the two tests performed at different stress range, cracks are nucleated in the two-hole row of the middle pattern. Nevertheless, it appears difficult to know the precursory mechanism between necking and crack initiation. A future investigation using digital image correlation may provide some answers.

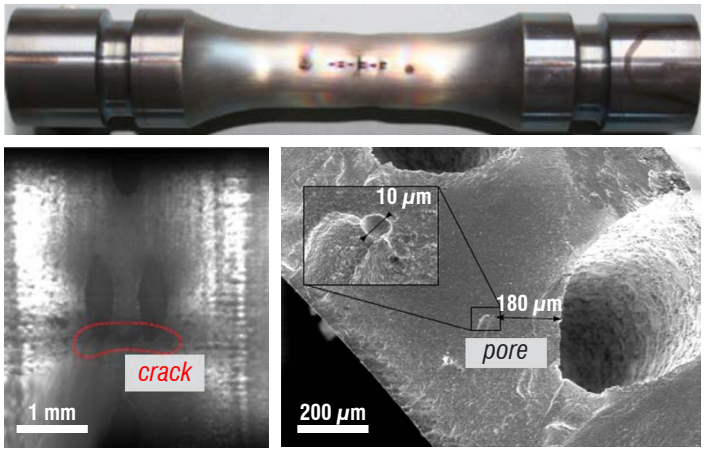
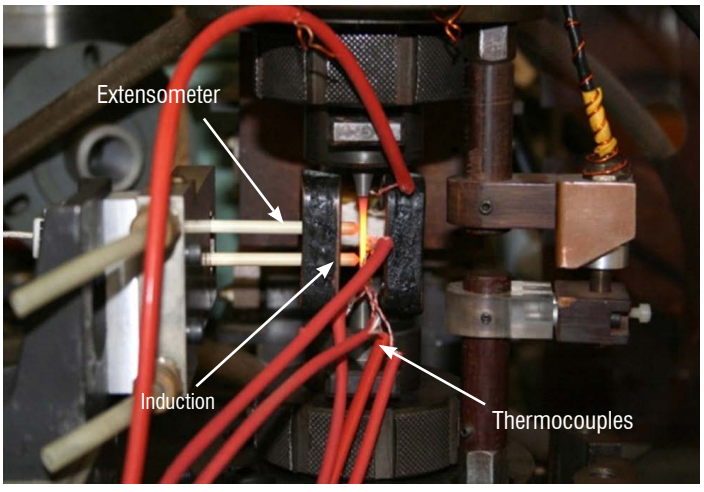


Figure 12 – Optical and SEM observations with localized necking and crack initiation linked to the interaction between holes (from [16])

The experimental simulation of the One-Engine-Inoperative event in turbine blades (OEI)

Although thermomechanical cycling is representative of the classically localized conditions experienced by components during the take-off, cruising, descent, and landing cycle of an engine, the rare and sudden appearance of an engine failure must also be explored. Indeed, during in-service operation of twin-engine helicopters, one of the two engines may unexpectedly be cut off. This implies a sudden temperature rise, up to 1200°C in some sections of the high pressure turbine blades for the remaining operating engine. This kind of event is called a One-Engine-Inoperative (OEI) rating, and is now included in helicopter turbo-engine certification procedures [15]. Consequently, fatigue-creep tests under isothermal conditions became insufficient to tackle the characterization of mechanical properties under non-isothermal loading conditions.

Here, the main challenge consists in reproducing this sudden temperature increase in the laboratory, by avoiding any temperature overshoot (Figure 13). In order to achieve this goal, a special inductor was required (i) to enable accessibility for an extensometer (ii) to allow good temperature homogeneity and (iii) to offer the good electrical efficiency needed for a high heating rate. The longitudinal gradient and its evolution during the OEI indicate that the gauge length is essentially heated by conduction. Indeed, the volumetric heat source distribution representing the induction heating is more localized outside of the gauge length, leading to a small overshoot outside of the gauge length. This is not an inconvenience due to the lower temperature. We can specify that this thermal response can be obtained only with temperature-loop control (Proportional-Integral-Derivative controller). Indeed, the microstructural evolution seems to significantly alter the electro-magnetic properties of the superalloy, modifying the thermal response. This special handling of the temperature control enables us to achieve temperature levels and heating/cooling rates like those encountered in a real OEI event.



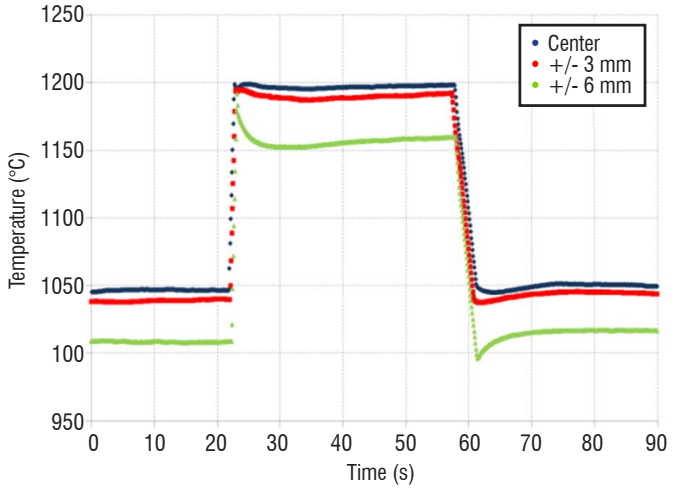


Figure 13 – Presentation of the OEI fatigue test device with induction coils and the thermal response during the OEI event (from [33])

The influence of an OEI event on mechanical responses has been well studied under creep conditions [34] and also in cyclic loadings [32]. The study of the creep-fatigue tests for <001> and <111> oriented specimens has shown that the plastic strain rate always decreases after one overheating. The results obtained provide some information on the effects of fine γ' precipitates on the mechanical behavior (Figure 14) and emphasize the kinetics of precipitation and dissolution, as detailed in [31, 33]. These tests revealing a strong crystallographic orientation effect on the anisothermal conditions are also extended in a multiaxial configuration (axial-torsional) [35]. Finite element simulations using a crystal-plasticity microstructure-sensitive model are necessary to understand the intricate experimental stress redistribution caused by the anisotropy of the material and are discussed elsewhere [15].

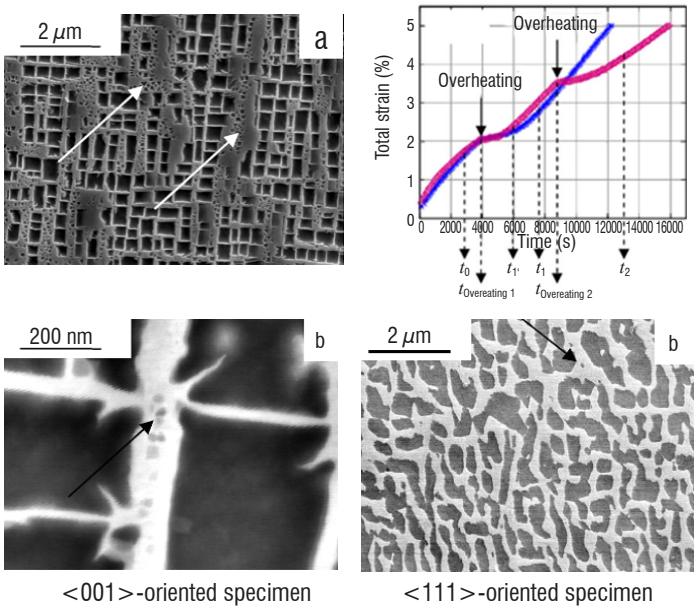


Figure 14 – Microstructures observed after overheating for 180s, with arrows indicating a g channel with fine precipitates (from [33]) – On the left: general view and blown-up focusing on the <001>specimen; on the right: Plastic strain evolution during creep-fatigue experiments and the associated microstructure for the<111> specimen

The peculiar case of thermal barrier coating

As a next step after the internal cooling system applied to turbine blades, Thermal Barrier Coatings (TBC) are often used on the turbine blades to decrease the service temperature down to over 100 °C. TBC are a multilayer system including a thermal insulating porous ceramic layer, mostly of 8 wt.% yttria partially stabilized zirconia, deposited on an alumina-forming metallic bond coat in contact with the single crystal superalloy substrate.

In service, the structure and composition of each layer can evolve, due to sintering of the ceramic layer, oxidation of the bond coat, and interdiffusion phenomena with the substrate. This affects the mechanical properties, such as the interfacial toughness, and consequently may cause the ceramic layer to spall off in the end. For more details on these degradation mechanisms, the reader is referred to [2]. For industrial partners, the lifetime assessment of TBC systems has been a challenge and a goal of prime importance since their introduction in aircraft engines. For this purpose, an energetic based model was developed at ONERA and calibrated by means of adhesion tests [43,45, 24, 12].

It is well know that, in the presence of a sufficiently large thermal gradient, cracks form and propagate on delamination planes in the TBC parallel to the interface, resulting in regions that spall away, leaving a thin layer of zirconia still attached to the substrate [23]. In the literature, experimental studies are often limited to thermal cycling without external mechanical loading [41,39], even if the degradation mode does not arise either when the system is thermally cycled within a furnace or when tested in a burner rig. In order to fill such a gap, the TGMF device previously presented was adapted for Thermal Barrier Coatings. The set-up evolution results from the knowledge of the thermo-physical properties of yttria partially stabilized zirconia. We decided to take advantage of its significant ionic conductivity at high temperatures [17]. This enables us to pursue the heating of a coated specimen by the induction system. From room temperature, only the substrate will heat by the Joule effect induced by the magnetic field and, given that the specimen temperature will increase, a consistent part of the heat will be directly generated in the TBC. In order to evaluate the temperature gradient in the wall thickness of a tubular specimen, thermal computations are performed with the following considerations:

- forced internal convection (exchange coefficient h provided by coupled aerothermal/solid simulation in the case of an uncoated specimen [16]),
- Joule effect in the superalloy (exponential curve of the heat source, linked to induction),
- Joule effect in the TBC (significant ionic conductivity in the TBC at high temperatures[17]),
- natural external convection.

Figure 15 clearly shows that, at a high temperature, ionic conductivity enables an adequate thermal gradient to be obtained in the coating, particularly when tests are performed with a sleeve. The temperature appears to be quite constant across the TBC, compared to the hypothetical case of a non-conductive coating.

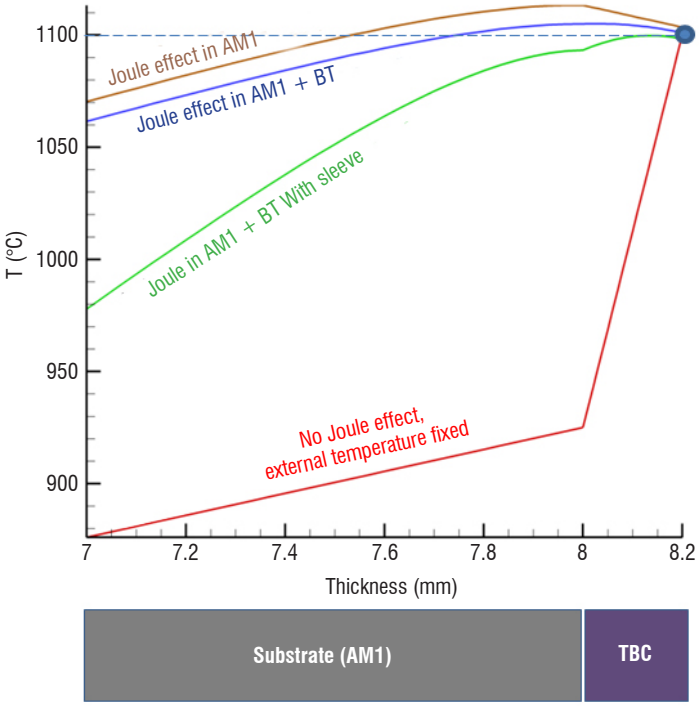


Figure 15 – Estimated thermal gradient in the coated specimen thickness, according to several assumptions regarding the heat generation for an external temperature of 1100°C

The temperature at the metal/oxide interface is well known to be a key parameter for the system lifetime assessment. Nevertheless, this interface temperature is not easily accessible by conventional measurement methods. Moreover, a thermocouple cannot obviously be welded on to the TBC due to its nature. Although Manara *et al.* [37] measured the optical properties of this semi-transparent medium, the real emissivity of the coating and its evolution remains unknown. In order to evaluate the interface temperature precisely during thermal cycling, an original thermal calibration procedure was proposed by a coupled experiment/calculation approach. On a dedicated specimen, an artificial spallation was made manually. The specimen is then meshed with the coating including the spalled zone. A set of thermal calculations enable us to obtain the relationship between the temperatures at different locations. First, the calibration test enables the temperature pointed by pyrometer at the thermal barrier coating and that measured by a thermocouple spot-welded in a spalled superalloy to be linked. For a fixed emissivity parameter, this link is a function of the internal cooling and of the external temperature. The numerical simulations offer the relationship between the exposed substrate and the interface temperature far from the spallation. Afterwards, a heating power chart can be achieved to assess the interface temperature.

Another challenging point is the determination of the service life of the thermal barrier; it is crucial to be able to distinguish the first steps of the TBC spallation. A simple way is to take advantage of the contrast of optical images, which is linked to the emissivity variation between the TBC and the natural oxide layer. Optical mirrors are added, in order to observe the entire circumference of the hollow specimen. Figure 16 shows an example of the spallation observed at a high temperature (1100°C).

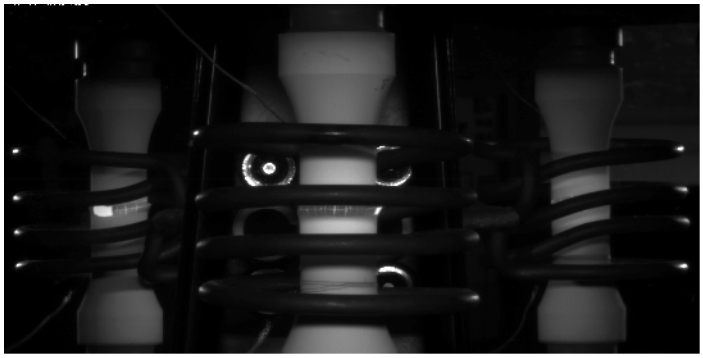


Figure 16 – Examples of the image captured at a high temperature (Observation in one shot of significant spallation in the left mirror)

First, preliminary tests are performed to validate the ability of the TGMF testing device for a coated specimen. The thermal cycles in their experiments consisted of a 60 minute exposure to a temperature of 1100°C at the interface between the TBC and the superalloy, followed by a 2 min air cooling prior to a 2 min reheating. The thermal gradient over the wall thickness introduces another source of complexity. In future studies, this device will allow us to analyze the degradation mechanisms of TBC under representative thermal gradient conditions. Obviously, a simulated thermal map deduced from experiments should be transferred into 3D Finite Element simulations, in order to perform lifetime calculations, which should be compared with lifetime assessment tools for thermal barrier systems.

Conclusions & Perspectives

The strong demand from aero-engine manufacturers and certification authorities to improve the fatigue life assessment of gas turbine blades has led to the development of advanced thermo-mechanical fatigue testing. ONERA's activities in this field have traditionally been associated with the prediction of the complex associated phenomena. These specific experimental developments have been performed to underpin the development of lifing methods and/or to gain a greater understanding of fatigue behavior and its associated mechanisms under realistic anisothermal conditions. The present work has focussed on the thermo-mechanical approaches developed at ONERA to study successively the superposition of thermal and mechanical loadings, the impact of a thermal gradient and of turbine blade overheating during a one-engine-inoperative event, and coated test specimen behavior under thermal gradient mechanical fatigue. A common feature of these complex experiments is that advanced numerical studies are required to analyze the effect of complex service thermal loading at either the component and/or the microstructure level.

It has been shown that axial-torsional thermo-mechanical fatigue testing enables the validation of the superalloy behavior under anisothermal conditions. Moreover, it is known that the cooling systems in gas turbine blades enhance the role of high temperature fatigue when compared to uncooled blades, where creep is dominant. The presence of holes on cooled blades also affects the component lifetime, and such behavior must be taken into account by life prediction methods. Models also need to be validated against the combined effects of temperature and stress gradients and, in some cases, thermal barrier coatings under these realistic conditions. Finally, an investigation into the implications of turbine malfunction ,such as the one-engine-inoperative event required by certification authorities, regarding the microstructural degradation and associated changes in high-temperature material behavior has enabled ONERA to adapt fatigue life models accordingly.

The representativeness of in-service loading conditions for the fatigue testing of gas turbine materials is always an experimental challenge. In the future, these complex tests will continue to be strongly underpinned by computational models. Future studies will be more driven to account for the ever-increasing complexity of thermo-mechanical loading conditions, which need to be reproduced at the laboratory scale. For instance, accumulative anisothermal fatigue with high cycle fatigue, like multiaxiality [8],will need to be properly evaluated when designing and predicting the life of engineering components. Furthermore, the extension of these complex thermo-mechanical approaches to incorporate crack growth is still in its infancy. By developing effective fatigue testing facilities and by enriching existing measurement methods with advanced imaging, thermography, and tomography techniques, etc., it is expected that the fatigue durability assessment of gas turbine engine components will become increasingly realistic in the future ■

Acknowledgments

These works are also the fruit of trainee and doctoral students. The authors wish to sincerely thank J. B. Le Graverend (currently at Texas A&M University, USA) for the OEI testing and analyzes, R. Degeilh (currently at EDF Lab Saclay, France) for the development of the TGMF device and J. C. Anglès (currently at Stanford University, USA) for the TBC study. The authors would like to thank J. D. Garaud and ANDHEO for their precious help in performing the coupled simulations. Furthermore, the authors are also particularly grateful to the SNECMA – SAFRAN group for providing the material and to the French Ministry of Defence for its financial support. The main part of these works was conducted under the French program "PRC Structures Chaudes" involving the SNECMA-SAFRAN group, the Turbomeca-SAFRAN group, ONERA and other laboratories.

References

[1] ASTM. E 2368-04 - *Standard Practice for Strain Controlled Thermomechanical Fatigue Testing*. Philadelphia (2004) PA, 3.01

[2] M.-P. BACOS, J.-M. DORVAUX, O. LAVIGNE, R. MEVREL, M. POULAIN, C. RIO, M.-H. VIDAL-SETIF - *Performance and Degradation Mechanisms of Thermal Barrier Coatings for Turbine Blades: a Review of ONERA Activities*. Aerospace Lab Issue 3, 2011.

[3] M. BARTSCH, B. BAUFELD, S. DALKILIC, L. CHERNOVA, M. HEINZELMANN - *Fatigue Cracks in a Thermal Barrier Coating System on a Superalloy in Multiaxial Thermomechanical Testing*. International Journal of Fatigue 30 (2008) 211-218.

[4] B. BAUFELD, M. BARTSCH and M. HEINZELMANN - *Advanced Thermal Gradient Mechanical Fatigue Testing of CMSX-4 with an Oxidation Protection Coating*. International Journal of Fatigue, 30:219-225, 2008.

[5] T. BECK, P. HAHNER, C. RAE, E. AFFELD, H. ANDERSSON, A. KOSTER, M. MARCHIONNI - *Standardization Thermo-Mechanical Fatigue – The Route to the Standardization (TMF-Standard Project)*. Materials and Corrosion, 57, N°1.

[6] T. BECK, K. RAU - *Temperature Measurement and Control Methods in TMF Testing – A Comparison and Evaluation*. Int. Journal of Fatigue, 30, p. 226-233, 2008.

[7] V. BONNAND - *Etude de l'endommagement d'un superalliage monocristallin en fatigue thermo-mécanique multiaxiale.* Thèse ENSMP-ONERA, 2006.

[8] V. BONNAND, J. L. CHABOCHE, P. GOMEZ, P. KANOUTÉ, D. PACOU - *Investigation of Multiaxial Fatigue in the Context of Turboengine Disc Applications*. International Journal of Fatigue, 33-8, p. 1006-1016, 2011.

[9] S. P. BROOKES, H.-J. KÜHN, B. SKROTZKI, H. KLINGELHÖFFER, R. SIEVERT, J. PFETZING, D. PETER, G. EGGELER - *Axial-Torsional Thermomechanical Fatigue of a Near-γTiAl-Alloy*. Materials Science and Engineering: A, Volume 527, Issues 16-17, 25 June 2010, p. 3829-3839.

[10] A. E. CARDEN and T. B. SLADE - *High-temperature Low-cycle Fatigue Experiments on Hastelloy X, Fatigue at High Temperature*. ASTM STP 459, 1969, p. 111-129.

[11] J.-L. CHABOCHE - *Calcul des déformations visco-plastiques d'une structure soumise à des gradients thermiques évolutifs*. Thèse de Doctorat de 3ème cycle, Orsay, 1972.

[12] J.-L. CHABOCHE, F. FEYEL, M. POULAIN, N. RAKOTAMALALA, A. ROOS, J.-R. VAUNOIS, A. LONGUET, AND P. KANOUTE - *Lifetime Assessment Tools for Thermal Barrier Systems in "Thermal Barrier Coatings IV"*. German Aerospace Center; ECI Symposium Series, 2015.

[13] J.-L. CHABOCHE, J. P. CULIE, F. GALLERNEAU, D. NOUAILHAS, D. PACOU AND D. POIRIER - *Thin Wall Thermal Gradient: Experimental Study F.E. Analysis and Fatigue Life Prediction*. 5th International conference on biaxial/Multiaxial Fatigue & Fracture, Cracow (Poland), 1997.

[14] L. F. COFFIN - *A Study of the Effects of Cyclic Thermal Stresses on a Ductile Metal*. Transactions of ASME, 1954, vol. 76, p. 931-950.

[15] J. CORMIER, F. MAUGET, J. B. LE GRAVEREND, C. MORICONI, J. MENDEZ - *Issues Related to the Constitutive Modeling of Ni-based Single Crystal Superalloys under Aeroengine Certification Conditions*. Aerospace Lab Journal vol.7, August 2015.

[16] R. DEGEILH - *Développement expérimental et modélisation d'un essai de fatigue avec gradient thermique de paroi pour application aube de turbine monocristalline*. thèse LMT Cachan-ONERA, 2013.

[17] J. W. FERGUS - *Electrolytes for Solid Oxide Fuel Cells*. Journal of Power Sources, Volume 162, Issue 1, 8 November 2006, p. 30-40.

[18] E. GLENNY - *A Study of the Thermal-Fatigue Behaviour Metals*. Journal of the Institute of Metals, Paper No 2010 ,1970, p. 449-461.

[19] P. HAHNER, C. RINALDI, V. BICEGO, E. AFFELDT, T. BRENDEL, H. ANDERSSON, T. BECK, H. KLINGELHOFER, H.-J. KUHAN, A. KOSTER, M. LOVEDAY, M. MARCHIONNI, C. RAE - *Research and Development into a European Code of Practice for Strain Controlled Thermomechanical Fatigue Testing.* Int. J. Fatigue, 30 (2008), p. 372-381.

[20] S. W. HOPKINS - *Low-Cycle Thermal Mechanical Fatigue Testing*. ASTM STP 612, 1976.

[21] N. X. HOU, Z. X. WEN and Z. F. YUE - *Tensile and Fatigue Behavior Of Thin-Walled Cylindrical Specimens Under Temperature Gradient Condition*. Journal of Materials Science, 43:1933-1938, 2008.

[22] N. X. HOU, Q. M. YU, Z. X. WEN, and Z. F. YUE - *Low Cycle Fatigue Behavior of Single Crystal Superalloy with Temperature Gradient*. European Journal of Mechanics - A : Solids, 29:611-618, 2010.

[23] J. W. HUTCHINSON, A. G. EVANS - *On the Delamination of Thermal Barrier Coatings in a Thermal Gradient*. Surface and Coatings Technology 149, p. 179-184, 2002.

[24] R. G. HUTCHINSON and J. W. HUTCHINSON - *Lifetime Assessment for Thermal Barrier Coatings: Tests for Measuring Mixed Mode Delamination Toughness*. Journal of the American Ceramic Society, 2011, 94:85-95.

[25] ISO/FDIS 12111 - Metallic Materials, Fatigue Testing, Strain Controlled Thermomechanical Testing Method.

[26] S. KALLURI, P. BONACUSE - *An Axial-Torsional Thermomechanical Fatigue Testing Technique, Multiaxial Fatigue and Deformation Techniques*. STP1280, p. 184, 1997.

[27] P. KANOUTE, D. PACOU, D. POIRIER, F. GALLERNEAU, J.-M. CARDONA - *Thermo-Mechanical Fatigue Life Analysis on Multiperforated Components.* Temperature-Fatigue Interaction, p. 341-350, 2002.

[28] A. KOSTER, G. LAURENT, G. CAILLETAUD and L. REMY - *Analysis of Thermal Fatigue Tests for Superalloy Components, Fatigue Under Thermal and Mechanical Loading: Mechanisms, Mechanics and Modelling*. Proc. of Symp., The Netherland, 22-24 May 1995.

[29] S. KRUCH, P. KANOUTE, V. BONNAND - *ONERA's Multiaxial and Anisothermal Lifetime Assessment for Engine Components*. Aerospace Lab Issue 09, August 2015.

[30] D. KULAWINSKI, A. WEIDNER, S. HENKEL, H. BIERMANN - *Isothermal and Thermo-Mechanical Fatigue Behavior of the Nickel Base Superalloy Waspaloy™ Under Uniaxial and Biaxial-Planar Loading*. International Journal of Fatigue, Volume 81, December 2015, p. 21-36.

[31] J. B. LE GRAVEREND, J. CORMIER, F. GALLERNEAU, P. PAULMIER, *Dissolution of Fine γ' Precipitates of MC2 Ni-Based Single-Crystal Superalloy in Creep-Fatigue Regime*. Eurosuperalloy 2010.

[32] J. B. LE GRAVEREND, J. CORMIER, F. GALLERNEAU, S. KRUCH, J. MENDEZ - *Strengthening Behavior in Non-Isothermal Monotonic and Cyclic Loading in a Ni-Based Single Crystal Superalloy*. IJF, 91 p. 257-263, 2016.

[33] J.-B. LE GRAVEREND, J. CORMIER, M. JOUIAD, F. GALLERNEAU, P. PAULMIER, F. HAMON - *Effect of Fine γ' Precipitation on Non-Isothermal Creep and Creep-Fatigue Behaviour of Nickel Base Superalloy MC2*. Materials Science and Engineering: A, Volume 527, Issue 20, 25 July 2010, p. 5295-5302.

[34] J.-B. LE GRAVEREND - *Étude et modélisation des effets d'incursion à très haute température sur le comportement mécanique d'un superalliage monocristallin pour aubes de turbine*. Thèse ENSMA-ONERA, 2013.

[35] J.-B. LE GRAVEREND - *Multiaxial Thermo-Mechanical Loading at High Temperature on a Ni-based Single Crystal Superalloy*. Oral presentation @ CREEP 2015. Article in preparation.

[36] M. A. McGAW, S. KALLURI, J. BRESSERS, S. D. PETEVES, editors - *Thermomechanical Fatigue Behavior of Materials*. ASTM STP 1428, vol. 4. West Conshohocken, PA: ASTM International; 2002.

[37] J. MANARA, M. AARDUINI-SCHUSTER, H.-J. RATZER-SCHEIBE, U. SCHULZ - *Infrared-Optical Properties and Heat Transfer Coefficients of Semitransparent Thermal Barrier Coatings*. Surface and Coatings Technology ,Vol. 203, Issue 8, p. 1059-1068, 2009.

[38] F. MAUGET, D. MARCHAND, G. BENOIT, M. MORISSET, D. BERTHEAU, J. CORMIER, J. MENDEZ, Z. HERVIER, E. OSTOJA-KUCZYNSKI, and C. MORICONI - *Development and Use of a New Burner Rig Facility to Mimic Service Loading Conditions of Ni-Based Single Crystal Superalloys*. Eurosuperalloys 14, 2014.

[39] F. MAUGET, F. HAMON, M. MORISSET, J. CORMIER, F. RIALLANT, J. MENDEZ - *Damage Mechanisms in an EB-PVD Thermal Barrier Coating System During TMF and TGMF Testing Conditions Under Combustion Environment*. IJF, accepted, 2016.

[40] K. C. MILLS, Y. M. YOUSSEF, Z. LI & Y. SU - *Calculation of Thermophysical Properties of Ni-based Superalloys*. ISIJ INTER., Vol. 46, N°. 5, p. 623-632, 2006.

[41] L. REMY, J. PETIT, editors - Temperature-fatigue interaction. ESIS Publication 29. Amsterdam: Elsevier; 2002.

[42] A. SCHOLZ, A. SAMIR, C. B ERGER - *Biaxial Thermomechanical Fatigue Experiments with Cruciform Test Pieces*. Proc. 7th of Int. Conf. on Biaxial and Multi-axial Fatigue & Fracture, Elsevier, Berlin, p. 555-560, 2004.

[43] J. SNIEZEWSKI, P. LOURS, Y. LE MAOULT - *Oxidation and Spallation of FeCrAl Alloys and Thermal Barrier Coatings: In Situ Investigation Under Controlled Temperature Gradient*. Materials at High Temperatures, Volume 27, Issue 2, 2010.

[44] D. A. SPERA, F. D. CALFO and P. T. BIZON - *Thermal Fatigue of Simulated Turbine Blades*. NASA TM X-67820, National Aeronautic Space Administration, May 1971.

[45] J.-R. VAUNOIS - *Modélisation de la durée de vie des barrières thermiques, par le développement et l'exploitation d'essais d'adhérence.* Thèse ONERA/Université de Grenoble, 2013.

[46] J.-R. VAUNOIS, J.-M. DORVAUX, P. KANOUTÉ, J.-L. CHABOCHE - *A New Version of a Rumpling Predictive Model in Thermal Barrier Coatings*. European Journal of Mechanics - A/Solids, Volume 42, November-December 2013, p. 402-421.

[47] J. ZIEBS, J. MEERSMANN, H.-J. KÜHN, H. KLINGELHÖFFER - *Multiaxial Thermo-mechanical Deformation Behavior of IN 738 LC and SC 16.* IV. Symposium on Thermo-mechanical Fatigue Behavior of Materials, ASTM STP 1371, West Conshohocken, USA, 2000, p. 257-278.

Acronyms

TBC	(Thermal Barrier Coating)
TMF	(Thermo-Mechanical Fatigue)
TGMF	(Thermal Gradient Mechanical Fatigue)

AUTHORS

Vincent Bonnand received his PhD in Materials Science and Engineering from the *Ecole des Mines de Paris* in 2006. He has been working at ONERA since 2007 in the Metallic Materials and Structures Department lab. His research topics concern the experimental mechanics at the intersection of different fields, such as solid mechanics, thermal analysis, associated measurements and non-destructive technologies.

Didier Pacou has been working at ONERA as an engineer in experimental mechanics in the Metallic Materials and Structures Department, since 1978. He is in charge of the complex fatigue devices and is involved in the development of specific devices and non-destructive technologies.

S. Giclais, P. Lubrina, C. Stephan
(ONERA)

E-mail : pascal.lubrina@onera.fr

DOI : 10.12762/2016.AL12-05

Aircraft Ground Vibration Testing at ONERA

During the development of a new aircraft, the Ground Vibration Test (GVT) is an important milestone because flutter computation relies on its results to determine the start of flight tests. Since the creation of ONERA after World War II, the GVT has always been the subject of fundamental research and application on industrial structures. On the occasion of the 70th anniversary of ONERA, an historical survey of the main methods and significant tests is proposed in this paper. After a brief introduction to the goals of a GVT, we highlight the continuous improvements made in this discipline in parallel with the progress in electronics, computers and algorithms, mainly focusing on internal studies at ONERA. Even though nowadays the GVT has reached maturity, aircraft dynamics is so challenging that several paths of research activities still remain. Keeping in mind the high quality standards expected by aircraft manufacturers, a vision of promising aspects for the future of GVT is finally proposed.

Introduction – A Brief History of Flutter

"*It is most probable that the first flutter accident of an aircraft occurred on December 8, 1903, when Professor Samuel Pierpont Langley, of the Smithsonian Institute, failed an attempted launch of his powered flying machine named "Aerodrome" from the Potomac River houseboat. That was only nine days before the Wright brothers' flight at Kitty Hawk*" [11]. Samuel P. Langley was nevertheless a renowned American astrophysicist and astronomer, who made airplane models powered by whirling arms and steam engines. In 1891, one of his unmanned models flew 3/4 of a mile before running out of fuel.

In this paper [11], one can read this description of flutter: "*From our present perspective, flutter is included in the broader term aeroelasticity, the study of the static and dynamic response of an elastic airplane. Since flutter involves the problems of interaction of aerodynamics and structural deformation, including inertial effects, at subcritical as well as critical speeds, it really involves all aspects of aeroelasticity (…). In man's handiwork, aeroelastic problems of windmills were solved empirically four centuries ago in Holland with the moving of the front spars of the blades from about the mid-chord to the quarter-chord position. We now recognize that some 19th century bridges were torsionally weak and collapsed from aeroelastic effects, as did the Tacoma Narrows Bridge in a spectacular fashion in 1940.*"

Why conduct a Ground Vibration Test on an aircraft?

Before going further, it is pertinent to recall the basic equation of flutter:

$$[M]\ddot{Z}(t)+[C]\dot{Z}(t)+[K]Z(t)= \left[C_{aero}\left(V^2\right)\right]\dot{Z}(t)+\left[K_{aero}\left(V^2\right)\right]Z(t) \quad (1)$$

where $Z(t)$ is the structure displacement vector. $[M]$, $[K]$ and $[C]$ are the positive definite mass, stiffness and damping matrices, respectively. They are on the left-hand side of equation (1) and represent the structural behavior. On the right-hand side of equation (1), $[C_{aero}(V^2)]$ and $[K_{aero}(V^2)]$ are the aerodynamic damping and stiffness matrices describing the unsteady aerodynamic forces.

The GVT is part of an aircraft certification process and its purpose is to measure the aircraft dynamic characteristics (natural frequency, mode shape, structural damping coefficient and generalized mass of the most important vibration modes).These results make the computation of flutter prediction possible, as well as the updating of the FEM (Finite Element Model) of the structure.

The fundamental equation is derived based on the assumption of linear second-order differential systems:

$$[M]\ddot{Z}(t)+[C]\dot{Z}(t)+[K]Z(t)=F(t) \quad (2)$$

Where the right member $F(t)$ represents the forces applied by means of excitation devices (illustrated in Figure 1). Thus, identifying the dynamics comes down to estimating the matrices $[M]$, $[K]$ and $[C]$ in a modal basis.

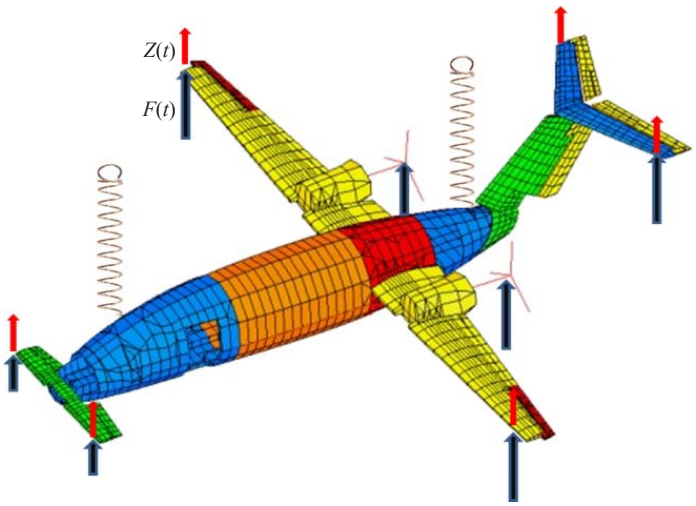


Figure 1 - Scheme of the Piaggio P180 Avanti FEM with forces, responses and soft suspensions

If real structures behaved as linear equation (2) predicts, it would be easy to characterize these matrices. Sensitive sensors and a few small forces well distributed over the structure would be enough to identify them. However, as the majority of vibratory modes are non-linear, in practice these matrices depend on response amplitudes and therefore on excitation forces. In general, only the damping and stiffness matrices present these peculiarities:

$$[M]\ddot{Z}(t)+\left[C\left(Z,\dot{Z}\right)\right]\dot{Z}(t)+\left[K\left(Z,\dot{Z}\right)\right]Z(t)=F(t) \qquad (3)$$

Nevertheless, for almost all vibration modes observed on aircraft, these non-linear parameters are small compared to the inherent linear ones. Then equation (3) can be written in the following form:

$$[M]\ddot{Z}(t)+[C]\dot{Z}(t)+[K]Z(t)+\varepsilon f\left(Z,\dot{Z}\right)=F(t) \qquad (4)$$

Where ε is a small scaling coefficient and f is the nonlinear function, also called the Restoring Force in the specialized literature.

As a consequence, for flutter prediction and FEM updating, aircraft manufacturers need linear modal parameters, obtained from the highest possible level of excitation forces. Of course, as the aircraft is expected to perform its first flight a few weeks later, the GVT has to be a nondestructive test.

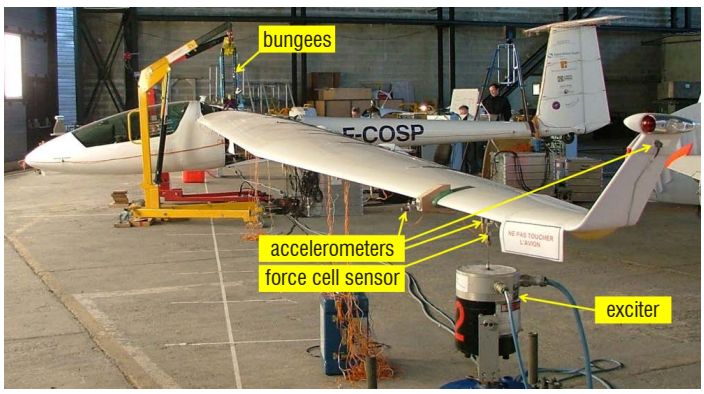


Figure 2 - GVT of a glider

Performing a GVT on an aircraft consists in applying external forces and measuring vibration responses. In addition, it should be noticed that, most generally, the aircraft will be uncoupled from the ground. To do that, the following devices are used nowadays (Figure 2).

The main difficulties come from nonlinear behavior, the high modal density and the possible lack of observability if the distribution of sensors is not optimal.

History of Ground Vibration Testing at ONERA: 1940s to 1970

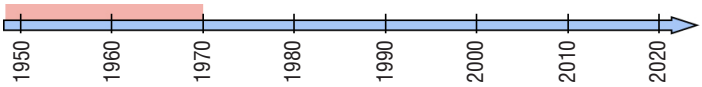


Figure 3 - GVT of the Do 17 R4 in 1939

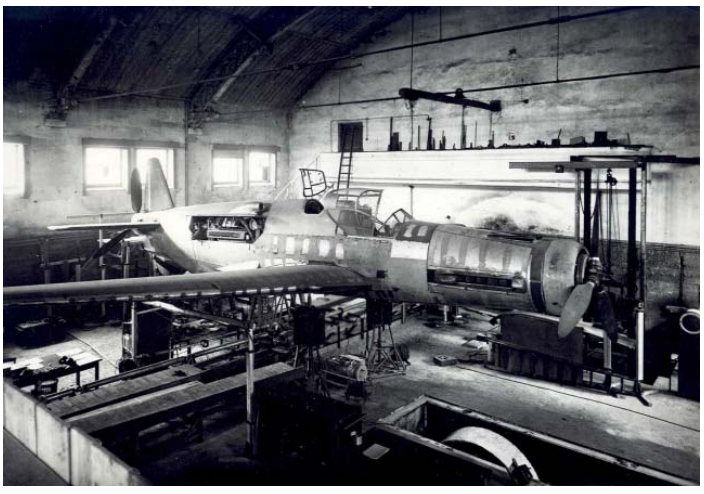


Figure 4 - GVT of the Do335 prototype at CTW

Just after the end of World War II, the French army based in the South of Germany created the "*Centre Technique de Wasserburg*" (CTW), particularly interested in aerodynamics and ground vibration testing. They performed an exhaustive analysis of the equipment, techniques and methods developed by the Dornier Company, based in

Friedrichshafen (Figure 3). One week after the official birth of ONERA, in May 1946, several top level civil and military personalities attended a demonstration of the GVT of an aircraft performed by Dornier (Figure 4). Some of those people, Mr. Mazet, Mr. Weber and Mr. Barrois, then became part of the management at ONERA.

At that time, excitation forces were delivered by unbalanced rotating systems and spring exciters controlled together with a sine signal, while measurements were performed with several fixed and mobile recorders, a tachymeter and a chronometer.

From Dornier technicians and engineers, most notably Mr. G. De Vries, and Wasserburg equipment, ONERA built its own team and applied the "Wasserburg" method for a GVT of the SNACASO SO-M1, Arsenal VG 70 and the SNCAN NC-271 aircraft in 1947. In parallel with the development and support on structural fatigue and structural static behavior, ONERA welcomed engineers of new French national aeronautic companies to train them in structural dynamics and flutter prediction.

Figure 5 - Mr. G. De Vries on the right hand side

Within the experimental and theoretical teams, French and German engineers developed new methods and techniques in a good atmosphere, as for instance the first developments and uses of electrodynamic shakers. Some of the Germans stayed in France, whereas some others went back and joined Mr. Dornier again at Friedrichshafen, working notably on the Dornier Do228 and the jet trainer Alphajet. Mr. G. De Vries [5] worked all along for the Structural department of ONERA until his retirement.

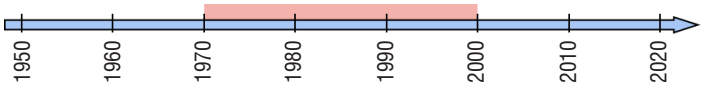
The demonstration of the GVT of a Wasserburg scale model in Meudon in December 1950 (Figure 5) gave national and international personalities in aeronautics and observers of many aircraft manufacturers an opportunity to focus on the benefit of flutter studies.

An example of test results is shown in Figure 6. It should be noticed that, although the Hertz (Hz) is nowadays the unit for frequencies, at that time the unit commonly used was cycles/minute (mn^{-1}).

The Phase Resonance Method, also called "Modal Tuning" or "Modal Appropriation", whose reputation remains linked to ONERA, was derived from the Wasserburg method. It relies on the Basile hypothesis, also called "Proportional generalized structural damping matrix hypothesis". Described by Mr. Basile, an engineer at ONERA, and mostly verified on aircraft structures, it makes the identification of all modal parameters possible, including structural damping coefficients.

This PRM, still used today, consists in measuring each vibration mode individually. It requires adjusting the excitation locations, force amplitudes and signs, and the frequency of the sine excitation signal. Using Lissajous ellipses between velocity responses and this excitation signal and using current controlled exciter amplifiers, developed by ONERA, make the use of this method easier.

The 1970s to 2000: the path to the modern era and ONERA as a major actor in GVT



Initially, the aircraft displacements were recorded by means of paper-band manual sensors (Figure 7a). These devices were travelled by operators at each measurement location and the quantities were manually evaluated from the paper-bands. Later, velocimetry sensors (Figure 7b) were used for vibration measurements. Their mobile parts were permanently linked to the aircraft, while their fixed parts needed to be supported by a huge number of scaffoldings.

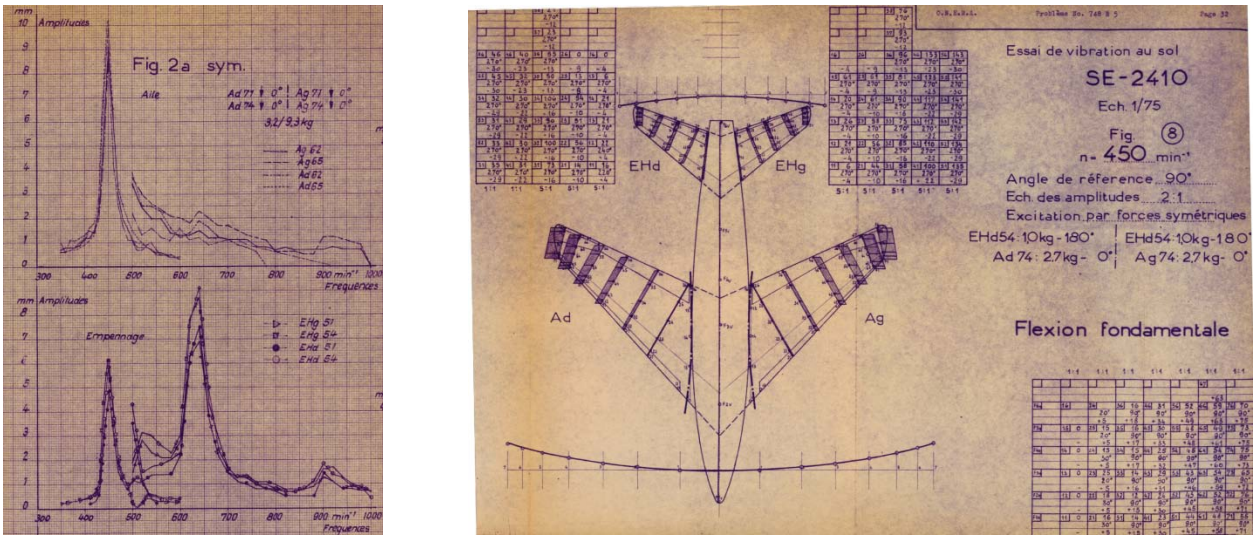


Figure 6 - Example of frequency responses and a mode shape plot (GVT of the SNCASE SE2410 in 1949)

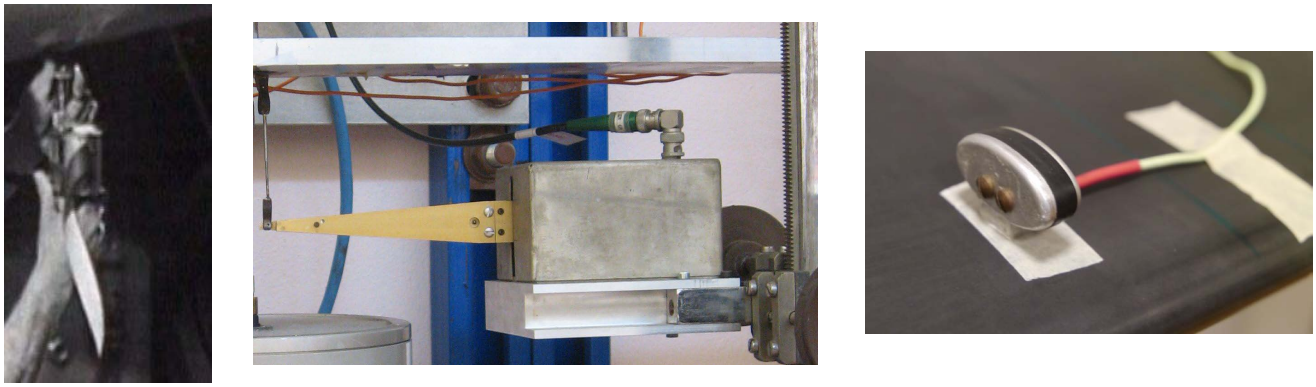


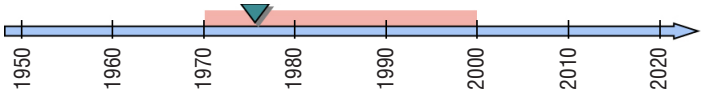
Figure 7 - ONERA displacement sensor, velocimetry sensor and accelerometer

They had the strong benefit of providing electrical signals recorded by an acquisition system. Thereafter, in the 1970s, ONERA developed small and affordable piezo-capacitive accelerometers (Figure 7c). Light and directly glued onto the aircraft, they did not need additional devices to support them. These were notably used for the first Airbus GVT in 1972 (see Figure 8).

With the first computers and FFT (Fast Fourier Transform) computation, one can say that the GVT entered the modern era.

The dissipation and generalized mass of a mode are now determined by the FQ (Force in Quadrature) method [5], which consists of superposing a dissipative force by means of an additional force in quadrature to the force making the aircraft vibrate at the resonance frequency. This method replaced the previous method consisting of adding local masses and deriving the generalized mass and the damping coefficient from the frequency deviations. Years later, to complement the FQ method, ONERA developed the CP (Complex Power) method [2].

Figure 8 - 1st Airbus (A300 B2) GVT in 1972



In 1977, for the GVT of the Ariane IV launcher components (see Figure 9), ONERA used the method called "Dat-Meurzec" for the first time. This was based on a curve-fitting process applied to the generalized coordinates and was, at ONERA, the first Phase Separation Method (PSM). Using frequency band responses measured, at that time, from stepped sine excitations, the method consists in applying a curve-fitting process from a first approximation of generalized coordinates in the near vicinity of each vibration mode.

Figure 9 - GVT of Ariane IV sections in 1977

Thanks to the continued increase in computer performance, this technique continued to improve and was

- tested in 1985 for a research GVT using a Dassault Aviation Falcon 20,
- used successfully in 1990 for a research GVT (Dassault Aviation, SOPEMEA, ONERA), dedicated to modal identification from excitation performed using control surface rotations on a Mirage 2000 [15],
- and finally used in 1999 during a research GVT (Airbus, DLR, SOPEMEA, ONERA) on an Airbus A340-300.

In parallel, other notable research activities were carried out on:

- the state of the art in EMA (Experimental Modal Analysis) in Europe, especially on damping estimation, coupled modes and nonlinearities
- GARTEUR (Group for Aeronautical Research and Technology in EURope) SM-AG19
  In the 1990s ONERA initiated a round robin survey of EMA in Europe, using a laboratory 2x2m model (see Figure 10), designed and manufactured by ONERA [1]. The model has

been sent to several laboratories and comparisons of results were made among GARTEUR associated partners. This model is today still a reference.

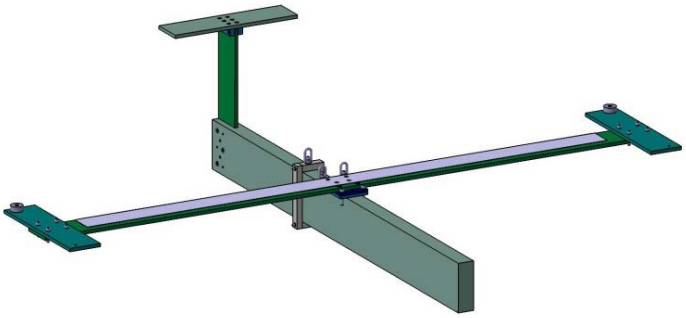
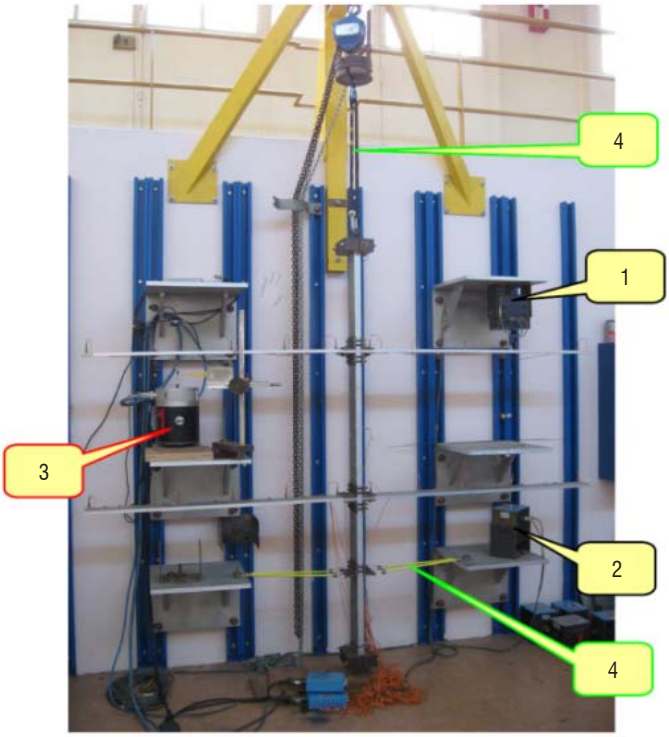


Figure 10 - GARTEUR SM-AG19 Computer-Aided Design

- New method for linear parameter estimation
  In the 1990s, Mr. C. Soize, ONERA Structural Department Director at that time, developed the Stochastic Linearization method with PhD students [10], [21], [22]. From time domain records of responses under random excitation, this method delivers the likeliest linear description of a structural behavior. Furthermore, the deviation between this linear model and the measurements provides information about nonlinearity contents. Applied to a real aircraft GVT, the method revealed its limits, mainly due to the low force levels delivered by random excitation with the 500 N max force exciters available at that time.

- Method for non-linear behavior estimation
  The idea of the POD (Proper Orthogonal Decomposition) method is to build an orthogonal basis dealing with the best energetic way for describing observed phenomena. PODs obtained from experimental measurements at a high force level are compared to PODs computed from analytical time-series, themselves obtained from measurements at a low force level (see the whole process in Figure 11). The best matching of PODs reveals the likeliest nonlinear model [4].



"Jagellios Blazon" Test Bed
(1) "external" force applied on beam #2
(2) "external" force applied on beam #4
(3) "internal" force applied on beam #1
(4) bungees for vertical and lateral suspensions

Figure 12 - Arm-tree lab model designed for method development and benchmarking

On the laboratory structure (Figure 12) excited by 2 shakers, the nonlinearity is well managed and measured by using a third shaker driven by the structure responses and considered as an internal component of the system. Without any use of the nonlinearity force measurements, the method is successfully applied when the nonlinearity location is known (Figure 13, Figure 14). The method suffers, for today, when multiple nonlinearity sites are involved and if their locations are not known [4].

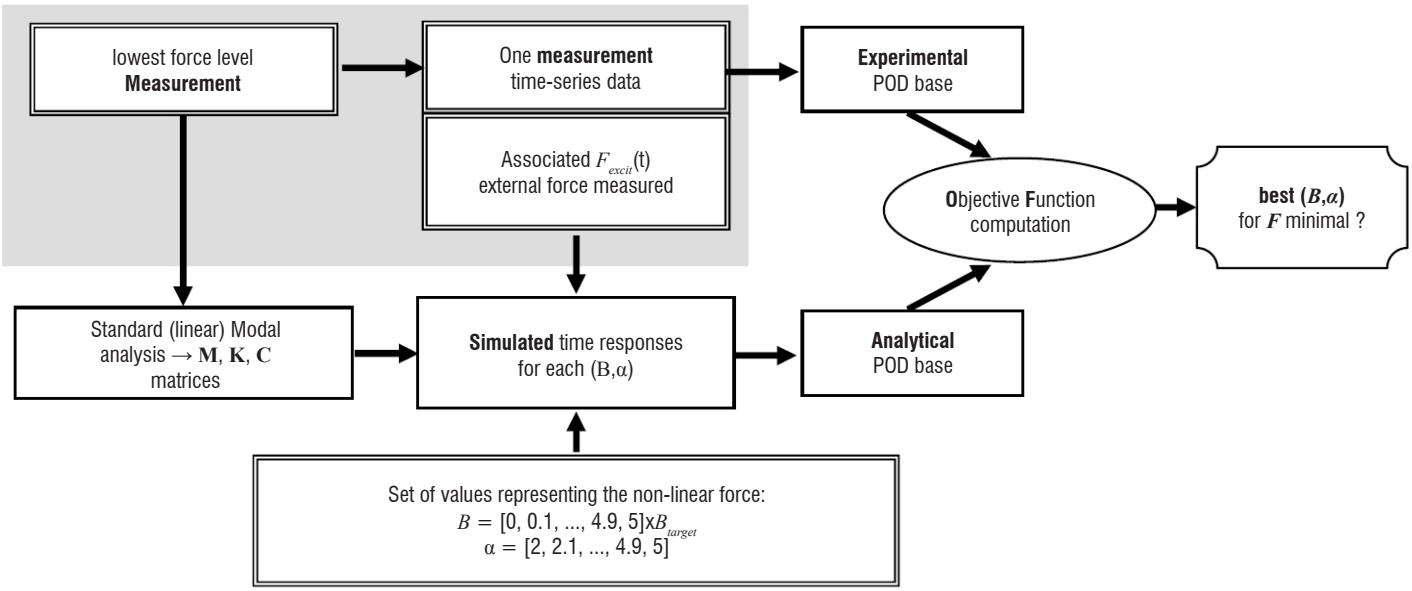


Figure 11 - Process for the identification of nonlinear parameters

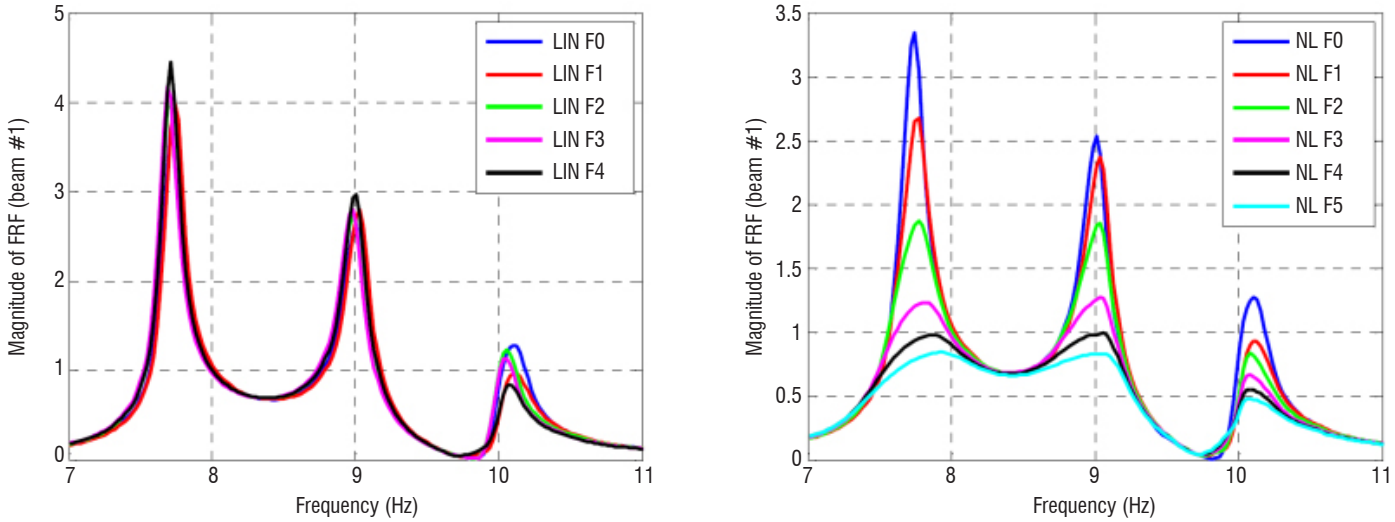


Figure 13 - FRFs (Frequency Response Function) of 1 sensor when a nonlinearity internal force is applied (right side) or not (left side)

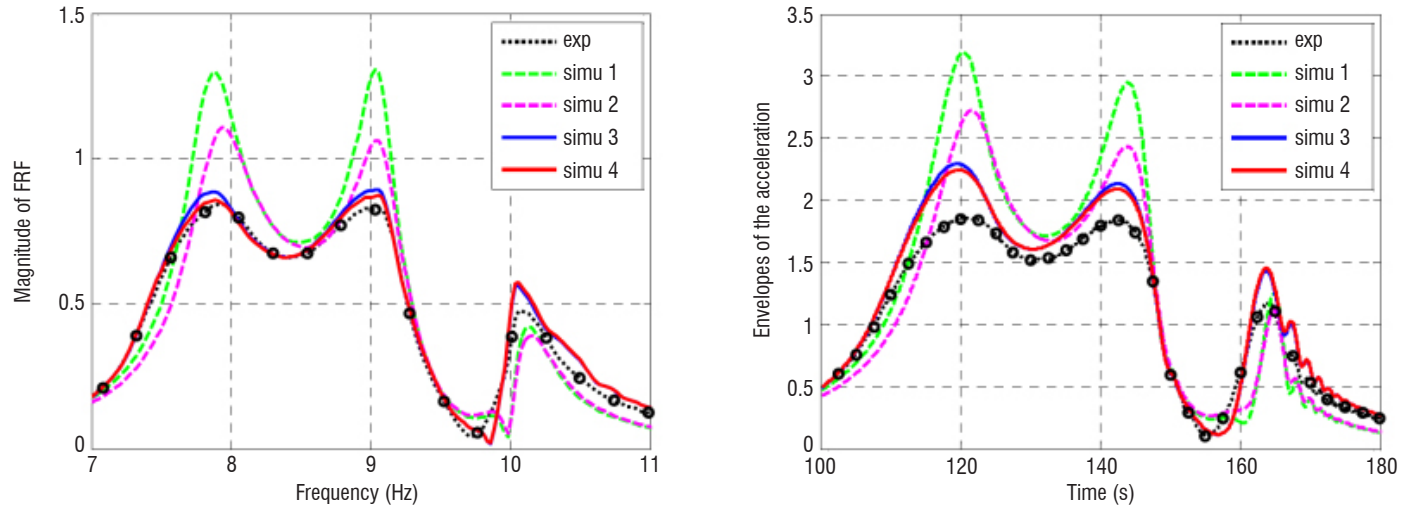
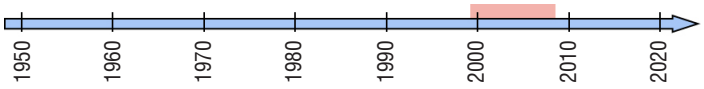


Figure 14 - Predictability: comparison between measured and computed FRFs from nonlinearity POD identification used lower force levels

1999 - 2008: The biggest GVT system in the World, the largest civil aircraft and a new test strategy emerges



1999 was an important milestone in GVT history at ONERA. It corresponds to the complete renewal of the ONERA GVT system, in coordination with SOPEMEA. The main motivations were new industrial requirements (test duration reduction, rapid transmission of final results, maintaining the result quality) and increasing dimensions of modern aircraft. A complete new acquisition system, based on HP VXI 16 bit resolution hardware was selected, able to go up to 1024 synchronous channels, running with SDRC I-DEAS Test software. New ICP TEDS PCB accelerometers and long-stroke 500N PRODERA exciters were specified and purchased to complete the system.

During the last four decades, numerous methods were created to process data coming from Phase Separation (PSM) [7], for instance:

- Time domain methods like LSCE (Least Squares Complex Exponential), Ibrahim Time Domain, etc.
- Frequency domain methods like Polyreference and Direct Parameter.

In practice, frequency domain methods gained attention because they allow the user to focus on a particular frequency band. ONERA started to use market standard EMA methods, like Polyreference (SDRC) and LSCE (LMS). Both algorithms rely on the decomposition of FRFs as a series of rational fractions

$$[H(s)] = \sum_{j=1}^{N} \left(\frac{[R_j]}{s-\lambda_j} + \frac{[R_j]^*}{s-\lambda_j^*} \right) + \frac{[U]}{s^2} + [L]$$

where the $[R_j]$ are the residues and the λ_j are the poles. Matrices $[U],[L]$ denote the upper and lower parts taking into account the influence of out-of-frequency band modes. These parameters are estimated by curve-fitting of experimental FRFs. The quality criterion of this step is to match analytical and experimental FRFs as closely as possible. Finally modal parameters are obtained in a second step. Frequency and damping values are easily computed with poles, while mode shapes estimation is based on the $[R_j]$.

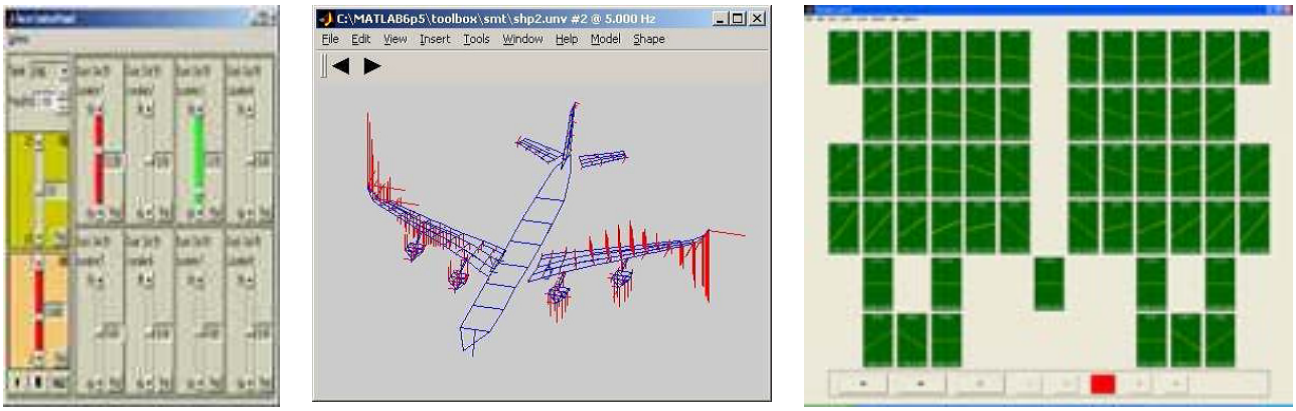


Figure 15 - Force setting interface, 3D mode shape representation and Lissajous displays of the ONERA modal appropriation software

Compared with PRM, these techniques permitted a strong increase in productivity, but without reaching the same levels of excitation. Anyway, throughout this period, they became more and more used in GVTs.

In parallel, based on VXI hardware, ONERA has developed a new PRM software suite and a new force control device, improving the ergonomics of the method. It became a new "standard" for PRM for the next 10 years (Figure 15).

An improved test strategy has also been developed, combining classical PRM with PSM. This strategy, investigated during a research GVT on an Airbus A340-300 in 1999 was applied during the GVT of the Airbus A340-600 prototype in 2001, both performed by ONERA/DLR/SOPEMEA GVT experts [9]. It was the first time that PRM and PSM were combined, reaping the benefit of their individual advantages, to obtain the best and most complete modal model of a tested aircraft in the shortest time (see the detailed test strategy in Figure 16). The use of a FEM in pre-testing, in order to help to define the right number of accelerometers and exciters and their locations, was also an additional way to improve the test efficiency.

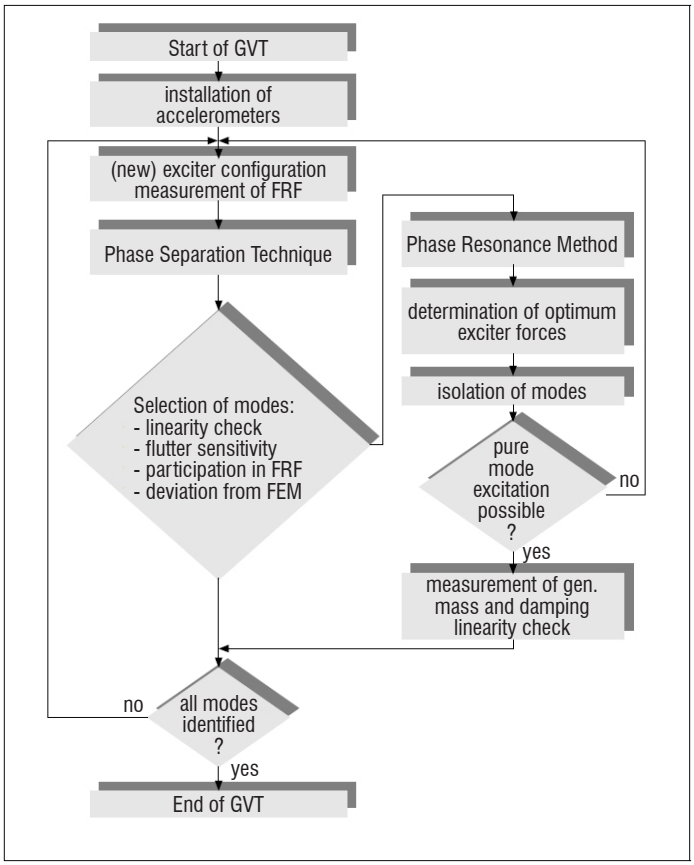


Figure 16 - Optimized test strategy applied during A340-600 GVT in 2001

Still in 2001, the Airbus A340-500 GVT (ONERA/DLR GVT team) was performed. For the first time, it had to be certified to address the post Fan Blade Off wind-milling event. In order to achieve the certification goals with respect to Airbus requirements, ONERA developed excitation stimuli and specific push-pull rods connecting the exciters to the engines [17].

In order to simulate the rotating excitation due to an unbalance, ONERA proposed the use of 2 electrodynamic exciters acting

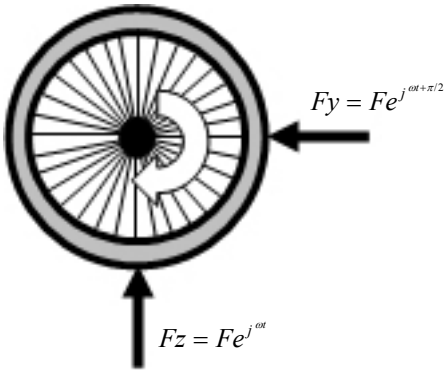


Figure 17 - Circular excitation by means of 2 orthogonal exciters

simultaneously in quadrature with exciters placed with a 90° geometrical deviation (Figure 17).

Excitation signals used for PSM became, most of the time, swept-sines. These swept-sine excitations are either symmetric or antisymmetric force patterns applied with two shakers, generally installed in a symmetric setup. Given that in this case the forces are by definition correlated, it is not possible to use the H_1 estimator directly on the frequency data:

$$\left[H_1(\omega)\right] = \left[P_{ZF}(\omega)\right]^{-1}\left[P_{ZZ}(\omega)\right] \quad (6)$$

Where $[P_{ZZ}(\omega)]$ and $[P_{ZF}(\omega)]$ are respectively the output and input-output power spectral densities.

One solution is to build augmented matrices from the combination of all runs, for instance two runs in the case of symmetric and anti-symmetric excitations. This solution had been adopted until this moment for classical PSM runs with swept-sines, but it could not be used for wind-milling excitations, where Y and Z excitations were correlated, and had to be applied, by principle, simultaneously.

This is why the SVDP (Single Virtual Driving Point) processing method was developed. The SVDP process defines a virtual driving point, which would give rise to vibratory responses strictly similar to those obtained with correlated forces. SVDP relies on the equivalent complex power:

$$CP(\omega) = \sum_{shakers} F_s(\omega)\dot{Z}_s(\omega) = F_V(\omega)\dot{Z}_V(\omega) \quad (7)$$

Where $F_s(\omega)$ is an excitation force acting on a driving point s, $\dot{Z}_s(\omega)$ is the velocity at driving point s, $F_V(\omega)$ is the virtual force and $\dot{Z}_V(\omega)$ is the velocity response of the virtual driving point. The virtual driving point does not exist physically. It is just an imaginary driving point of the virtual force, which has the same excitation energy as the multi-shaker setup and produces the same response.

Once the SVDP process has been applied, SIMO (Single Input Multiple Output) FRFs with regard to the virtual driving point are obtained and classical curve-fitting can be directly used on them. As no hypothesis is necessary about the purity of symmetry of the mode shapes, as from that time the SVDP method has been standardized for swept-sines runs for all subsequent GVTs.

Figure 18 - General view of the A380 in the painting hall during the GVT

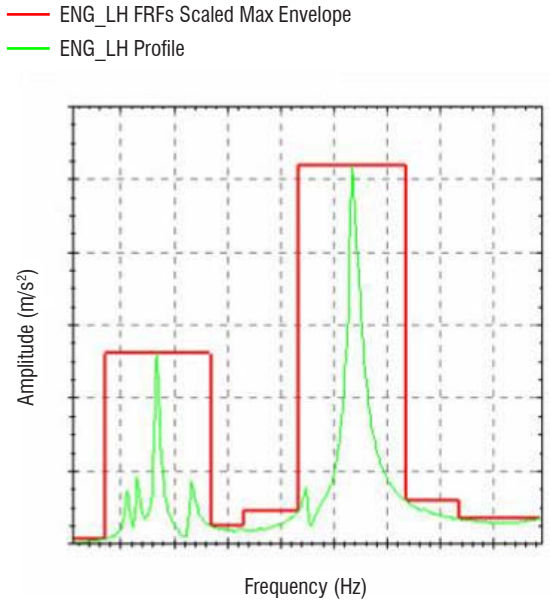
In 2005, the GVT of the Airbus A380-800, the largest civil aircraft ever built, was performed (ONERA/DLR GVT team) [13]. To illustrate the size of this test campaign (Figure 18), we can mention:

- 850 accelerometers
- 50 different excitation locations
- 25 km of cables

Pursuant to the technical requests made by Airbus, the test strategy promoted by ONERA and DLR since 1999 was applied:

- Modal identification method (PSM) based on mathematical curve fitting on measured FRFs for the majority of modes
- Classical modal tuning method (PRM) for modes of special importance only.

The development of specific devices and processing methods contributed to the success of this test campaign. One can recall the improvement of the mode filtering and correlation tools, with various criteria, such as the generalized modal mass, MIF (Mode Indicator Function), participation factor, and MAC (Modal Assurance Criterion) to help us to discriminate and to sort modes. The use of home-made



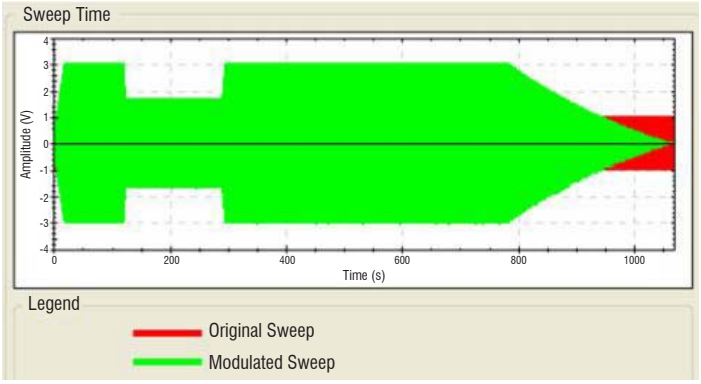


Figure 19 - Maximizing the force level over frequencies and example of a modulated excitation time signal

seismic exciters to excite the two decks in the fuselage can also be recalled. SVDP processing was enhanced and generalized for all PSM runs.

For this test, a force notching process was also developed, in order to maximize the level of the force excitation provided over the frequency band of interest. The classical broadband excitation signals for electrodynamic shakers are random and swept-sine signals. While random signals can be used to achieve a quick insight into the structural dynamic behavior at a very low level of input energy, as the total energy is distributed over the entire frequency range of excitation, swept-sine excitation signals are more appropriate to achieve higher response levels.

In practice, the frequency band is automatically split into several subbands, ensuring a constant excitation force amplitude around vibration modes (Figure 19). From this force pattern computation, an excitation template is generated for the sweep-sine, which maximizes the force level, with respect to limitations (maximum acceleration levels, maximum exciter strokes, and maximum voltage of amplifiers). The resulting excitation signal is a swept-sine whose amplitude is modulated over time (Figure 19).

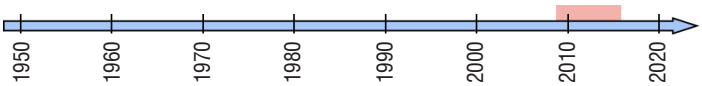
In the context of PSM modal identification, linearity plots, also called impedance plots, have been standardized (Figure 20) to obtain the eigenfrequency and damping coefficient of a mode as a function of the excitation level. This excitation level is given in terms of either a displacement, or the excitation complex power, or a generalized force:

$$P_{Gen} = \sum_{i=1}^{n} p_i u_i / u_{max} \qquad (8)$$

where p_i are the individual forces, u_i are the driving point amplitudes, and u_{max} is the maximum displacement for the target mode.

Figure 20 shows an example of a linearity plot for a control surface mode. Typically, a drop in the eigenfrequency is observed while the excitation level increases and a saturation effect appears at high force and amplitude levels. It is expected to reach this saturation range during the test.

2009 - 2016: A new GVT system, a 21st Century design aircraft and a mature test strategy optimizing productivity



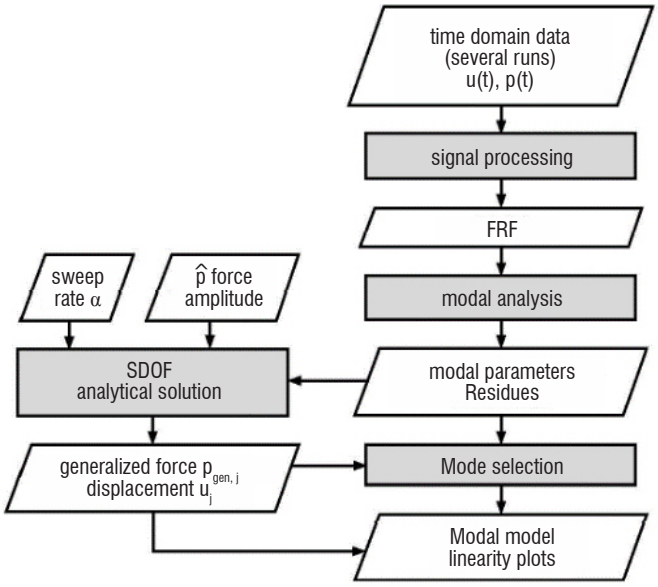
A new GVT system was acquired in 2009, in partnership with the DLR. It occurred with the arrival of new people at ONERA and the DLR to renew and to complete both GVT teams. This new system, able to go up to 768 synchronous channels, is based on LMS SCADAS III 24 bit resolution hardware. The input and output modules are plug and play, the transducer conditioners are embedded in the input modules and, thanks to optical fiber connections between frontends, the architecture can be distributed all around an aircraft. This system runs with LMS Test.Lab and the PolyMAX method is used for PSM modal identification. ONERA and the DLR worked in collaboration with LMS to significantly improve their Normal Modes Testing solution dedicated to PRM.

A research GVT on the first Airbus A340-600 was undertaken in 2011, within a framework founded by the DGAC and LUFO. In this context, new methods and means were developed and new techniques were proposed. One can speak about the use of long-stroke 1000N exciters (dedicated to engine excitation), control surface rotation excitation, taxi vibration testing, fuel sloshing sensor, and new rigid Airbus platforms for exciter support.

On the subject of methods, ONERA developed a new method to optimize sensor placement named ARISPO (Anti-Redundancy Information Sensor Placement Optimization). In the latter, the selection of sensor positions is performed using an algorithm based on the Fisher information matrix, I. This matrix I gathers the weights of possible positions by a sum of modal vector products:

$$I = \sum_{k=1}^{N_s} \Phi_k^T \Phi_k \qquad (9)$$

with Φ_k denoting the k[th] row of the modal matrix, Φ. Removing or adding a sensor directly yields its contribution to the identification process. It is quantified by a matrix norm of I, generally the determinant.



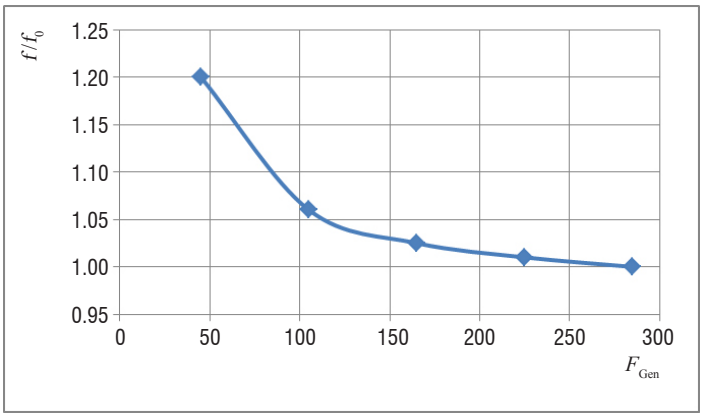


Figure 20 - Computing linearity plots from swept-sine excitations and result example of a control surface rotation mode

In practice, a set of sensors implies about 400-600 accelerometers. In addition, the planning of sensor placement is made through modal vectors computed by the FEM, which can have several thousands of degrees of freedom (each DOF is a candidate for a sensor). Hence, the number of possible combinations is huge, and it would not be possible to test the Fisher information matrix for all of them.

Furthermore, the Fisher information matrix I does not quantify the amount of information shared by two sensors. We proposed in [23] to quantify the redundancy between two potential DOFs by the following formula:

$$R_{kl} = 1 - \frac{\|I_k - I_l\|}{\|I_k + I_l\|} \quad (10)$$

If two potential DOFs k and l have a redundancy R_{kl} close to 1, then they are redundant and only one should be kept. If R_{kl} is close to 0, then each of them contributes its own information. This measure is used as a second criterion to select the most relevant placement for sensors.

Finally, we proposed to use an expansion algorithm in two steps:

- Place the next sensor where it maximizes the Fisher information matrix
- Delete those sensors that are redundant with this sensor

This algorithm is stopped when all sensors have been placed on the structure. An example of sensor placement performed with this algorithm is shown for an A340-600 in Figure 21.

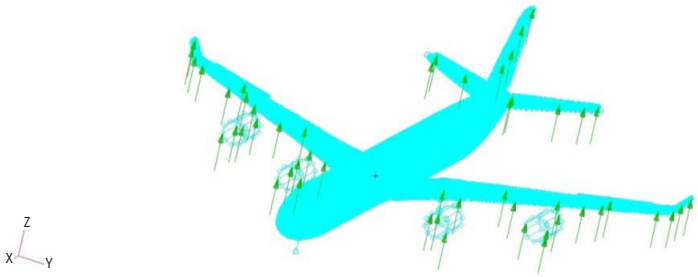


Figure 21 - Example of sensor placement optimization (only Z sensors represented)

In parallel, ONERA also worked on controllability. Two criteria were selected for the optimization of exciter positions. They are aimed at meeting the ONERA quality standards for modal identification:

- A good MIF,
- A significant amplitude.

First, let us define the MIF (Mode Indicator Function). The structural response, Z, for a purely real force vector, F, is given by:

$$Z(\omega) = H(\omega)F(\omega) \quad (11)$$

Where $H(\omega)$ is the frequency response function matrix. Expanding into real and imaginary components and dropping the frequency notation for the sake of clarity, the equation becomes:

$$Z_r + iZ_i = H_r F + iH_i F \quad (12)$$

If a normal mode can be excited at a particular frequency, a force vector F must be found such that the real part Z_r of the response vector is as small as possible compared to the total response. Then, the MIF is computed by:

$$MIF = 1 - Z_r Z^T / \|Z\|^2 \quad (13)$$

The basic idea is to compute the MIF for different exciter set and different force patterns. The force excitation system that yields the best MIF will be considered as the best one. A force excitation is defined by three characteristics:

- The number of exciters
- The force pattern
- The instrumented degrees of freedom (DOF)

For example, the force pattern is a unit force in the case of an exciter, and can be [1 1] or [1 -1] in the case of two exciters.

If two exciter sets can give similar or very close MIFs, it might not be relevant to choose the one among them that gives the most significant value. As our goal is not to fine tune the mode, but rather to excite it under reasonably good conditions, both of them could be suitable. The highest amplitude criterion should enable us to retain the most robust solution between them. It is expected too that it will help to confirm the exciter sets from previous experiences and from mechanical nous.

An example of such an application is presented in Figure 22 for three exciters. The goal was to reveal the fuselage modes of the A340-600. The algorithm proposed two vertical exciters on the HTP (Horizontal Tail Plane) and a lateral one on the nose landing gear.

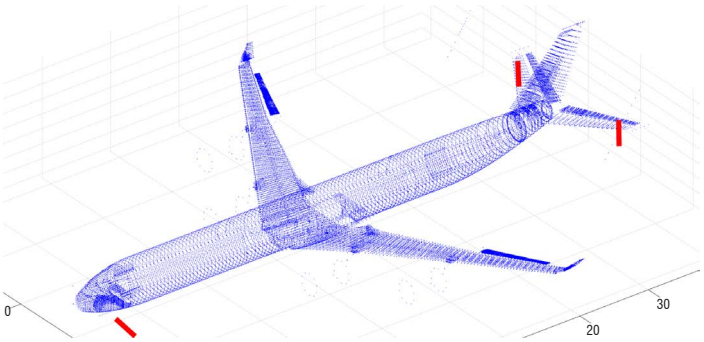


Figure 22 - Exciter placement (red lines) optimized to reveal fuselage modes

Developments were carried out on new excitation stimuli, such as multi-sine sweeps [12], tested during the research GVT on the A340-600. Their benefits have been evaluated in terms of nonlinearity detection, data quality and test productivity. The following combinations were tested:

- Decomposition of the frequency range of a single sweep sine in a combination of two sweep-sine signals running simultaneously in sub-frequency ranges (the complete frequency range covered remains unchanged). The principal objective was to reduce the duration of the acquisition run.

- Cover the same frequency range with two different levels of force (Figure 23). The principal objective was to detect nonlinearities and to have, in the same run, at least two different sets of information (frequency, damping and energy applied) for different modes to build their corresponding impedance curves quicker. Figure 24 shows the results obtained with the SVDP processing for a symmetric excitation. It clearly displays nonlinear behaviors.

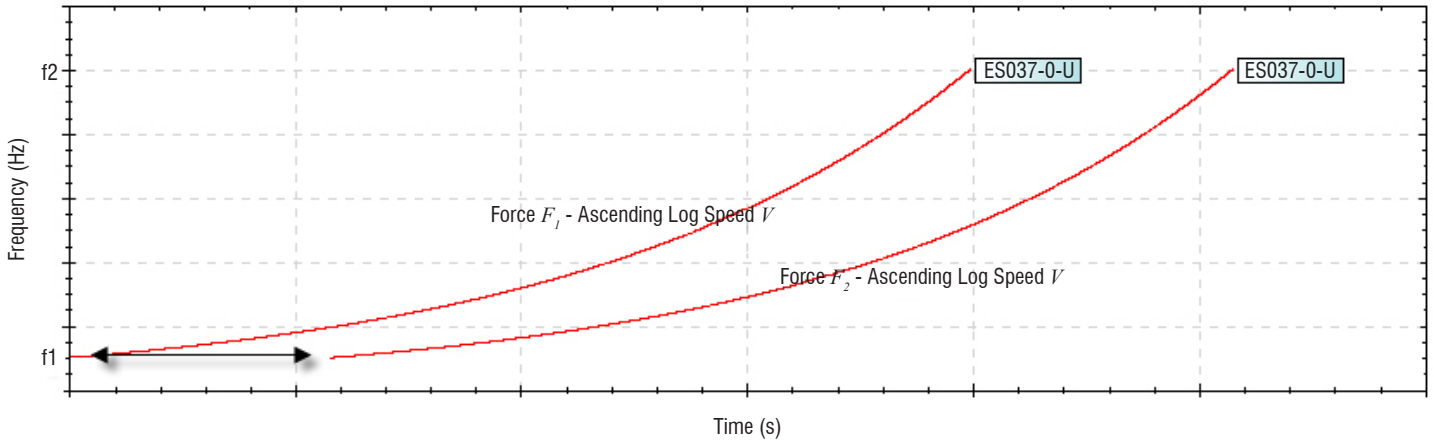


Figure 23 - Sweep sine combination to cover frequency range f_1-f_2 with two different levels of force

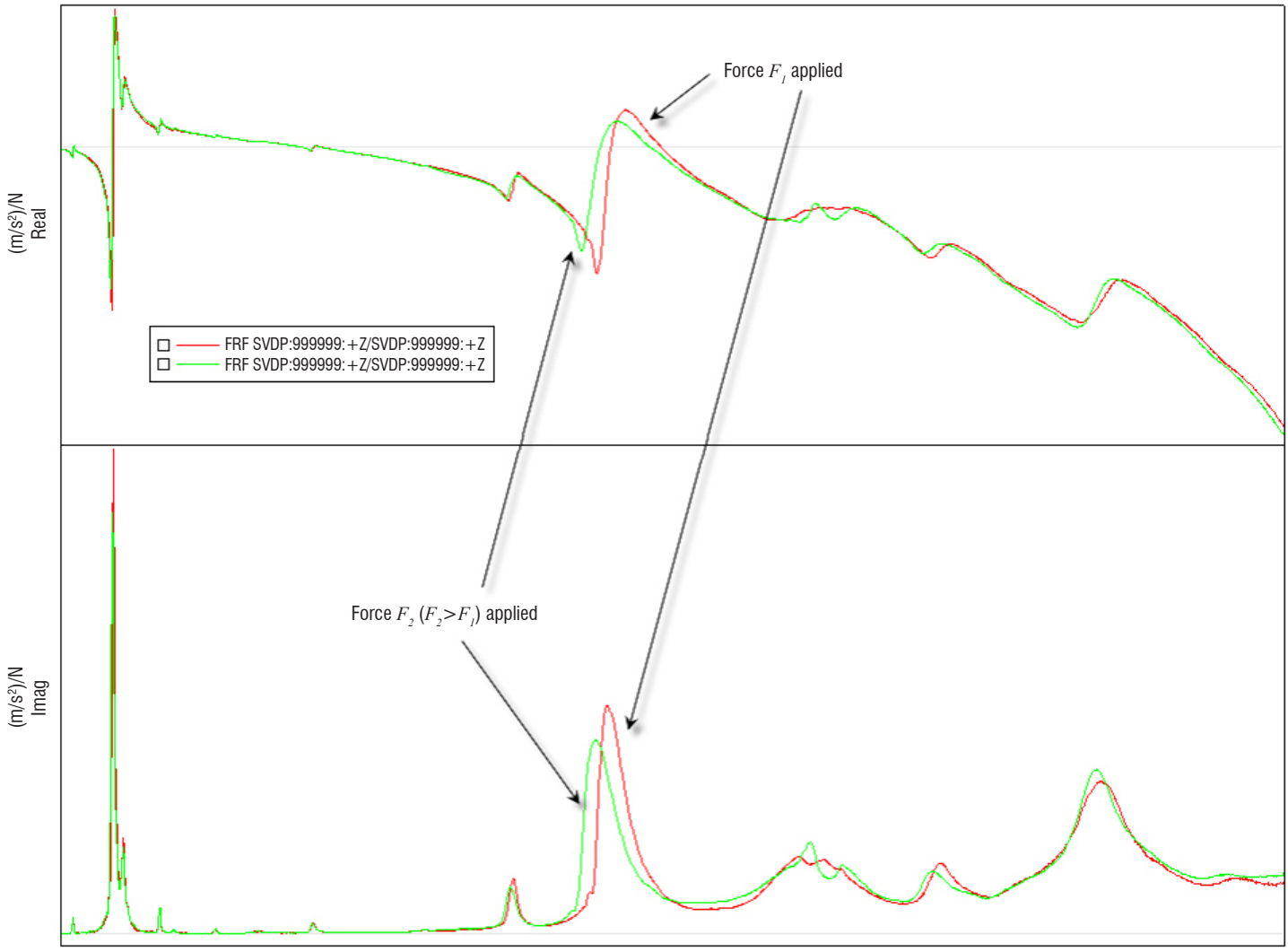


Figure 24 - FRFs of the virtual driving point for each sweep sine of the combination for the symmetric excitation

All of these tools, integrated into the ONERA GVT Tools software suite, are a real benefit to meet the expectations of a modern GVT. The modal model has also been improved, by completing the modal parameter delivery, especially by adding nonlinearity plots for damping.

In 2013, ONERA and DLR performed GVT campaigns for the new Airbus A350 XWB-900 [18] (Figure 25). A first GVT was performed on the first aircraft prototype and a second one on the third prototype exclusively dedicated to the nose landing gear dynamics. The very short time devoted to those test campaigns (9 measurement days for the complete aircraft GVT and 2 measurement days for the nose landing gear GVT) was imposed by a challenging specification from the Airbus A350 XWB FAL (Final Assembly Line), making it necessary to optimize our test strategy.

Figure 25 - GVT campaign of the A350 XWB-900 in Airbus facilities (Toulouse, France)

PSM was used most of the time during the first GVT campaign, since it was the best compromise between time-consumption and mode providing. PRM, known as the most accurate method when non-linear structural behaviors are encountered, was applied for engine modes and for all of the nose landing gear modes of the second GVT campaign.

For speeding up the PSM use, we set up a specific workflow (Figure 26), from excitation signal definition and time data acquisition, to EMA, passing by SVDP and force notching processing. The DLR Correlation Tool was used for data delivery (modal model and non-linearity plots).

In the end, schedules were respected, with 180 modes placed in the final modal model propitious to be used for the FEM updating and flutter computation. Productivity, without any negative impact on quality, was improved (Figure 27 and Figure 28), demonstrating the maturity of our test strategy for GVT campaigns.

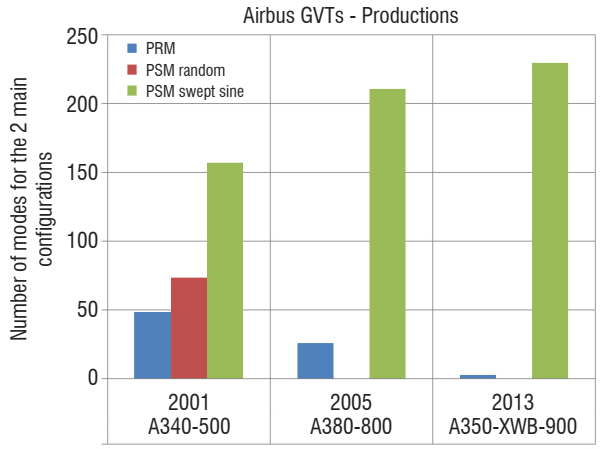


Figure 27 - Diagram of the mode numbers from the different methods for the last major Airbus GVTs

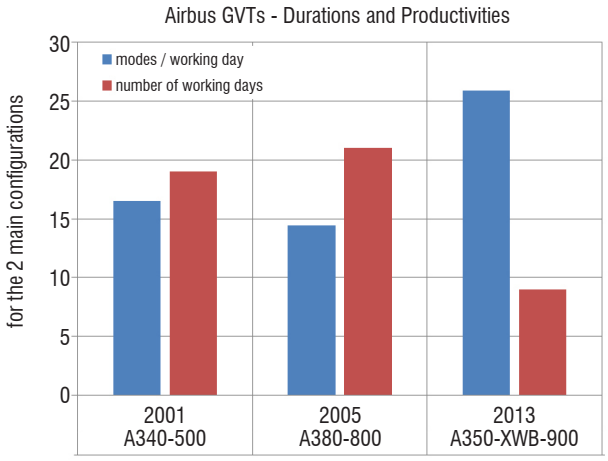


Figure 28 - Diagram of the GVT duration and productivities for the last major Airbus GVTs

In 2014, it was the time for the A320 NEO GVT [25] (Figure 29). The new A320 NEO family incorporates two new engines and sharklets. Here again, schedule constraints (only 7 days dedicated to measurements) and the need for data delivery (mode shapes) quickly after a

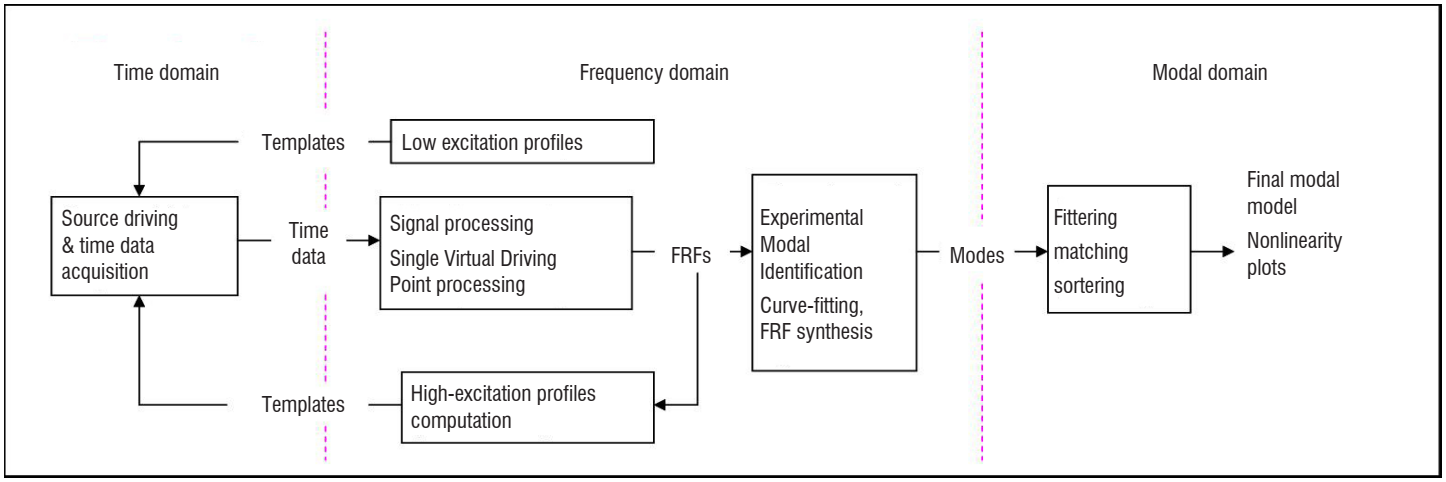


Figure 26 - Data workflow for PSM

measurement (within the following hour) required the improvement of our test techniques and methods.

Figure 29 - GVT campaign on the A320 NEO PW MSN 6101 in Airbus facilities (Toulouse, France)

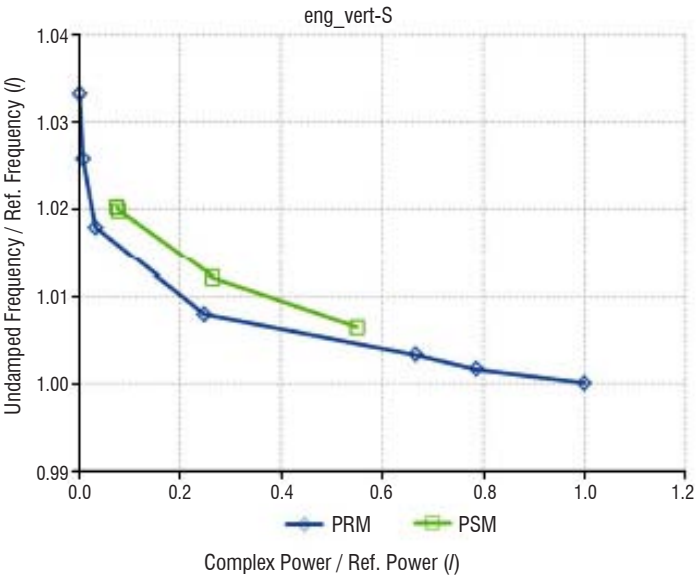
In order to meet these requirements, the DLR Correlation Tool was improved, giving Airbus the capability to access the current modal model online. During the A320 NEO GVT, the task of modal correlation was a specific challenge, and online access enabled the user to do it on-site. Finally, the huge amount of data was condensed down from about 3321 poles identified from all FRF datasets to only 78 master modes in the final modal model for the main configuration.

In addition, ONERA and the DLR shared the work progress table of all individual work stations (exciter preparation, data acquisition, modal identification and modal correlation) in an online multi-user access worksheet.

The data acquisition and processing status could be tracked easily by everybody involved in the GVT, including the customer. This visibility allows instantaneous decisions from the customer to orientate the test.

Finally with regard to this test, we can also add that a MIF per component was provided for each mode and that synergy between PRM and PSM results was enhanced to complete our test process (Figure 30). The A320 NEO GVT was completed in time, despite very challenging specifications and with all expected results delivered with the required quality.

Figure 31 - Photogrammetry on the "Paris" aircraft at ONERA (S1Ch, Meudon)



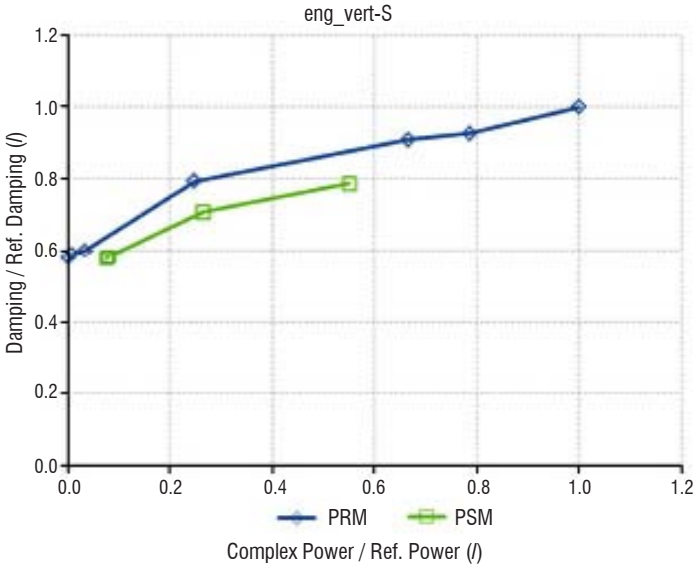
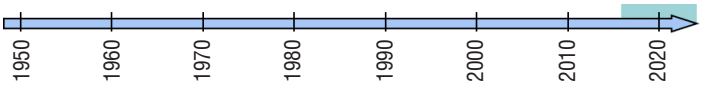


Figure 30 - Example of a comparison between linearity plots for the symmetrical vertical engine mode: PSM and PRM

...and in the future



The ONERA research strategy is closely linked to the current challenges in structural dynamics. This results in:

- A constant enhancement of our tools and means, both software and hardware
  As a matter of technological survey on measurement, we can mention new technologies such as photogrammetry (Figure 31). Recent studies on our Paris aircraft showed promising results but, up to now, far short of the accuracy required at ONERA.

- Data processing: FRF Minimum Variance Method
  In modal analysis, the FRF estimation is generally made through a two-step process: first densities of spectral power (PSDs) are computed using the Welch method (averaged series of periodograms), then these are combined in a least-squares inverse to obtain FRFs. Although this signal processing is now well-mastered, it does not take advantage of the large number of sensors used during a GVT, like spectral methods developed in radar literature. Here we studied the advantages of the Capon spectral method compared to the Welch method [19], [24].
  Contrary to the Welch method, the Capon method estimates all spectra at the same time, not one after the other. In fact, it assumes that, in a set of sensors, noise is uncorrelated among channels. Thus, the more channels there are, the less influenced by noise the FRFs are. Although the difference is not obvious for a set of 5-10 sensors, FRFs are significantly improved for more than 50 sensors. Finally, it is even possible to directly obtain FRFs from signals, skipping the spectra step for computation efficiency, if they are not required for modal analysis.
  An example is presented in the following pictures (Figure 32 and Figure 33) from the GVT of an UAV (Unmanned Aerial Vehicle) named EOLE. 145 sensors and two shakers (one on each rudder) were installed. It can be observed that FRFs are significantly less noisy using the Capon method.

Figure 32 - EOLE UAV GVT (CNES, ONERA, Aviation Design, PERSEUS Student Project www.onera.fr\focus\eole)

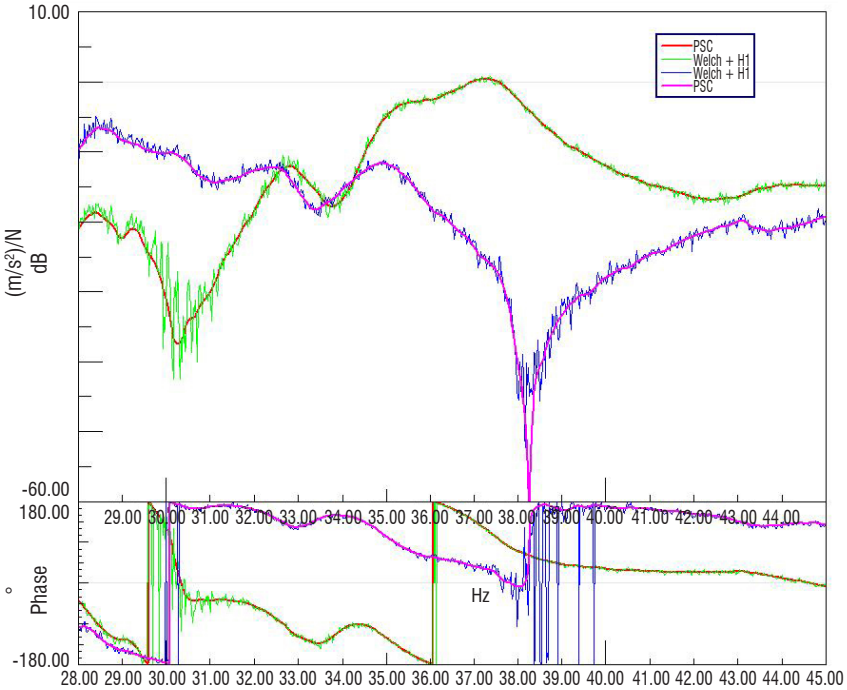


Figure 33 - FRFs computed through Welch signal processing (green and blue curves) and Capon method (red and magenta curves)

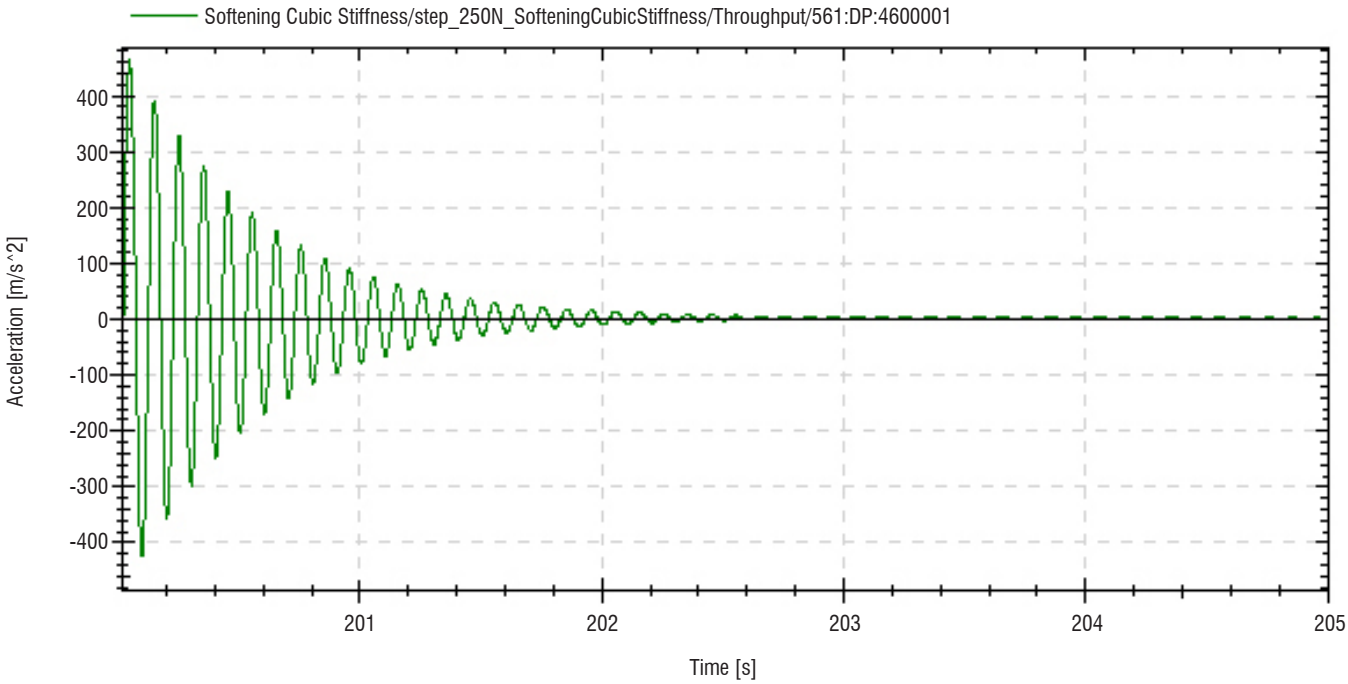


Figure 34 - Time response of a static relaxation

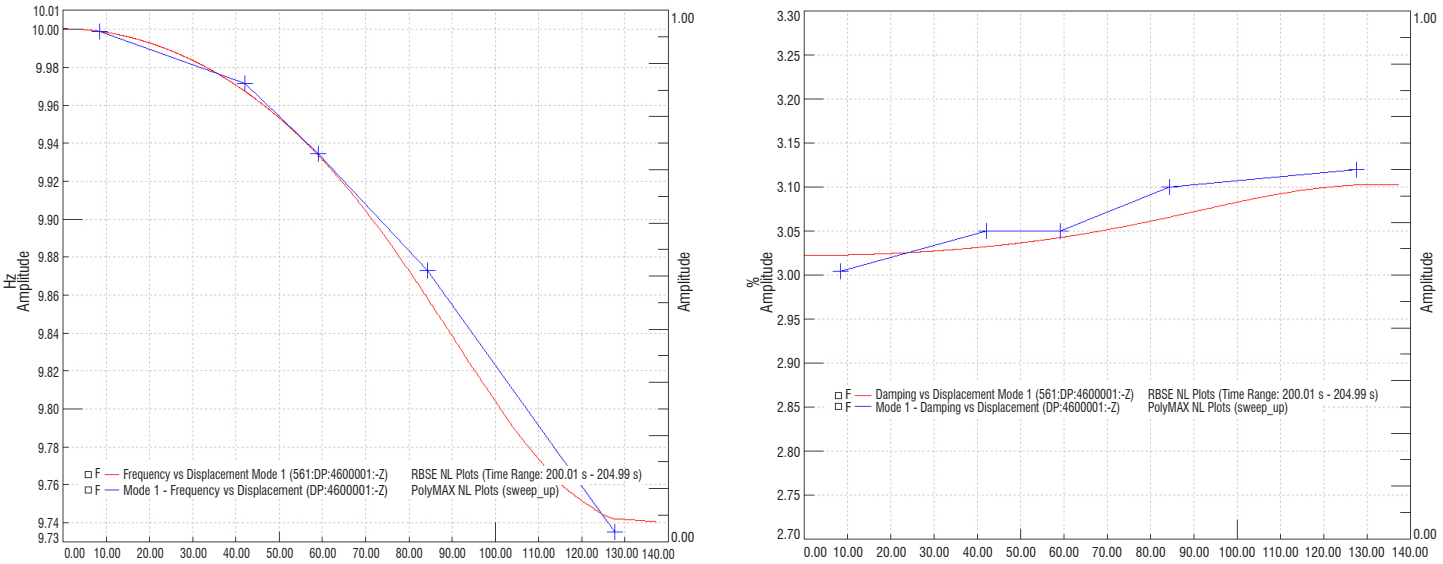


Figure 35 - Nonlinearity plots (frequency and damping) with PSM (cross in blue) and RBSE (line in red)

- Data analysis: Recursive Bayesian-Stephan Estimation
  ONERA is also working on the RBSE (Recursive Bayesian-Stephan Estimation) method. An original method was proposed in [26] for the identification and the tracking of poles of a weakly nonlinear structure from its free responses. Let us consider a structure harmonically excited close to a resonance frequency. Suddenly the excitation is shut down. We suppose that its free response will reveal the whole dependency of its modal parameters (frequency and damping) during its decrease. Then, the idea is to process the measured signals and to extract this dependency from them.

This signal processing method is based on a model of multi-channel damped sines whose parameters evolve over time. Their variations are approximated in discrete time by a nonlinear state space model. States are recursively estimated by a signal process that couples a two-pass Bayesian estimator with an EM (Expectation-Maximization) algorithm. An iterative procedure between them allows an accurate and robust tracking of poles. As a result, equivalent modal parameters such as frequency and damping are obtained as functions of amplitudes.

An application of this method to a nonlinear (cubic stiffness) 1 DOF system is shown below (Figure 34 and Figure 35). We can see that the evolution of modal parameters given by this technique is close to the values obtained by classical curve-fitting of FRFs.

This method has been also applied to a real fan blade specimen. Comparison with results obtained more classically with closed loop control step sine is really satisfying (Figure 36).

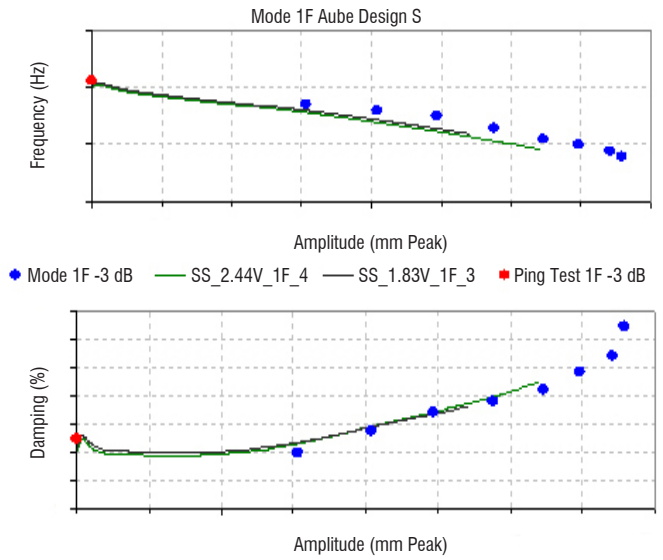


Figure 36 - Linearity plots of a fan blade mode: RBSE (in green), PSM (blue points)

- Nonlinear mode identification
  From the concept of nonlinear modes, new techniques have been proposed by the UL (University of Liège) to compute and to identify invariant manifolds that can gather all of the necessary information for describing the dependency of structures on force levels. Within this framework, ONERA shared the FEM and data tests made on the Paris aircraft with the UL, in order to assess the benefits of nonlinear modes compared to linear ones [16]. Promising results were obtained and work is still in progress to apply them during a GVT.

- Result quality
  Recent developments have been achieved to provide uncertainties on modal parameters from PSM, exhibited between modal synthesis and real measurements. This new piece of information may complement the existing criteria for the selection of the most accurate identification.

Some other research topics should be investigated to apply them in a real industrial context:

- automatic PRM [20],
- combination of external (shakers) and internal (control surfaces) excitation forces,
- productivity and dynamic test program follow up (Real Time Modal Analysis),
- GVT by sub-structuration [3], [14].

With regard to industrial applications, in addition to the long standing and strong cooperation between ONERA and the DLR in performing GVTs for Airbus, in November 2016 ONERA and SOPEMEA signed a partnership to perform GVTs for other aircraft manufacturers, as well as for other sectors outside the aeronautical field.

Conclusion

Throughout its history, Ground Vibration Testing at ONERA has always been strongly related to modal testing and modal analysis, hardware improvements and aircraft innovations. In parallel to the progress made in numerical predictions, several aspects still need to be addressed to render the tests easier, more accurate and more productive. Bearing this in mind, the Morane-Saulnier MS-760 "Paris" aircraft owned by ONERA is a strong advantage for testing new methods and means on an industrial scale.

Quality for identified modes is always linked to the expense of time for testing and data processing. Within the high pressure context of GVT, future studies will be increasingly driven by test purposes, i.e., FEM updating and flutter calculation, in order to balance the involvement of effort and the required accuracy on results. By developing research studies on the one hand, and by enriching the interactions between test suppliers and manufacturers on the other hand, it is the authors' opinion that future GVTs will not only be performed within a still challenging period of time, but will also provide information more focused on the purposes ■

Acknowledgements

The authors would like to thank the DGA (French Armament Procurement Directorate) and the DGAC (French Civil Aviation Authority) for funding research studies and equipment that contributed to making and keeping the ONERA GVT team at the highest world level in this domain.

The authors thank Mr. Marc Rapin, former engineer at ONERA, for his inputs about the history of GVT at ONERA.

Among other past and present partnerships, we especially thank the University of Cincinnati, the FEMTO-ST Institute of Besançon, the University of Liège and, last but not least, the DLR Institute for Aeroelasticity GVT team from Göttingen.

References

[1] E. BALMES, J. WRIGHT - *GARTEUR Group on Ground Vibration Testing. Results from the Test of a Single Structure by 12 Laboratories in Europe*. IMAC, Orlando USA, 1997

[2] E. BONNEAU - *Determination of the Vibration Characteristics of a Structure from the Expression of the Complex Power Supplied*. La Recherche Aérospatiale 130 - 1969

[3] M. BÖSWALD, D. GÖGE, P. LUBRINA, B. SCHULZE - *Recent and Future Improvements in Ground Vibration Testing of Large Aircraft*. International Forum on Aeroelasticity and Structural Dynamics, Stockholm, 2007

[4] S. DEQUAND, P. LUBRINA - *Experimental Nonlinear System: Application of the Proper Orthogonal Decomposition Method*. IMAC-XXVI - Conference and Exposition on Structural Dynamics, Orlando, 2008

[5] G. DE VRIES - *Analyse des réponses d'une structure mécanique dans l'essai global de vibrations*. Symposium IUTAM - Paris - 1965

[6] G. DE VRIES - *Remarques sur l'analyse des courbes d'admittance d'une structure mécanique*. la Recherche Aérospatiale n°95 - 1963

[7] D. EWINS - *Modal Testing: Theory, Practice and Application, Baldock*. Publisher: Research Studies Press, 2000
[8] P. FARGETTE, P.-M. HUTIN, B. LEVADOUX - *Modal Identification Testing of Weakly Nonlinear Structure. Application of a Stochastic Linearization Method*. 2nd DLR-ONERA Aerospace Symposium, Berlin, 2000
[9] P. FARGETTE, U. FÜLLEKRUG, G. GLOTH, B. LEVADOUX, P. LUBRINA, H. SCHAAK, M. SINAPIUS - *Tasks for Improvement in Ground Vibration Testing of Large Aircraft*. International Forum on Aeroelasticity and Structural Dynamics, Madrid, 2001
[10] O. FILATRE - *Identification of Weakly Non-linear Dynamical Systems by Means of Random Excitations (in French).* Thesis Ecole Centrale Paris, Malabry, 1991
[11] I.E. GARRICK, W.H. REED III - *Historical Development of Aircraft Flutter*. Nasa Langley Research Center, Hampton, VA, AIAA 81-0491R, 1981
[12] S. GICLAIS, P. LUBRINA, C. STEPHAN, M. BÖSWALD, Y. GOVERS, J. UFER, N. BOTARGUES - *New Excitation Signals for Aircraft Ground Vibration Testing*. International Forum on Aeroelasticity and Structural Dynamics, Paris, 2011
[13] G. GLOTH, P. LUBRINA - *Ground Vibration Test of the Airbus A380-800*. International Forum on Aeroelasticity and Structural Dynamics, Munich, 2005
[14] D. GÖGE, M. BÖSWALD, U. FÜLLEKRUG, P. LUBRINA - *Ground Vibration Testing of Large Aircraft – State-of-the-Art and Future Perspectives*. IMAC-XXV - Conference and Exposition on Structural Dynamics, Orlando, 2007
[15] A. GRAVELLE, M. LEPART, P. LUBRINA - *Aircraft Ground Vibration Test by Means of Flight Control Surfaces*. International Forum on Aeroelasticity and Structural Dynamics, Aachen, 1991
[16] G. KERSCHEN, M. PEETERS, J.C. GOLINVAL, C. STEPHAN - *Nonlinear Modes of a Full-Scale Aircraft*. AIAA Journal of Aircraft, vol. 50, p.1409-1419, 2013
[17] P. LUBRINA - *Ground Vibration Experiments on Large Civil Aircraft for Engine Imbalance Purpose.* International Forum on Aeroelasticity and Structural Dynamics, Amsterdam, 2003
[18] P. LUBRINA, S. GICLAIS, C. STEPHAN, M. BÖSWALD, Y. GOVERS, N. BOTARGUES - *Airbus A350 XWB GVT: State-of-the-Art Techniques to Perform a Faster and Better GVT Campaign*. IMAC-XXXII - Conference and Exposition on Structural Dynamics, Orlando, 2014
[19] S. L. MARPLE Jr - *Digital Spectral Analysis with Applications*. Prentice-Hall Signal Processing Series, 1987
[20] L. RATSIFANDRIHANA - *Amélioration des procédures d'identification modale des structures par appropriation automatique et utilisation de forces non contrôlées*. Thèse de Doctorat, Université de Franche-Comté, 23 mai 1995
[21] C. SOIZE - *Stochastic Linearization Method with Random Parameters for SDOF Nonlinear Dynamical Systems: Prediction and Identification Procedures*. Probabilistic Engineering Mechanics, 10 (3), p.143-152, 1995
[22] C. SOIZE, O. LEFUR - *Modal Identification of Weakly Non-Linear Multidimensional Dynamical Systems Using a Stochastic Linearization Method with Random Coefficients*. Mechanical Systems and Signal Processing, 11(1), p.37-49, 1997
[23] C. STEPHAN - *Sensor Placement for Modal Identification*. Mechanical Systems and Signal Processing, 27, 461-470. 2012
[24] C. STEPHAN - *Evaluation of Filter Banks Spectral Estimators for Modal Analysis (in French)*. Third Seminar "Experimental Vibration Analysis", Blois, 2012
[25] C. STEPHAN, T-P. VO-HOANG, S. GICLAIS, Y. GOVERS, P. LUBRINA, M. BOESWALD, A. LAPORTE - *Airbus A320 NEO GVT and FEM Updating: State-of-the-art Techniques to Perform an Industrial Vibration Test Campaign and a Rapid Process to Update Renewed FEM for Clearance of First Flight Test*. International Forum on Aeroelasticity and Structural Dynamics, St Petersburg, 2014
[26] C. STEPHAN, H. FESTJENS, F. RENAUD, J.-L. DION - *Poles Tracking of Weakly Nonlinear Structures using a Bayesian Smoothing Method*. Mechanical Systems and Signal Processing, doi:10.1016/j.ymssp.2015.05.028

Acronyms

ARISPO	(Anti-Redundancy Information Sensors Placement Optimization)
CTW	(Wasserburg Technical Center (*Centre Technique de Wasserburg*))
DGA	(French Armament Procurement Directorate (*Délégation Générale à l'Armement*))
DGAC	(French Civil Aviation Authority (*Délégation Générale de l'Aviation Civile*))
DLR	(German Aerospace Center (*Deutsches Zentrum für Luft- und Raumfahrt*))
EM	(Expectation-Maximization)
EMA	(Experimental Modal Analysis)
FAL	(Final Assembly Line)
FEM	(Finite Element Model)
FFT	(Fast Fourier Transform)
FRF	(Frequency Response Function)
FQ	(Force in quadrature)
FRF	(Frequency Response Function)
GARTEUR	(Group for Aeronautical Research and Technology in EURope)
GVT	(Ground Vibration Testing)
HP	(Hewlett-Packard)
HTP	(Horizontal Tail Plane)
ICP	(Integrated Circuit Piezoelectric)
LMA	(Applied Mechanics Laboratory (*Laboratoire de Mécanique Appliquée*))
LMS	(Leuven Measurement Systems)
LSCE	(Least-squares Complex Exponential)
LUFO	(German Federal Research in Aeronautics (*Luftfahrtforschung*))
MAC	(Modal Assurance Criteria)
MIF	(Mode Indicator Function)

NEO	(New Engine Option)
ONERA	(The French Aerospace Lab (*Office National d'Etudes et de Recherches Aérospatiales*))
POD	(Proper Orthogonal Decomposition)
PRM	(Phase Resonance Method (also called Modal Appropriation, Modal Tuning, Normal Mode Testing))
PSM	(Phase Separation Method (also called Global Method))
RBSE	(Recursive Bayesian-Stephan Estimation)
SIMO	(Single Input Multiple Output)
SDRC	(Structural Dynamics Research Corporation)
SDRL	(Structural Dynamics Research Lab)
SVDP	(Single Virtual Driving Point)
TEDS	(Transducer Electronic Data Sheet)
UAV	(Unmanned Aerial Vehicle)
XWB	(eXtra Wide Body)

Nomenclature

Z	Displacement at one location of a structure
$\dot{Z}$	Velocity at one location of a structure
$\ddot{Z}$	Acceleration at one location of a structure
$[M]$	Structural mass matrix
$[C]$	Structural damping matrix
$[K]$	Structural stiffness matrix
$[C_{aero}]$	Aerodynamic damping matrix
$[K_{aero}]$	Aerodynamic stiffness matrix
V	Velocity of a structure in a fluid
F	Excitation force
$f\left(Z,\dot{Z}\right)$	Restoring force
$[H_1(\omega)]$	H_1 estimator FRF matrix
$[P_{ZZ}(\omega)]$	Output power spectral density matrix
$[P_{ZF}(\omega)]$	Input-output power spectral density matrix
CP	Complex power
P_{Gen}	Generalized force
p_i	Individual force at excitation point i
u_i	Individual amplitude at excitation point i
u_{max}	Maximum displacement amplitude
I	Fisher information matrix
Φ_k	kth row of the modal matrix
R_{kl}	Redundancy between DOFs k and l
PRM	Phase Resonance Method (also called Modal Appropriation, Modal Tuning, Normal Mode Testing)
PSM	Phase Separation Method (also called Global Method)

AUTHORS

Stéphane Giclais, Supméca (Institut supérieur de mécanique de Paris) engineer, Stéphane Giclais joined ONERA (DADS/ADSE) in 2010 as project leader in aircraft GVT. Prior to joining ONERA, he has worked as a Project Engineer for Brüel & Kjaer for 6 years, on acoustic and vibration applications for automotive and aerospace industries, and as an Application Engineer, specialized in structural dynamics solutions, for SDRC and MTS Systems for 7 years.

Pascal Lubrina, INSA Lyon Mechanical Construction Engineer, project leader in aircraft GVT at ONERA since 1982.

Cyrille Stephan, after having graduated from the ENSMM (Ecole Nationale Supérieure en Mécanique et Microtechniques) in 2004, Cyrille Stephan got his PhD in Sciences for Engineering at the University of Franche-Comté in 2009. He then joined ONERA and is charge of developments of methods for structural dynamics, especially on signal processing and nonlinearities.

L. Jacquin, V. Brion, P. Molton, D. Sipp, J. Dandois, S. Deck, F. Sartor, E. Coustols
(ONERA)

D. Caruana
(ISAE)

E-mail: laurent.jacquin@onera.fr

DOI: 10.12762/2016.AL12-06

Testing in Aerodynamics Research at ONERA: the Example of the Transonic Buffet

The paper reviews research conducted at ONERA over the last thirty years on the transonic buffet. We first present the transonic buffet phenomenon and we explain its importance for aeronautical applications. Then, a distinction is made between the 2D buffet produced by an airfoil and the 3D buffet that characterizes swept wings of finite span. The 2D buffet amounts to a pure oscillation of the shock phase-locked with the detachment and reattachment of the boundary layer downstream, whereas the 3D buffet takes the form of a pocket of broadband perturbations located in a limited portion of the wing. We recall that these mechanisms were first studied in the 1980s through a series of tests conducted in the transonic wind tunnel ONERA T2 at Toulouse and in the large transonic wind tunnel ONERA S2Ma at Modane. Since this pioneering work, progress in the measurement techniques has led to the constitution of a comprehensive database of the 2D buffet that we describe. This database, obtained in the wind tunnel ONERA S3Ch at Meudon, has been extensively used to validate various CFD tools, with the latter being used in turn to investigate the buffet physics. We illustrate this collaboration between simulation and physics by recalling that a linear stability analysis of accurate Reynolds-Averaged-Navier-Stokes (RANS) solutions made it possible to prove that the buffet on a 2D airfoil stems from a global instability mechanism. We also review more recent tests done in the case of a laminar airfoil, which reveal very distinct behaviors of the buffet flow. This illustrates how sensitive the buffet is to the nature of the boundary layer. The last section of the paper gives a short overview of advanced simulations for these different test cases. In the conclusion, we list research perspectives, which include some more general topics such as data assimilation.

Introduction

In fundamental science, testing is often aimed at confirming, improving, or disproving the predictions of a given theory[1]. In aerodynamics, there is no new theory to be tested a priori, the basic equations (Navier-Stokes) having been established for a long time and still providing satisfactory results. Nevertheless, this science faces a fundamental obstacle, which is turbulence. Turbulence is an unavoidable consequence of the non-linear structure of the fluid mechanics equations, which produces a cascade of flow scales down to the limits of the continuum, constantly challenging the experiment, metrology and computation, see [2]. Since the 1980s, the advent of Computational Fluid Dynamics (CFD) has deeply transformed aerodynamic engineering, which now relies heavily on computations. There is a consensus today that a critical area in this field is the ability to adequately predict viscous turbulent flows with the possible presence of a boundary layer transition and flow separation [3]. As CFD progresses, experiments become scarce and more comprehensive. However, they remain unavoidable for exploring the borders of the parameter space that they previously secured. Thus, testing will long remain essential to CFD, the success of which still depends on the best compromise between efficient software, physical knowledge and reliable measurement availability. The case of the transonic buffet described in this paper provides a good illustration of the collaboration between these various research domains, which has shaped thirty years of aerodynamics in an institution like ONERA.

[1] At the time of the 70th anniversary of ONERA, an emblematic example is the satellite MICROSCOPE launched by Soyuz on April 29th 2016 from the French Guiana Space Center, and which contains some world-class ONERA technologies. This satellite is aimed at improving the confirmation of the equivalence between inertial mass and gravitational mass, an important verification of Einstein's relativistic theory of gravitation, see [1].

Transonic buffet

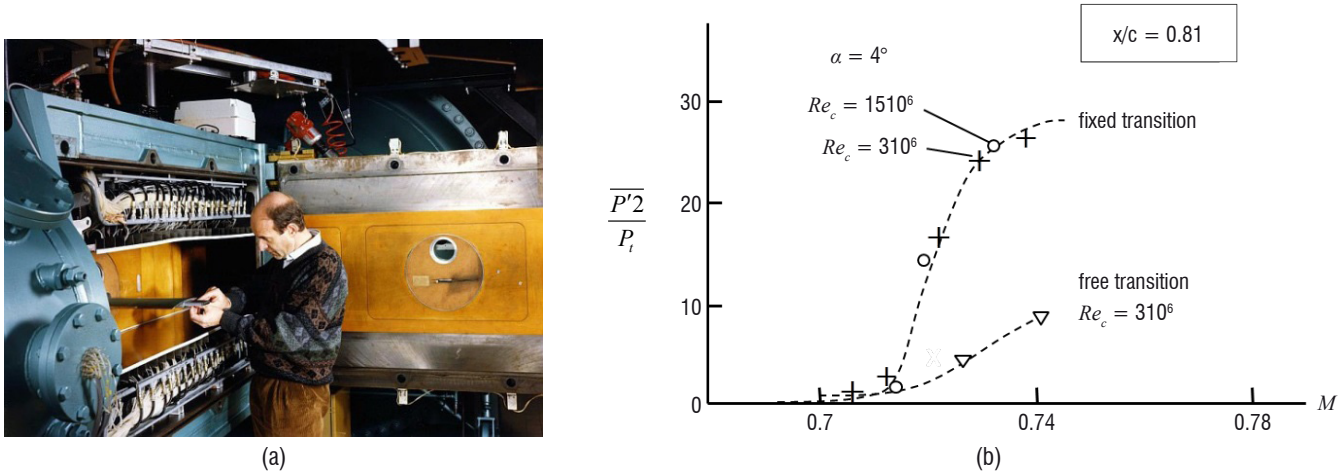
When seeking a subject that could illustrate the role of testing in aerospace research in this 70th anniversary volume launched by ONERA, the transonic buffet came as a natural choice. This is because this topic illustrates well how aerodynamics has progressed over the last twenty years, thanks to advances in flow diagnostic techniques and fruitful cooperation between theory and numerical simulation. It is also relevant to question the future of wind tunnels. Buffet refers to an aerodynamic excitation leading to unsteady forces. Transonic buffet refers to cases where a shock wave is involved; it is characterized by self-sustained displacements of the shock wave location and periodic boundary-layer separation downstream from the shock. This occurs, for instance, on the upper surface of a highly loaded airfoil, i.e., when the flow Mach number and the airfoil angle of attack are both sufficiently large. As far as aeronautical applications are concerned, transonic buffet is one of the most important compressibility-based problems that limit the load capacity and efficiency of cruising aircraft. It can develop in all flow regions where shock waves are present, i.e., on airframe lifting surfaces and on rotating blades. Its control could enlarge flight envelopes and lead to significant energy savings. Conducting relevant research in that domain calls for controlling complex experiments in high-speed wind tunnels, but it also requires mastering advanced metrologies, advanced fluid mechanics software and cutting-edge theoretical approaches in compressible flow physics. Significant efforts have been made by the aeronautical community on the topic, especially at ONERA, which devoted both fundamental and applied type research in every relevant domain to this problem, such as flow stability, unsteady CFD and experiments in transonic/supersonic facilities. This article is aimed at presenting a synthetic view of the research conducted at ONERA in this field since the mid-1980s. The first set of described test cases concerned a 2D supercritical airfoil. We start with cases where the 2D transonic buffet develops under purely turbulent conditions, which is naturally the case on real aircraft but which necessitate a transition tripping of the airfoil boundary layers tested in wind tunnels. The buffet flow in that case takes the form of a single and global flow oscillation, of which the onset and structure turned out to be predictable by linear hydrodynamic stability, as detailed below. We then consider the case where the boundary layer is laminar, a subject of current investigations and an important issue for the topic of laminar wing aircraft. Recent results, which confirm pioneering work conducted in the 1980s, reveal that the 2D laminar buffet is less severe than the turbulent one. Finally, turbulent buffet on 3D airfoils have been also documented several times at ONERA. In this case, the buffet develops in a limited portion of the wing. The physics of this buffet is complex and remains largely a matter of questioning.

Pioneering work

Pioneering experimental studies were conducted since the mid-1980s by the Aeroelasticity and Structural Dynamics Department and by the Aerodynamics and Energetics Models Department of ONERA. The first team studied a transonic airfoil named RA16SC1 in the ONERA S3Ma wind tunnel [4] from 1984 to 1987. In the same period, the second team launched a series of experiments at ONERA-Toulouse in the Cryogenic Induction Tunnel ONERA T2 on two models, the RA16SC1 and the OAT15A airfoils, see [5]. This team also coordinated a campaign on the 3D buffet on a swept wing in 2000 in the S2Ma wind tunnel at ONERA-Modane, see a review in [7]. Steady and unsteady pressure transducers on the model surfaces were used to characterize these different flows. Part of these experiments did not focus particularly on the flow physics, but rather on the control of the buffet by open-loop or closed-loop strategies. Also, the instrumentation used at that time was too modest to appraise the physics with enough detail, or for the building up of a sufficiently comprehensive database for fine CFD-experiment confrontations. Nonetheless, nearly all of the key elements regarding the buffeting phenomenon were already available at the end of these early works, and almost all of the bases of the subsequent work done at ONERA on the topic were developed at that time.

Airfoil buffet

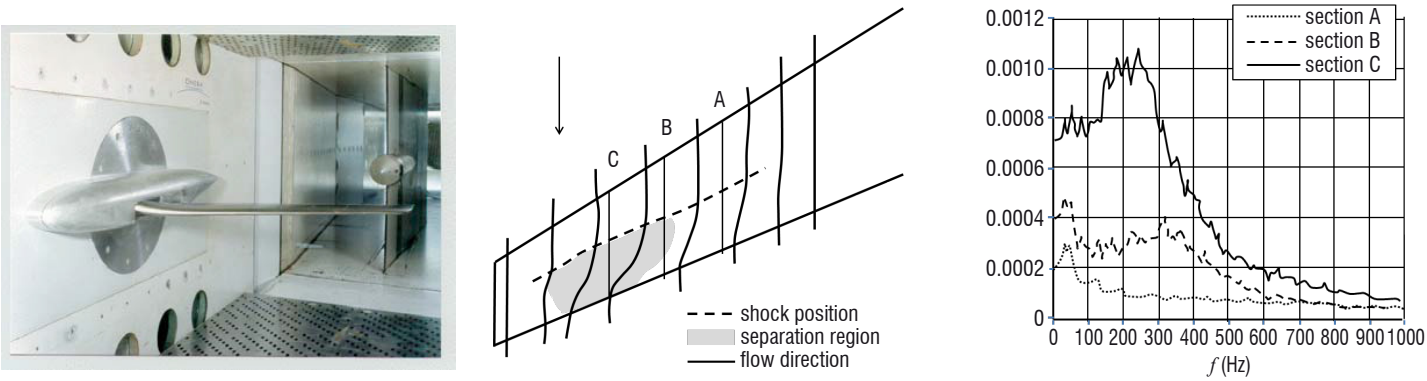
A first example of those pioneering contributions is shown in Figure 1 which was extracted from an experiment on the 2D transonic buffet conducted in the ONERA T2 wind tunnel in the mid 1980's [5]. "T2" was a transonic wind tunnel with a test section that was 0.39 m wide and 0.37 m tall, equipped with 2D adaptive walls, see Figure 2(a). The tunnel was cryogenic and pressurized, operating within a total temperature range from 300 K to 100 K and within a total pressure range from 1.5 bars to 4 bars[2]. The model was a supercritical OAT15A of 0.15 m chord length, designed for cryogenic flows and equipped



(a) (b)

Figure 1 – (a) Airfoil buffet experiment on the OAT15A airfoil in the cryogenic induction tunnel T2 at ONERA-DMAE, (b) pressure fluctuation variance on the downstream upper side ($x/c = 0.81$) of the airfoil for an angle of attack $\alpha = 4°$: effects of the Mach number, the Reynolds number and the boundary layer state from [5].

[2] The T2 tunnel closed in the 2000s due to lack of funding and activity.



(a) (b) (c)

Figure 2 – (a) 3D buffet experiments on a 3D model in the S2MA tunnel. Case $M_0 = 0.82$, $\alpha = 3.7°$, $Re_{cm} = 8.3\ 10^6$: (b) mean surface flow organization with separation area, (c) spectrum of the shock position deduced from unsteady pressure measurements in three sections of the wing, [7].

with a single transducer to measure the unsteady wall pressure on the downstream part of the upper side, at 81% of the chord. Figure 1(b) shows the variations with the Mach number of the pressure-fluctuation variance for a fixed angle of attack equal to 4°. Two different states of the upper boundary layer flow were considered: a natural transition case and a forced transition one obtained by adding a tripping device on the upstream suction side. In the first case, the boundary layer is assumed to be laminar (at least down to the shock location); in the second case, the boundary layer is turbulent. Even though no specific diagnostic was applied to prove these assertions, they look reasonable. The figure shows that when the boundary layer is turbulent, the transonic buffet is characterized by a sudden increase of the pressure fluctuations, here around $M = 0.72$ for this particular airfoil and angle of attack ($\alpha = 4°$). Analyses of the corresponding signal have shown that the pressure fluctuations are characterized by a single frequency close to 100 Hz, see [5]. However, Figure 1(b) also shows two other important results. The first concerns the negligible influence of a variation by a factor five of the Reynolds number on the turbulent boundary layer cases. This suggests that the underlying dynamics is essentially fixed by the mean turbulent flow. It took two decades to confirm that the onset of the transonic buffet actually results from a single global instability mode of the mean turbulent flow, see below. The second interesting result found in Figure 1 is the absence of any clear buffet when a natural transition is considered. As shown also later on, this radical change in the flow properties in the case of a laminar boundary layer is the subject of active research today. Concerning CFD, the methods used at that time to compute such flows were unsteady viscous-inviscid interaction methods, see [6].

3D buffet

3D buffet refers to flows over transonic swept wings. As illustrated below, the 3D buffet strongly differs from the 2D airfoil buffet. As mentioned in the introduction, early studies were devoted to this problem at ONERA in the early 2000s (see a review in [7]). They were based on the use of a 3D half wing/fuselage body similar to a civil transport aircraft. At that time, the model was designed with the help of viscous-inviscid coupling computations. Its principal dimensions are 1.25 m spanwise and between 0.25 and 0.45 m chordwise. The model was equipped with 6 accelerometers, 60 steady and 103 unsteady pressure transducers, and the global efforts were measured by means of a 6-component wall balance. It was tested in 2000 in the ONERA S2MA facility. "S2Ma" is a transonic or supersonic, pressurized closed circuit wind tunnel with a test section that is 1.77 m tall, 1.75 m wide and 3.75 m long, equipped with fixed top and bottom perforated walls, see Figure 2(a). The tests were performed with a mean aerodynamic chord Reynolds number of $Re_{cm} = 8.3\ 10^6$ and free stream Mach numbers between 0.80 and 0.84. Figure 2(b) and (c) illustrate the main properties found for the 3D buffet flow on the upper side of the model wing, here for $M_0 = 0.82$ and $\alpha = 3.7°$. This flow differs from that of a 2D airfoil in two main points: (i) the buffet develops in a restricted portion of the wing, here around 75% of the span, where a finite area separated region appears; this leads to a corrugation of the mean shock in that region, see Figure 10(a); (ii) the dynamics of this buffet is no longer modal but rather broadband. In the case of Figure 2(c), energetic frequencies arise at around 200 Hz. Once normalized by the mean chord length, this corresponds to a Strouhal number $St = fc_m/U_\infty = 0.24$, which is nearly three times larger than that obtained on a 2D airfoil. More detailed investigations have shown that at that Mach number the shock location became displaced over less than 5% of the chord in Section C identified in Figure 2(b) when the buffet was established. Consequently, in comparison with the 2D buffet phenomenon, the shock location is less displaced for 3D buffeting.

A second campaign was conducted ten years later, in 2010, using the same model and the same facility. As reported in [8], thanks to a deeper pressure signal analysis some progress has been made on the physics of the 3D buffet, which consists in the combination of the streamwise oscillation of the shock location, like the 2D buffet with the phenomenon of "stall cells" (named "buffet cells" by Iovnovich *et al.* [9]). These "stall cells" are clearly evidenced by the presence of a spanwise wavelength at the shock location. On a swept wing, these stall cells are convected towards the wing tip. The 3D buffet frequency increases with the sweep angle from St = 0.07 in 2D to St = 0.26 for a 30° sweep.

The physics of the turbulent airfoil buffet

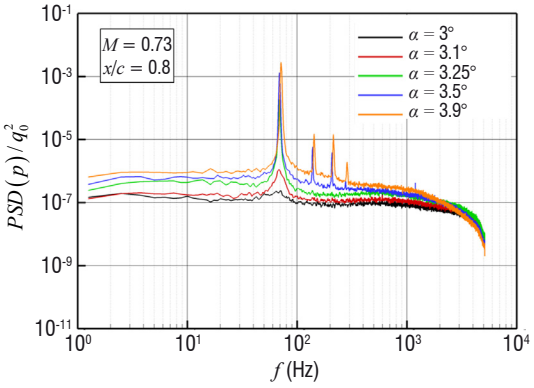
We now focus on the period since the early 2000s, when detailed experiments, including flow characterization by means of advanced optical methods, were made possible on a 2D airfoil and where CFD had made significant progress.

Mean flow and dynamics

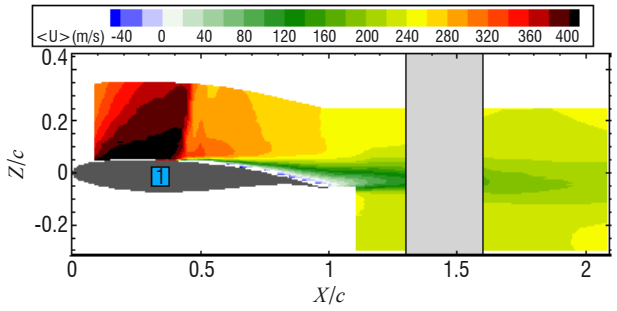
A fully documented experiment on the topic was conducted at ONERA in the mid-2000s in the continuous closed-circuit transonic wind tunnel S3Ch of the Fundamental and Experimental Aerodynamics Department at Meudon, [10]. The test section of the S3Ch wind tunnel is 0.76 m tall and 0.8 m wide equipped with 2D adaptive top and bottom walls like the T2 wind tunnel in Toulouse, see Figure 3(a). As in the previous work conducted in the T2 wind tunnel, a tripped supercritical airfoil OAT15A was again considered for studying turbulent regimes. Here, the model had a chord length of 0.23 m and a 0.8 m width (actually the width of the wind tunnel). This gives $AR \approx 3.4$ for the aspect ratio. This was sufficient to avoid any 3D effects due to interactions of the airfoil flow with the boundary layers of the wind tunnel side walls, which were a source of concern during the T2 experiments for which $AR \approx 2.6$. In the S3Ch experiment, surface flow visualizations proved that the flow was perfectly 2D over a large portion of the model [10]. The average chord-based Reynolds number was $Re_c = 3 10^6$ with laminar-turbulent transition fixed on the airfoil by means of a Carborundum strip located at $x/c = 7\%$ from the leading edge. The Mach number M_0 was varied between 0.70 and 0.75 and the flow incidence α, controlled by means of deformable walls, could be varied between 2.5° to 3.9°. The measurements comprised surface flow visualizations by oil and sublimating products, steady and unsteady pressure (68 pressure taps and 36 unsteady Kulite™ pressure transducers in the central section of the wing), flow-field characterizations by Schlieren films, and velocity fields by means of a two-component laser-Doppler velocimeter. As revealed by the pressure spectra of Figure 3(b), measured at $x/c = 0.8$ for $M = 0.73$, the onset of the buffet occurs for an angle of attack AOA = 3.2° and it is characterized by flow oscillations that induce very energetic and harmonic pressure oscillations. These oscillations are felt over the entire range of transducers on the upper surface, from the shock foot down to the trailing edge, revealing the global character of the phenomenon. The frequency of the phenomenon is $f \approx 70$ Hz giving a chord-based Strouhal number $St = fc/U_\infty = 0.068$. Similar low frequencies were obtained previously in other airfoil buffet flows: a value of $St \approx 0.081$ was found by Lee [11] for a supercritical airfoil, and McDevitt and Okuno [12] found $St \approx 0.045$ for a NACA0012 airfoil. This variability in the transonic buffet Strouhal number indicates that the chord is not the only length scale involved; others, such as the mean shock position, the boundary layer thicknesses or the shock interaction region sizes can be also considered, but no consensus has yet emerged today. The value $\alpha = 3.5°$ was selected for a full characterization of the unsteady velocity field with Laser-Doppler velocimetry. A phase-averaged technique has been applied to this data. The technique can separate the "coherent" motion, related to a periodic excitation (possibly a global mode, see below), from the random fluctuating part.

A component of the velocity, $u(\underline{x},t)$ for instance, is decomposed into three contributions, $u(\underline{x},t) = \overline{u}(\underline{x}) + \tilde{u}(\underline{x},t) + u''(\underline{x},t)$, where $\overline{u}(\underline{x})$ is the temporal-average, $\tilde{u}(\underline{x},t)$ is the cyclic component and $u''(\underline{x},t)$ is a random fluctuating component. The phase-averaged velocity is defined as $\langle u(\underline{x},t)\rangle = \overline{u}(\underline{x}) + \tilde{u}(\underline{x},t)$, the remaining fluctuating component u'' being a residue characterizing events that are not in phase with the reference signal. In the present case, the reference signal has been chosen as the pressure signal measured by the Kulite™ transducer located close to the mean shock location. Phase averages of the longitudinal and vertical components of the velocity (u,v) were then determined following a procedure described in [10]. The structure of the phase-averaged longitudinal velocity flow $\langle u(\underline{x},t)\rangle = \overline{u}(\underline{x}) + \tilde{u}(\underline{x},t)$ is illustrated in Figure 3(c) and (d), which shows two phases (among the 20 covering the buffeting flow period) corresponding to the extreme locations of the oscillating shock, respectively. The oscillation of the shock phase-locked with the oscillating of the shear layer downstream is shown.

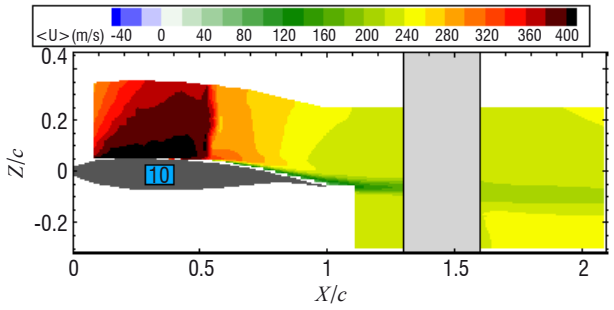
(a)



(b)



(c)

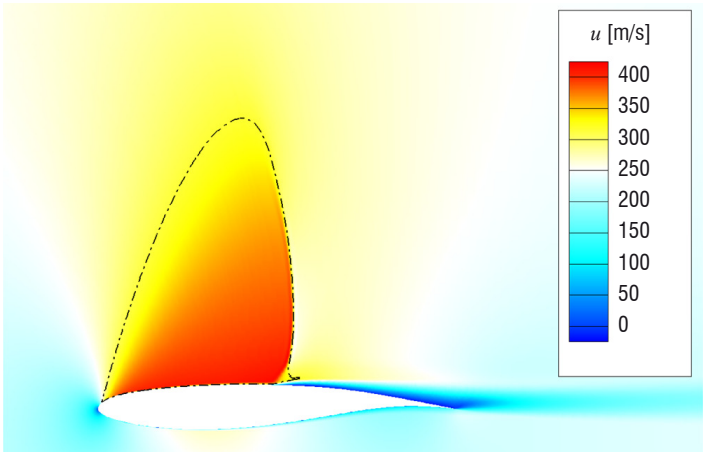


(d)

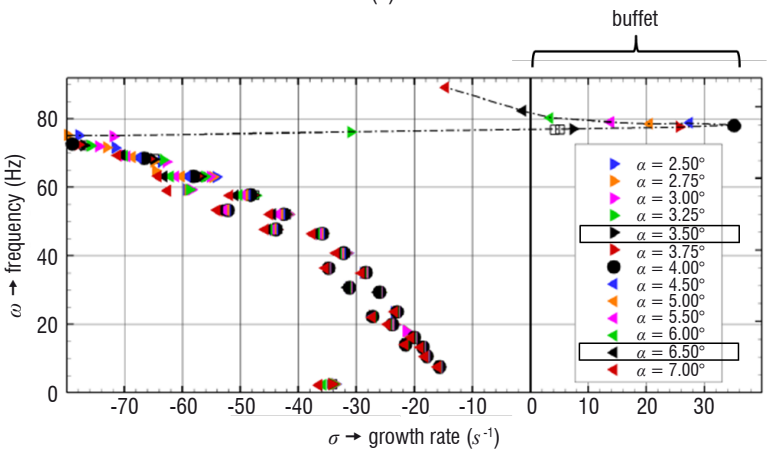
Figure 3 – (a) Airfoil buffet experiment on the OAT15A airfoil in the S3Ch wind tunnel at ONERA-DAFE, (b) Pressure power spectra at, $x/c = 0.8$ for different angles of attack, (c) LDV phase-averaged longitudinal velocity $\langle u(\underline{x},t)\rangle = \overline{u}(\underline{x}) + \tilde{u}(\underline{x},t)$ (see text) with shock upstream (phase 1/20), (d) with shock downstream (phase 10/20) (from [10]).

A global instability mechanism

As shown by Crouch *et al.* [13], both the onset conditions and the spatial and temporal structure of the buffet flow on 2D airfoils appear to be well described by a global-stability approach. This milestone, at the crossroads of theoretical fluid mechanics and computational sciences, was achieved thanks to the rising of CFD in the 2000s. A short synthesis of the accomplishments of ONERA in this field applied to transonic buffet will be made later. The theoretical framework of the analysis of Reference [13] were the unsteady Reynolds Averaged Navier-Stokes equations (URANS). The main ingredients of this theory are the following. Consider the Navier-Stokes equations governing the flow dynamics in the form $\partial_t q = R(q)$, where the state vector $q = (\rho, \underline{u}, T, \nu)^T$ represents the aerodynamics (i.e., the density, the velocity field, the internal energy and the viscosity fields). Let q_0 denote a base flow, meaning an equilibrium point such that $R(q_0) = 0$, and let q' denote superimposed small perturbations; the latter are governed by $\partial_t q' = A(q')$, where A is the Jacobian operator linked to the R by the relation $A = \partial R / \partial q_{q=q_0}$. The perturbation is then sought in the form of normal modes $q'(\underline{x}, t) = \hat{q}(\underline{x}) \exp\{(\sigma + i\omega)t\}$, where σ and ω represent the temporal growth rate and the frequency of the global mode $\hat{q}$, respectively. When $\sigma > 0$ (resp. <0), the base flow is unstable (resp. stable). Substituting q' leads to an eigenvalue problem for $\lambda = \sigma + i\,\omega$, which is written as $A\hat{q} = \lambda \hat{q}$. When the Reynolds number is large, the flow is fully unsteady and it is usually turbulent. A choice must be made for the base flow q_0. After splitting the total flow q into a steady-state flow $\overline{q}$ and an unsteady perturbation q', an approximation of q_0 by $\overline{q}$ can be proposed. The steady-state flow $\overline{q}$ is a solution of the Reynolds Averaged Navier-Stokes equations (RANS) $RANS(\overline{q}) = 0$. This calls for a turbulent model to close additional turbulent flux terms, *i.e.*, the averaged Reynolds stresses. Using an eddy viscosity formulation for this model, the global stability analysis described above can finally be applied to the state vector $q = \overline{q} + q'$ with $\overline{q} = (\overline{\rho}, \overline{u}, \overline{T}, \overline{\tilde{\nu}})^T$ and $q' = (\rho', \underline{u}', T', \tilde{\nu}')^T$, where $\tilde{\nu} = \overline{\tilde{\nu}} + \tilde{\nu}'$ is the eddy viscosity. The ability of an eddy-viscosity formulation of the turbulence model to reproduce such an unsteady separated flow as close as possible to the test case was proven by Thierry and Coustols [14]. Crouch *et al.* [13] used a combination of the global mode analysis of the RANS equations with the Spalart-Allmaras turbulence formulation of the eddy viscosity (RANS-SA) and successfully tested the method with the NACA0012 airfoil database of McDevitt and Okuno [12]. The periodic behavior of the flow illustrated in Figure 3 suggests that a transonic buffet flow should be dominated by a single normal mode. This was confirmed by the theory, further applications to the ONERA-OAT15A airfoil having led to the same conclusion, see Sartor *et al.* [14]. Some of the latter authors' findings are reproduced in Figure 4. Figure 4(a) shows the base flow horizontal velocity field, a solution of the RANS-SA equations for $M_0 = 0.73$ and $\alpha = 3.5°$, and Figure 6(b) shows the eigenvalue spectra obtained for various angles of attack. The theory predicts a destabilization of a single mode of frequency $f = 77$ Hz for an angle slightly larger than $\alpha = 3.25°$, against $f \approx 70$ Hz and $\alpha = 3.1°$ in the experiment. Interestingly, when α is increased further (up to very large angles, such as 7°),



(a)



(b)

Figure 4 – Buffet at Mach number $M = 0.73$ (tripped supercritical airfoil OAT15A) – Global stability analysis: (a) RANS-SA horizontal velocity field for $\alpha = 3.5°$, (b) eigenvalue spectra for various angles of attack, [15].

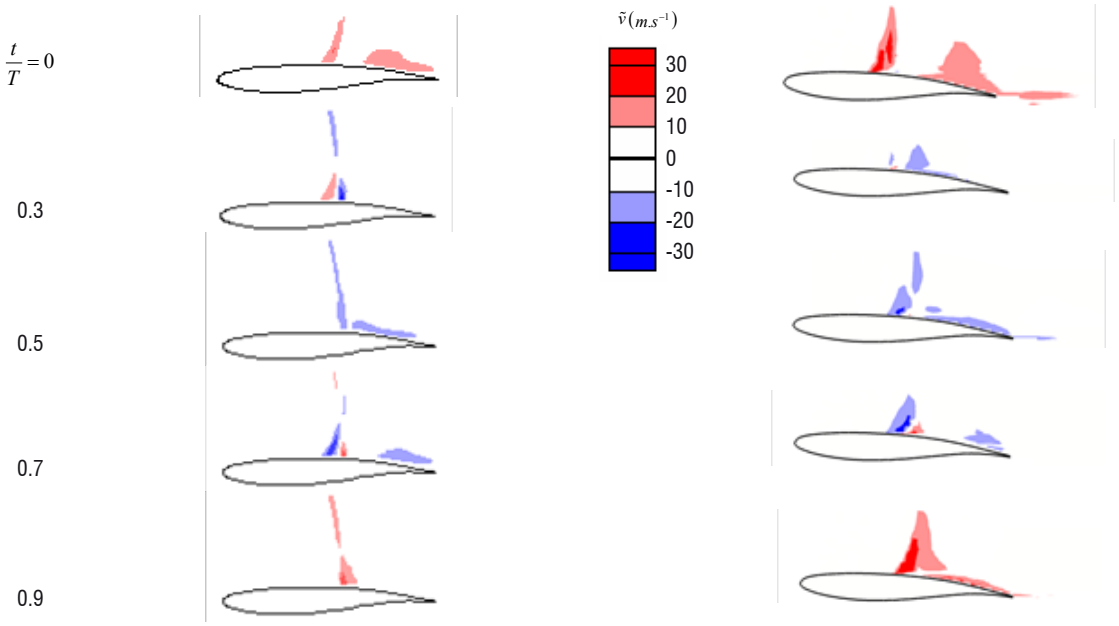


Figure 5 - Buffet at $(M_0,\alpha)=(0.73,3.5°)$ (tripped supercritical airfoil OAT15A) - Unsteady vertical velocity fluctuations. Left: computation (global mode), right: experiment (cyclic component of the vertical velocity $\tilde{v}(\underline{x},t)$), [16].

the fully-detached flow obtained is no longer buffeting; this suggests that a periodic reattachment of the downstream flow is a key feature to obtain a buffet flow.

In the framework of the global stability theory recalled above, when the dynamics are dominated by a single global mode the latter should correspond to the cyclic component $\tilde{q}$ of a triple decomposition. This was emphasized by Crouch *et al.* [15] by comparing the structure of the global mode q' of the RANS-SA flow with the cyclic component $\tilde{q}$ of the OAT15A flow described in Figure 3(c) and (d). Figure 5 shows the vertical-velocity fluctuations using contours chosen to allow a direct comparison of the two flow-fields.

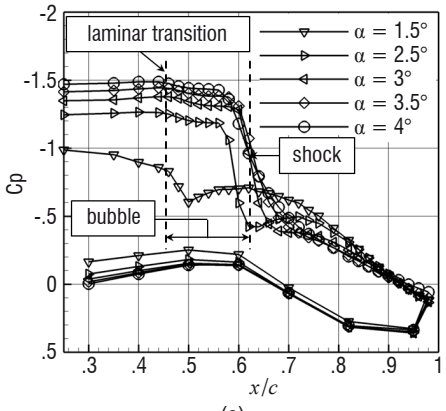
Laminar airfoil buffet

Laminar versus turbulent mean flows

The experiments described above were reproduced recently in the same facility (ONERA S3Ch), using an OALT25 airfoil designed to promote a laminar boundary layer ahead of the shock wave in the absence of tripping, see [17]. The chord length and the span of the OALT25 airfoil are identical to those of the OAT15A, their relative thicknesses being very close (12.18% compared to 12.3%). The Mach number M_0 was varied between 0.70 and 0.75 and the flow incidence α between 1.5° and 4°. The same experimental protocols and techniques were used, LDV being replaced by PIV, with the addition of thermographic diagnostics obtained by using an Infra-Red (IR) camera to check that the upstream laminar flow was 2D and free from any spurious transition up to the shock. Given that the adaptive walls prevent a direct view from above, the IR camera is installed outside the test section and provides a limited view through the side windows of approximately a quarter of the span. Figure 6(b) shows an IR photo, where the change from dark to light grey indicates a transition. A transition cone caused by an added small roughness is also visible. The photo shows that the transition is 2D (at least in the observed region) and that it is caused by the shock. Figure 6(c) then shows the pressure coefficient C_p-distributions provided by pressure taps located in the central section

(a)

(b)



(c)

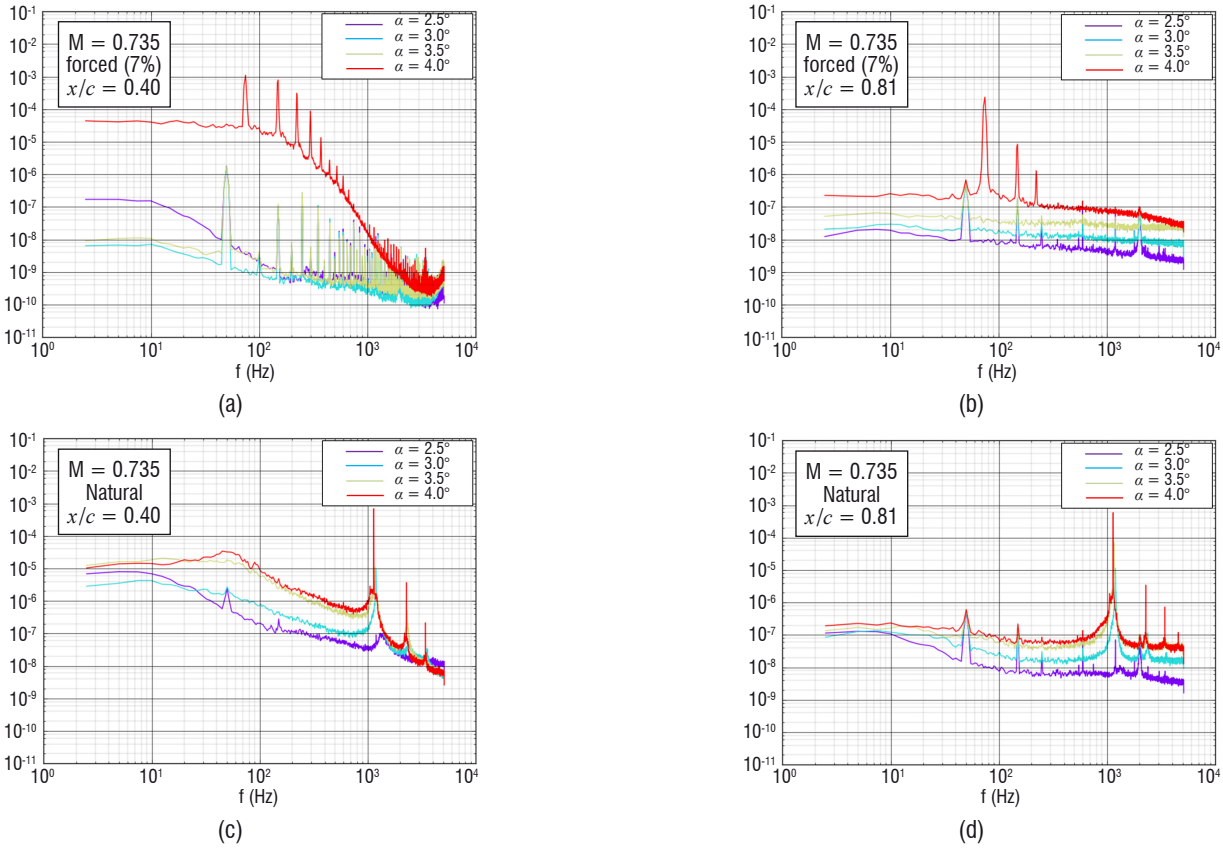
Figure 6 – (a) Airfoil buffet experiment on the supercritical OALT15A airfoil without stripping (natural transition) in the S3Ch wind tunnel at ONERA-DAFE, (b) IR image for $M_0 = 0.73$ (see text), (c) Mean pressure coefficient, [17].

of the wing at 25% $\leq x/c \leq$ 100% for $M_0 = 0.73$ and for different angles of attack. It shows that a shock forms for and settles at a position of approximately 60% of the chord length when α is increased. It can be noted that the C_p curves of Figure 6(b) are marked by a slight compression ahead of the shock, at about 40% of the chord. This feature is absent in the case where the boundary layer is turbulent, as in the OAT15A data commented earlier, as well as in the present experiments when transition was forced. The most likely reason for this is the existence of a laminar separation bubble under the shock foot. This bubble forms as soon as the compression occurs, as a result of the little resilience of a laminar boundary layer to an adverse pressure gradient. The presence of a laminar bubble was confirmed with the help of numerical RANS results, where the laminar-turbulent transition in the boundary layer was accounted for by means of a dedicated criterion, [17]. These experiments reveal that the dynamics of the laminar airfoil buffet flow differ radically from those of a turbulent airfoil. First, the spreading of the C_p - recompression region due to shock unsteadiness in Figure 6(b) shows that, in the laminar regime, the shock oscillation amplitude does not exceed 5% of the chord length against 20% in the case of the turbulent OAT15A airfoil, see Figure 3(c)-(d).

Importantly, this suggests that the laminar buffet is less severe than the turbulent buffet. Tripping the boundary layer on this airfoil confirmed these conclusions. This is shown in Figure 7, where the global footprint of the shock obtained in the laminar case is compared to the turbulent case obtained with forced transition at 7% of the chord for $(M_0, \alpha) = (0.735, 4°)$. These pictures were obtained by taking the minimum pixel intensity over a set of images that compose the Schlieren movies. The extent of the shock movements in the turbulent case represents about 24% of the chord, while in the laminar case the amplitude of the shock movement does not exceed 5% of the chord (the vertical extent of the dark turbulent shear layer region behind the shock is the mark of the periodic separated flow). The evolution of the flow dynamics towards buffeting was then analyzed by inspecting pressure spectra at increasing angles of attack for a given Mach number, here $M_0 = 0.735$. Figure 8(a) and (b) show the pressure

(a) (b)

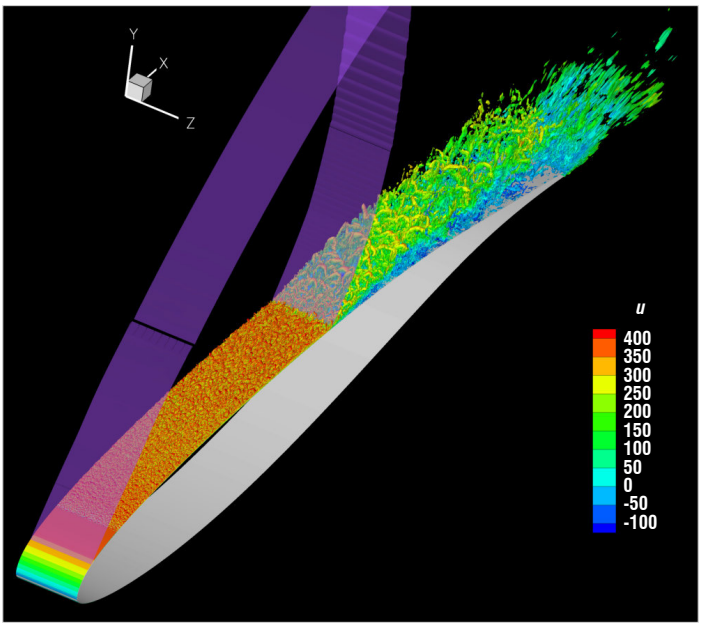
Figure 7 – Buffet at $M_0 = 0.73$ (supercritical airfoil OALT25) – Minimum intensity projection of the Schlieren movies for $(M_0, \alpha) = (0.735, 4°)$: (a) forced transition cases (7%) and (b) natural transition cases, [17].



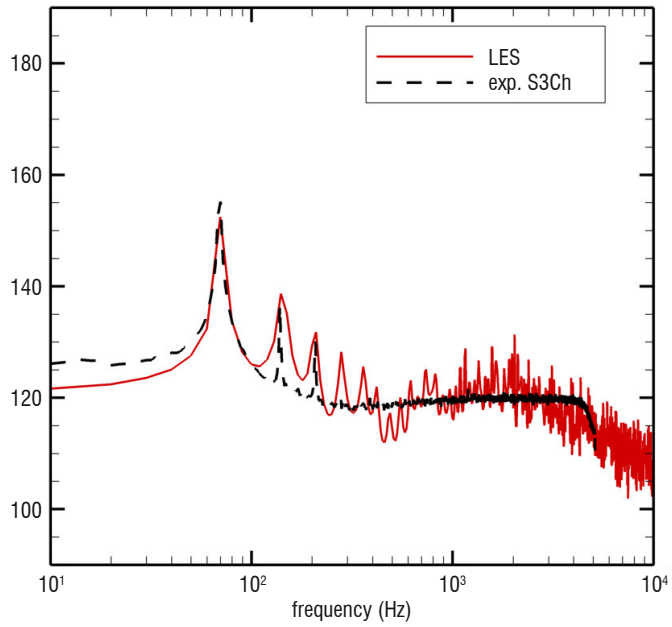
(a) (b)

(c) (d)

Figure 8 – Buffet at $M_0 = 0.735$ (supercritical airfoil OALT25) – Spectra of the pressure fluctuations for a selection of angles of attack α: (a) $x/c = 0.40$, turbulent case (tripping 7%), (b) $x/c = 0.85$, turbulent case (tripping 7%), (c) $x/c = 0.40$, laminar case, (d) $x/c = 0.85$, laminar case, [17].



(a)



(b)

Figure 9 – LES of the buffet flow at Mach number $M = 0.73$ (tripped supercritical airfoil OAT15A) – (a) An iso-value of the Q criteria colored by the longitudinal velocity and one iso-value of the pressure (purple) to mark the shock location, (b) PSD of pressure fluctuations (SPL in dB) at x/c=0.9 [19].

spectra obtained in the tripped boundary layer case at $x/c = 40\%$ (close to the mean shock position at $x/c = 60\%$) and at $x/c = 81\%$; Figure 8(c) and (d) show the pressure spectra in the natural transition case, at $x/c = 60\%$ (at the mean shock position) and at $x/c = 81\%$. The turbulent case at a high angle of attack ($\alpha > 3.5°$) exhibits buffet at 75 Hz, a value very close to that found in the OAT15A experiment. The global nature of the phenomenon is evidenced again by the presence of the same mode at $x/c = 0.40$ in Figure 8(a) and at $x/c = 0.81$ in Figure 8(b). As shown in Figure 8(c)-(d), when the transition is set free the flow behaves very distinctly, exhibiting a global buffeting frequency that is an order of magnitude higher, at about 1130 Hz (incidentally, this corresponds to a chord-based normalized frequency $St = f\,c/U_\infty \approx 1$). This reinforces the conclusion that a laminar buffet is less severe than a turbulent buffet. Low-frequency unsteadiness is felt at $x/c = 0.40$, see Figure 8(c), but it is not modal and it is not retrieved further downstream, see Figure 8(d)[3].

This energetic footprint should be related to the dynamics of the laminar separation bubble identified in Figure 6(b). These results are promising for aeronautical applications, because they suggest that more laminar aircraft could combine two advantages: less drag, due to delayed transition, and a larger flight envelope, thanks to a less severe buffet[4].

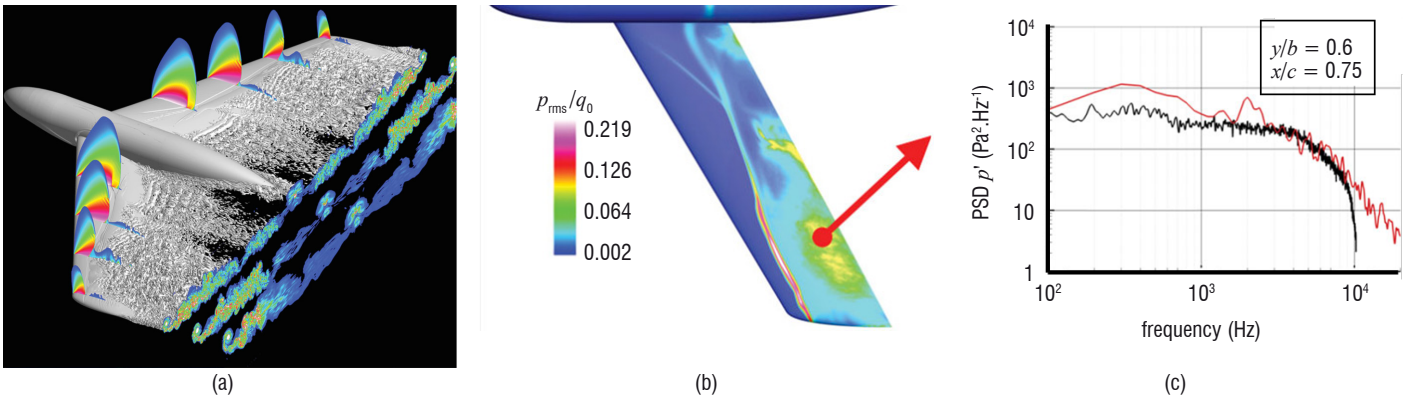
A tribute to CFD

Although this review paper has primarily concerned testing, CFD has accompanied a large part of these experimental studies. After the cited pioneering works based on unsteady viscous-inviscid interaction methods used to guide experiments in the 1980s [6], efficient Navier-Stokes solvers came into play at ONERA as in all other institutions. An example was given in Figure 4(a) based on the use of the ONERA software package for a compressible flow around complex geometries, named elsA [18], for which the above experiments were precious test cases. Reference was also made to studies on the role of the turbulent models in simulations of the transonic buffet by means of Unsteady Reynolds-Averaged-Navier-Stokes (URANS) solvers, [14]. Beyond RANS or URANS approaches, fully turbulent computations became the main challenge in the 2000s. Large Eddy Simulations (LES), Detached Eddy Simulation (DES) and hybrid techniques (ZDES) mixing RANS approaches (for the wall flows) and LES (in the detached regions) were confronted to those test cases. The capacity of a LES to capture the turbulent buffet of the 2D airfoil in Figure 3 is illustrated in Figure 9 [19]. This LES used 42 million cells to compute a domain width of 7.3% of the chord, see Figure 9(a). Figure 9(b) shows that the pressure spectrum obtained on the upper trailing edge at $x/c = 0.9$ compares well to the experiment. The LES method is currently being tested on the case with a natural transition described in Figures 6 to 8. The progress made compared to the 2010s are those provided by high-performance computing (HPC) optimization (code FastS). This leads to a reduction in the computation time by a factor between five and ten for this type of computation [20].

A full LES of a representative 3D aircraft geometry is largely beyond the reach of actual computers. However, the ZDES method can be used, as shown in Figure 10, where the 3D buffet flow in Figure 2 has been computed using 190 million cells, [21]. Results are in rather good agreement with wind tunnel tests, although the low-frequency fluctuations are slightly overestimated due to lack of duration of the computed signal compared with the wind tunnel data, see [21].

[3] All of the data show the presence of a peak at $f = 50$ Hz with a small amplitude around $10^{-6}\,q_0^2$ (where q_0 denotes de dynamical pressure). It is an electrical artifact.

[4] Other experiments were conducted in 2012 in the ONERA S2MA wind tunnel on laminar airfoils designed by Dassault-Aviation within the framework of JTI Clean Sky SFWA-ITD. The results, protected by privacy rules, are generally in line with the current research findings described in this chapter.



(a) (b) (c)

Figure 10 – ZDES of transonic buffet over a civil aircraft configuration ($Re_{mc} = 2.810^6$, $M = 0.82$) – (a) Isosurface of Q-criterion featuring the coherent structures resolved with ZDES and vorticity and Mach contours. (b) Pressure RMS fluctuations on the suction side, showing the location of the shock motion. (c) pressure spectrum at the location identified in (b), [21].

Buffet control

Interestingly, as mentioned above, it has been demonstrated that the transonic buffet is easy to control by open-loop or closed-loop strategies. Early investigations [4] were very successful in showing that vortex generators deployed upstream of the shock suppress the 2D buffet and that moving small flaps located at the trailing edge of a swept wing can significantly reduce the 3D buffet. Further investigations, including fluidic actuation tests, have confirmed the efficiency of such strategies. This topic is beyond the scope of this paper. The reader will find in [22] and [23] comprehensive reviews of various contributions by ONERA to buffet flow control.

Conclusion and perspectives

This article summarizes nearly three decades of research based on wind tunnel tests conducted by ONERA on the transonic buffet, an important basic phenomenon for aeronautical applications. Thanks to significant progress made since the 1980s, the case of a 2D airfoil in a fully turbulent regime is now better understood. The two next steps are: (i) understanding the sensitivity to the boundary layer state revealed by the most recent experiments, (ii) elucidating the transition mechanisms from 2D to 3D buffet. These objectives are now clearly within reach. The success of such researches will require the best possible utilization of the various wind tunnels and CFD platforms available at ONERA. This will otherwise require a strengthening of cross-interactions between simulation and experiment, in particular by considering data assimilation. We will end this review by addressing this challenging research perspective, to which the buffet problem could contribute. Data assimilation is the process of incorporating observations into a mathematical model of a real system, by combining numerical, control and optimization methods (see [24]). Its objective for fluid mechanics is the combination of experimental and numerical analysis to obtain more reliable and more complete information for a real flow. Its applications to aerodynamics are only beginning and the transonic buffet appears to be particularly well suited for producing such techniques. Indeed, as seen above, the 2D buffet flow dynamics strongly depend on the state of the boundary layer in the region of the shock wave. However, in a wind tunnel, this region is too thin and too sensitive to be fully characterized conveniently by means of the available measurement techniques, especially in the laminar regime. Data assimilation techniques combining CFD models of variable complexity (from RANS to LES) with different turbulence and transition models and different sets of flow measurements can be tested carefully using the above described research environment. Extension of the method to the 3D buffet will have to deal with a reduced ability to control the experimental conditions and with the manipulation of a system of much larger dimensions. Meanwhile, adjoining data-driven techniques based on inverse modeling and machine learning can enable, for instance, the construction of accurate models of turbulence and transition, by allowing what is missing in the closure to be inferred and converting that inference into modeling knowledge [25]. Transonic buffet data would challenge such a machinery by physics that are altogether global, local, compressible and highly receptive to details in the shock boundary layer interaction region. To begin with, machine-learned transition and turbulence models, which could easily manage both turbulent and laminar 2D buffet flows, would constitute a breakthrough. All of these constitute promising subjects for the topic of testing in aerodynamics ■

Acknowledgments

Part of this research has been supported by the French Ministry of Defense, which awarded grants for the OAT15A and OALT25 campaigns in the ONERA T2 and S3Ch wind tunnels (special thanks to P. Moschetti) and by the European FP7 Project BUTERFLI (FP7-AAT-2013.8-1-RTDRUSSIA) program, which funded the experiments on the laminar airfoil in 2013.

References

[1] P. TOUBOUL M. RODRIGUES, G. METRIS and B. TATRY - *MICROSCOPE, Testing the Equivalence Principle in Space*. Comptes Rendus de l'Académie des Sciences, Paris, Série IV, Vol. 2, pp 1271-1286, Nov. 2001.

[2] L. JACQUIN - *Scales in Turbulent Motions*. AerospaceLab, 1, 2009.

[3] J. SLOTNICK - *CFD Vision 2030 Study: A Path to Revolutionary Computational Aerosciences.* Report NASA/CR 2014-218178, 2014.

[4] B. BENOIT and I. LEGRAIN - *Buffeting Prediction for Transport Aircraft Applications Based on Unsteady Pressure Measurements*. AIAA Paper 87-2356, 1987.

[5] J.-B. DOR, A. MIGNOSI, A. SERAUDIE and B. BENOIT - *Wind-Tunnel Studies of Natural Shock Wave – Seperation Instabilities for Transonic Airfoil Tests*. Symposium Transsonicum III, IUTAM Symposium Göttingen 25-27.5.1988, J. Zierep and H. Oertel (Eds.), Springer Verlag, 1989.

[6] J.-C. LE BALLEUR and P. GIRODROUX-LAVIGNE - *Unsteady Viscous-Inviscid Interaction Method and Computation of Buffeting over Airfoils*. Proceedings Joint IMA/SMAI Conference, Computational Methods in Aeronautical Fluid Dynamics, Springer-Verlag Ed., 1987.

[7] D. CARUANA, A. MIGNOSI, M. CORREGE, A. LE POURHIET and A.M. RODDLE - *Buffet and Buffeting Control in Transonic Flow*. AST 9, 605-616, 2005.

[8] J. DANDOIS - *Experimental Study of Transonic Buffet Phenomenon on a 3D Swept Wing.* Physics of Fluids, 28, 2016.

[9] M. IOVNOVICH, M. and D. RAVEH - *Numerical Study of Shock Buffet on Three-Dimensional Wings*. AIAA Journal, Vol. 53, No. 2, 2015.

[10] L. JACQUIN, P. MOLTON, S. DECK, B. MAURY and D. SOULEVANT - *Experimental Study of Shock Oscillation Over a Transonic Supercritical Profile*. AIAA Paper 2005-4902 (2005) and AIAA Journal, vol. 47, 9, 2009.

[11] B. H. K. LEE - *Transonic Buffet on a Supercritical Aerofoil*. Aeronautical Journal, pp. 143-152, 1990.

[12] J. B. McDEVITT and A. F. OKUNO - *Static and Dynamic Pressure Measurements on a NACA0012 Airfoil in the Ames High Reynolds Number Facility*. NASA Tech. Paper No. 2485, 1985.

[13] J. D. CROUCH, A. GARBARUK and D. MAGIDOV - *Predicting the Onset of Flow Unsteadiness Based on Global Instability*. J. Comput. Phys., 2007.

[14] M. THIERY and E. COUSTOLS - *Numerical Prediction of Shock Induced Oscillations Over a 2D Airfoil: Influence of Turbulence Modelling and Test Section Walls*. International Journal of Heat and Fluid Flow 27, pp. 661-670, 2006.

[15] F. SARTOR, C. METTOT & D. SIPP - *Stability, Receptivity, and Sensitivity Analyses of Buffeting Transonic Flow over a Profile*. AIAA Journal, Vol. 53, 7, 2015.

[16] J. D. CROUCH, A. GARBARUK, D. MAGIDOV and L. JACQUIN - *Global Structure of Buffeting Flow on Transonic Airfoils*. IUTAM 2009.

[17] V. BRION, J. DANDOIS, J.-C. ABART and P. PAILLART - *Experimental Analysis of the Shock Dynamics on a Transonic Laminar Airfoil*. EUCASS, 2016.

[18] L. CAMBIER, S. HEIB, S. and S. PLOT - *The ONERA elsA CFD Software: Input from Research and Feedback from Industry.* Mechanics and Industry, Vol. 14, No. 1, pp. 159-174, 2013.

[19] E. GARNIER and S. DECK - *Large-Eddy Simulation of Transonic Buffet over a Supercritical Airfoil*. 45ème Colloque d'Aérodynamique Appliquée, AAAF, 22-24 Mars, Marseille, 2010.

[20] Y. MARY and J. DANDOIS - *Private communication*.

[21] S. DECK, F. GAND, V. BRUNET and S. B. KHELIL - *High-Fidelity Simulations of Unsteady Civil Aircraft Aerodynamics: Stakes and Perspectives*. Application of zonal detached eddy simulation. Phil. Trans. R. Soc. A, 372, 2013.

[22] P. MOLTON, J. DANDOIS, A. LEPAGE, V. BRUNET and R. BUR - *Control of Buffet Phenomenon on a Transonic Swept Wing*. AIAA J., Vol. 51, No. 4, April 2013.

[23] J. DANDOIS, P. MOLTON, A. LEPAGE, A. GEERAERT, V. BRUNET, J.-B. DOR and E. COUSTOLS - *Buffet Characterization and Control for Turbulent Wings*. AerospaceLab, 6, 2013.

[24] S. ANDRIEUX - *Inverse Problems and Experiments: a Fruitful Symbiosis*. Ibid

[25] K. DURAISAMY - A Framework for Turbulence Modeling using Big Data. NASA Aeronautics Research Mission Directorate (ARMD), LEARN/Seedling Technical Seminar, January 13-15, 2015.

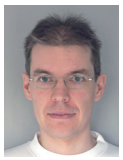
AUTHORS

Laurent Jacquin received his Master's Degree in Mechanical Engineering (1980) and his PhD (1983) from Aix-Marseille University. He obtained his Research Habilitation Degree (*thèse d'état*) from the University of Lyon in 1987 and joined ONERA as a research engineer. He led the Fundamental and Experimental Aerodynamics Department (DAFE) from 2003 to 2015. He is currently the Scientific Director of the Fluid Mechanics and Energetics Branch of ONERA and a Professor in the Mechanics Department of the *Ecole Polytechnique*. His research topics include: turbulence, hydrodynamic stability, vortex dynamics, compressible flows, aerodynamics and experimental methods.

Vincent Brion graduated from the *Ecole Polytechnique* and Supaero (2005), and obtained his PhD (2009) from the Ecole Polytechnique. He has been a research engineer at ONERA in the Department of Fundamental and Experimental Aerodynamics since 2009. His research interests include experimental aerodynamics, boundary layer shock interactions, vortex dynamics and flow control. He has led the Buterfli project on laminar buffet since 2013, most particularly taking charge of the wind tunnel tests at the ONERA S3Ch wind tunnel.

Pascal Molton received a DUT in Thermal Engineering and Energetics from the Ville d'Avray University in 1982. Since then, he has been a research engineer at ONERA in the Fundamental and Experimental Aerodynamics Department. He works in the fields of subsonic/transonic/supersonic flows, jets, vortex flows and flow control.

Denis Sipp graduated from the *Ecole Polytechnique* and obtained a PhD degree from this institution in 1999. He has been a research engineer at ONERA since 2002 and obtained his Research Habilitation Degree in 2009 at the Pierre and Marie Curie University in Paris. He has been the head of the Department of Fundamental and Experimental Aerodynamics since 2015. He has been an Associate Professor in the department of Mechanics at the *Ecole Polytechnique* from 2003 to 2016. His research focuses on hydrodynamic stability, flow control, turbulence, and data-assimilation.



Julien Dandois graduated from the *Ecole Nationale Supérieure d'Arts et Métiers* (ENSAM) in 2003 and received a PhD in Fluid Mechanics from Paris 6 University in 2007. Since then, he has worked in the Civil Aircraft Unit of the Applied Aerodynamics Department, in the fields of high-lift, separation control, buffet control, closed-loop control and aeroacoustics.

Sébastien Deck is a senior scientist in the Applied Aerodynamics Department. His background is: Master's Degree in Mechanical Engineering from the *Ecole Supérieure de l'Energie et des Matériaux* (Orléans) in 1999, PhD from the University of Orléans in 2002, Habilitation Thesis (HDR) from the Pierre and Marie Curie University, Paris 6 in 2007. He joined ONERA in 2002. His Research topics include: CFD, advanced turbulence modeling (hybrid RANS/LES, LES), aerodynamics, flow control, compressible flow, and unsteady data post processing.

Fulvio Sartor received his Master's Degree in Aeronautical Engineering from the University of Rome "*La Sapienza*" in 2011. Following his work on shock/wave-boundary/layer interaction in the Fundamental and Experimental Aerodynamics Department of ONERA, he obtained his PhD in Fluid Mechanics from Aix-Marseille University in 2014. After a short period at the University of Liverpool, where he worked as a research associate on the numerical simulations of transonic buffet, he joined ONERA as a research engineer in the Applied Aerodynamics Department.

Eric Coustols graduated from ENSEEIHT Toulouse in 1979. He received a Master of Sciences Degree from the University of California San Diego in 1980 and defended his PhD thesis at the ENSAE (ISAE/SUPAERO) in 1983. He has been working at ONERA in the Department of Modeling for Aerodynamics and Energetics since 1983, in various research fields aimed at improving civil aircraft performance. For the last twenty years or so, he has been coordinating the involvement of several ONERA research units in EU-funded projects; from Jan. 2007 to Dec. 2013, he led the ONERA joint research project BUFET'N Co related to Buffet Control on a 3D turbulent transonic wing. Currently, he is the Director of the Clean Sky 2 Program at ONERA.

Daniel Caruana is a graduate of the C.N.A.M. (*Conservatoire National des Arts et Métiers*) and is currently a Senior Research Engineer at ISAE-SUPAERO school. At ONERA, he was in charge of plasmas for flow control and also involved in many aeronautics studies for aerodynamic and aero-acoustic control, flow physics, actuators definition and design, control technique. He has been the coordinator of the PLASMAERO European project (2009-2012). He joined ISAE/SUPAERO school in June, 2015. He is the lead of Technical Team and deputy director of the Aerodynamic, Energetic and Propulsion Department. He has in charge the management of the experimental studies, research test benches and wind tunnels including the new Aeroacoustics Wind Tunnel facility.

**B. Leclaire, C. Brossard,
R. Courtier, F. David,
S. Davoust, A. Gilliot, L Jacquin,
J.-M. Jourdan, O. Léon,
S. Masseboeuf, J.-C. Monnier,
S. Mouton, E. Piot, A. Ristori,
D. Sebbane, F. Simon**
(ONERA)

E-mail: benjamin.leclaire@onera.fr

DOI: 10.12762/2016.AL12-07

Planar Particle Image Velocimetry for Aerospace Research at ONERA

Planar Particle Image Velocimetry (PIV) is a measurement technique that yields the instantaneous spatial structure in a planar cut of a flow. It exists in several forms (two-component, three-component or Stereo, and time-resolved), all of which are now considered to be mature, and as production tools for flow investigation in experimental facilities. This contribution is aimed at showing, through various examples of research activities, how it is currently being used at ONERA, with the idea of extending knowledge on complex flows, and therefore on the performance of complex systems. After a short reminder of the operating principle of planar PIV, applications are presented in the contexts of jet mixing physics, helicopter fuselage drag reduction, reacting flows, industrial wind-tunnels, and acoustics.

Introduction

Particle Image Velocimetry (PIV) is today one of the most widely used measurement techniques in fluid mechanics, in a large variety of situations ranging from fundamental flow physics to various different application fields, such as geophysics, industrial processes, and aerospace sciences. Its success is explained both by its non-intrusive nature, and by the wealth of spatial and, possibly, temporal information that it may provide. The basic idea behind PIV is to spread a fog of solid particles or liquid droplets into the flow to be measured, with constraints on the particle density and size so that they can be considered to be passively entrained by the flow. Illumination of these particles is then performed at two or more consecutive instants separated by short time intervals (possibly as short as less than 1 μs, for the highest flow velocities), and images of these particles at each instant are acquired using CCD or CMOS cameras. Tracking of the particle motion through suitable algorithms then enables the corresponding flow displacement to be determined, and then the velocity field.

For roughly three decades, this measurement technique has motivated very important research efforts in institutes, laboratories and universities worldwide, which have contributed to significantly enriching its capabilities in terms of spatial and temporal flow characterization. This progress has also been made possible by simultaneous advances in light sources and imaging systems. To name but a few, and with no intention of providing an exhaustive account of these research efforts (see for instance [1] or [2], for reviews), PIV started out as a two-component, two-dimension (2C2D), low-frame rate technique, yielding instantaneous snapshots at a low repetition rate (typically 5 to 10 Hz) of two flow velocity components in a plane, and is now available as a three-component, three-dimensional (3C3D), high-frame rate technique, providing volumic snapshots of the three velocity components with repetition rates as high as typically 10-20 kHz [3]. In parallel, dedicated processing approaches have made possible the addition of new analysis features to the initial statistical and instantaneous characterizations. For instance, specific time-resolved (see, e.g., [4], [5], [6]) or Lagrangian tracking approaches [7] have emerged, and, building upon this wealth of information, algorithms able to reconstruct pressure fields from the velocity fields, as well as efforts on objects placed in the flows, are now available (see for instance [8] for a review). In the same spirit, several new strategies for processing 3C3D (possibly partially) time-resolved data are inspired on computational fluid dynamics, thereby bridging the gap between experiment and numerical simulation (see [9], [10], [11], [12], among others). The above mentioned developments and studies, though particularly significant, still only represent a fraction of a very large variety of uses, processing, and applications of the PIV technique among the fluid mechanics community.

In this paper, we choose to focus on illustrating how the most well-mastered form of PIV, i.e., planar PIV (possibly Stereo and/or time-resolved), is used at ONERA as a key ingredient in the aerospace research performed there. Three-dimensional techniques, both of the PIV and Particle Tracking Velocimetry (PTV) type, are also available at ONERA but are excluded from this paper, because they are covered by a companion paper of this volume [13]. Selected examples from recent works at ONERA, each representing a particular context and objective, will be presented, together with the achievements and

results made possible by the use of PIV. Note that for conciseness purposes, and in order to focus mostly on applications, only a slight account will be given of the setup and processing details. With regard to the latter point, it can be mentioned that in all examples traditional PIV algorithms based on the cross-correlation of interrogation windows (IW) will be used. With this processing, each of the vectors obtained in the result snapshots corresponds to the average motion of particle patterns contained in the IW defined around this vector. The applications shown will encompass the physical analysis of turbulent jet mixing, helicopter fuselage drag reduction, an overview of the capability of PIV in the context of reacting flows and in industrial wind-tunnels, together with the constraints. A last focus will finally show how PIV can be used to perform aeroacoustic measurements.

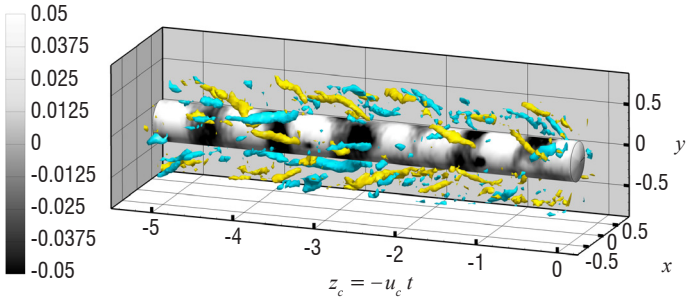


Figure 2 – Pseudo-spatial flow reconstruction in the z direction from time-resolved data. Grey contours identify the axial velocity fluctuation u_z' at $r/D = 0.24$, while yellow (resp. blue) iso-surfaces identify $\omega_z' = 4$ (resp. -4) iso-values of the axial vorticity fluctuation.

TR Stereo PIV for the study of jet turbulence and its control

Jet generated turbulence is a fundamental subject of foremost importance, because it plays a strong role in applications such as mixing, combustion and propulsion, for instance. In the latter domain, it is closely linked to the generation of aeroacoustic disturbances, and underlies the overall furtivity of systems. Several decades of research have enabled the largest scale and the most coherent unsteady perturbations occurring in the mixing layer (the thin region separating the jet from the quiescent surroundings) to be singled out as key ingredients in the global mixing. Among these, axial vorticity fluctuations have received a lot of attention, since they are directly responsible for injections and ejections of fluid, thereby achieving mixing. Understanding their origin and behavior is therefore crucial to propose flow control strategies aimed at enhancing mixing and obtaining complete jet dissipation over a smaller axial distance, for example.

In order to complete the knowledge on mixing mechanisms and to propose more efficient strategies for its control, ONERA recently embarked on experimental studies based on Time-Resolved Stereo PIV on turbulent round jets, at an exit diameter-based Reynolds number of 210,000 [14] [15]. Jets are generated in the open-circuit subsonic wind-tunnel R4Ch, specially designed for fundamental studies on jets and turbulence (see Figure 1). Depending on the contraction ratio of the final nozzle generating the jet, its operating velocities may range from a few m.s^{-1} up to 50 m.s^{-1}. Upstream of the contraction, a portion of the cylindrical pipe, in which a honeycomb with 1 mm-wide cells is placed, can be set in rotation. This enables a component of solid-body rotation to be added to the flow, and the instabilities and unsteadiness present in propulsive jets to be modeled more finely, for instance. Such a capability will be exploited in forthcoming works, given that the present study focuses on non-rotating jets. As depicted in Figure 1, measurement is performed in a cross-sectional plane of the jet, located 2 diameters (D) downstream from its exit. Results are obtained here using the GPU-implemented software FOLKI-SPIV, developed by the departments DAFE and DTIM of ONERA [16].

Aside from direct three-dimensional measurement, which involves a still complex setup, one way to characterize and investigate the three-dimensional structure of flows can be to rely on Taylor's "frozen turbulence" hypothesis, i.e., in the present case, to assume that turbulent coherent structures do not change while they pass through the measurement plane. Suitable post-processing based on this hypothesis [17] then enables a so-called pseudo-spatial flow reconstruction to be obtained, as shown in Figure 2. This figure evidences the presence of streamwise vorticity fluctuations in the mixing layer region (yellow and blue contours), and of their spatial synchronization with axial velocity fluctuations in the jet core (grey contours), the latter being mostly associated with the well-known axisymmetric Kelvin-Helmholtz vortices [14].

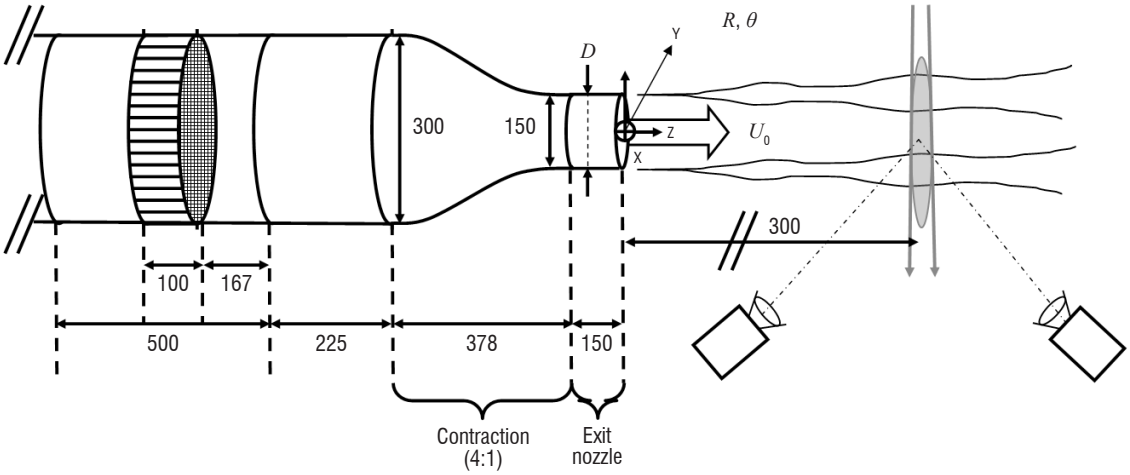


Figure 1 – R4Ch wind-tunnel test section, generated jet and Stereo PIV setup with laser sheet (grey) and cameras.

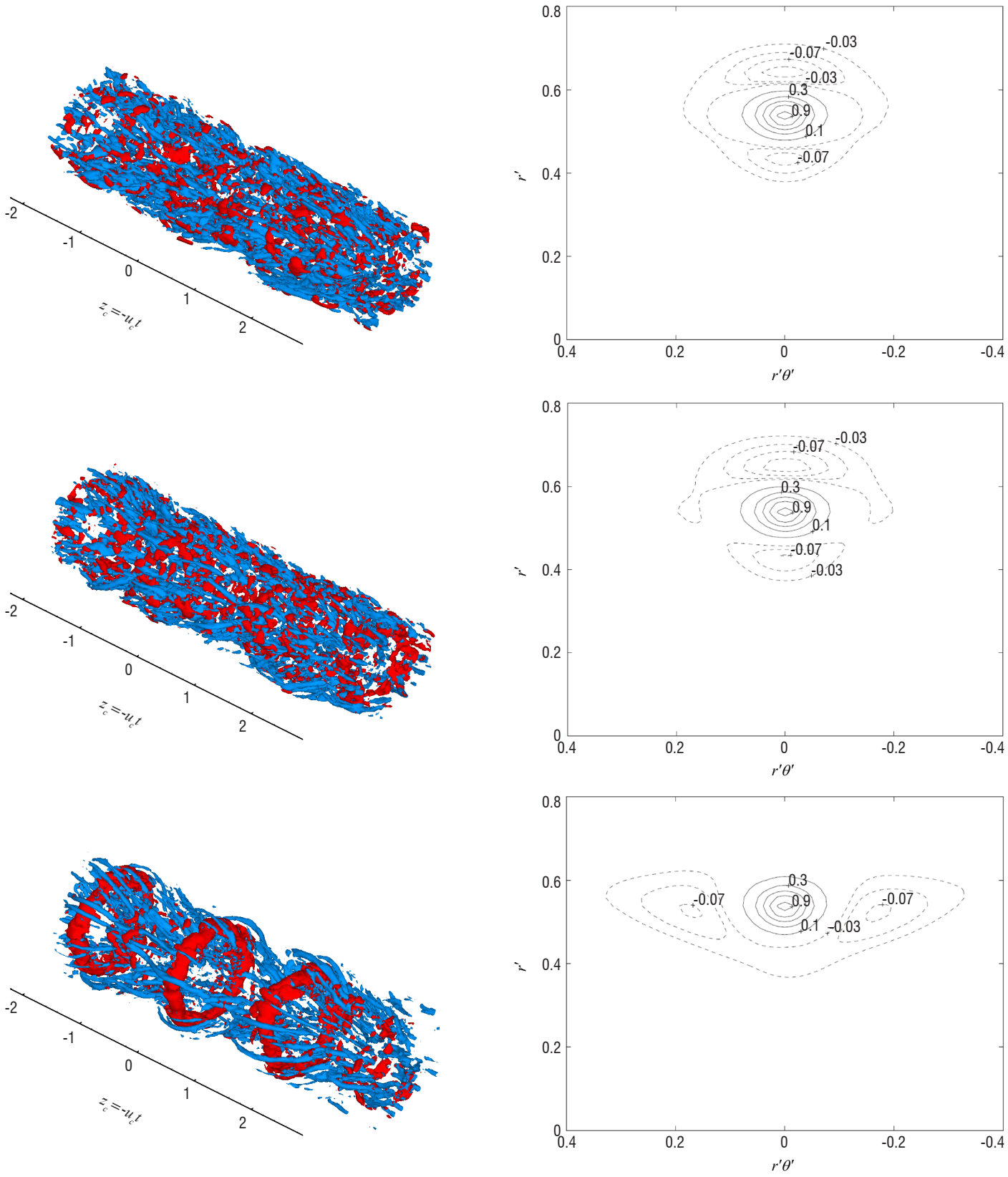


Figure 3 – Jet pseudo-spatial reconstructions (left column) and correlation of the axial vorticity fluctuation taken at reference point $r = 0.52$ (right). Unforced jet (top row), intermediate forcing (medium row), strong forcing (bottom row).

A more in-depth analysis, using spatial auto-correlation of the axial vorticity fluctuation ω_z', helps the statistical spatial structure of the streamwise vortices responsible for mixing to be characterized more finely. The right column of Figure 3 shows iso-contours of this auto-correlation, the sampling point being located in the middle of the shear layer, at a radius $r/D = 0.52$, and distances from that point being denoted by primes. The left column of this figure shows the corresponding pseudo-spatial reconstructions. Three situations are considered: the natural, unforced

jet (top row, as in Figure 2), and two intensities of axisymmetric forcing obtained by operating a loud-speaker at the natural jet preferred mode frequency in the wind-tunnel plenum chamber (medium row: medium intensity, and bottom row: high intensity). In the natural situation, one observes that streamwise fluctuating vortices of alternate signs are preferentially piled in the radial direction, which is new and specific to the present Reynolds number [14]. Indeed, past works, on lower Reynolds number jets, have reported piling in the azimuthal direction. As Figure 3 shows, this new spatial organization is robust and strongly linked to the global flow structure. Indeed, it survives a medium level of forcing, as seen in the medium row, and a sufficiently high forcing intensity of the axisymmetric Kelvin-Helmholtz vortices is necessary to retrieve a structure more similar to that of lower Reynolds number flows, as shown in the bottom row. In this case, correlations indeed show negative and positive patches along the azimuthal direction θ', and the pseudo-spatial reconstruction is strongly dominated by Kelvin-Helmholtz vortices (in red).

All of this evidence shows that in sufficiently high Reynolds number jets, streamwise fluctuating vortices and the mixing in general are less influenced by Kelvin-Helmholtz azimuthal vortices than at lower Reynolds numbers. As described in more detail in [14] [15], further pursuing the analysis of flow characterizations and reconstructions offered here by TR Stereo PIV makes it possible to show that the mean shear of the mixing layer plays a larger role, and to propose scenarios for the existence and location of these vortices.

PIV for active flow control: helicopter fuselage drag reduction

The reduction of helicopter parasite drag has become a significant issue as emphasis is placed on speed achievement and on minimizing the impact of the growth of rotorcraft traffic on the environment. The primary source of drag of helicopters is bluff body drag (pressure drag) from pylon/hub/fuselage/landing gear flow separation and interference. Fuselage drag is mainly due to the massive flow separation that occurs at the rear of the fuselage, especially for helicopters with a very pronounced aft loading ramp. A comprehensive experimental investigation of helicopter blunt fuselage drag reduction using active flow control is being carried out within the CleanSky project [18] [19]. The objective is to demonstrate the capability of several active technologies to decrease fuselage drag by alleviating the flow separation occurring in the backdoor area of some helicopters (with pronounced ramp for backdoor loading). The work is performed on a simplified blunt fuselage at model-scale. Several active flow control actuators are considered for evaluation: steady blowing, unsteady blowing (or pulsed jets), and zero-net-mass-flux blowing (or synthetic jets).

The experimental tests of this research program are conducted in the ONERA L1 wind-tunnel located in Lille. This large facility is an Eiffel type wind-tunnel with a return hall. It can reach wind speeds of up to 75 m/s ($Tu \approx 0.3\%$) and its test section is circular with a diameter of 2.4 m. The model is a generic helicopter fuselage (see Figure 4). Its backdoor region is equipped with 8 removable actuators in a U-shape arrangement (see Figure 5). The model is mounted upside down on two struts in the test section, so that the supports do not interfere with the ramp region. Drag, lift and pitching moments are measured, by means of a Monin balance, at various angles of attack. The model is also instrumented with pressure taps and unsteady pressure transducers. In order to complete this usual instrumentation, PIV has been used to obtain quantitative information on the instantaneous and mean topology of the flow around the generic helicopter fuselage.

These PIV investigations have only been focused on the backdoor region of the model and were recorded for different experimental configurations (baseline, with continuous, pulsed and synthetic fluidic actuators). For PIV 2C and 3C measurements, the tracers used are olive oil particles produced by a Multi-Laskin nozzle device located in the return circuit of the wind tunnel. With this seeding system, only a few seconds are necessary to obtain a good and homogeneous concentration of tracers in the test section and, more particularly, in the PIV measurement regions. For illumination of these tracer particles, the light source is a frequency-doubled Nd:YAG laser with two oscillators each supplying two consecutive pulses at 5 Hz frequency with about 400 mJ per pulse. This equipment is located above the test section of the wind tunnel. The light beam emitted by each laser passes through a same set of spherical and cylindrical lenses. It is directed, by means of mirrors, towards the measurement areas, through a window glass located in the roof of the test section. In the laser sheet generator, the angular position of the cylindrical lens allows the measurement plane to be oriented in a direction parallel or perpendicular to the upstream flow. Two high-resolution CCD cameras (2048 x 2048 pixels^2, 12 bits) were used for the recording of PIV images. The frequency of acquisition is limited to 5 double-frames per second. In the pulsed configuration of the actuators, the recording of PIV images is synchronized on

Figure 4 – Simplified blunt fuselage model in the test section of the wind-tunnel

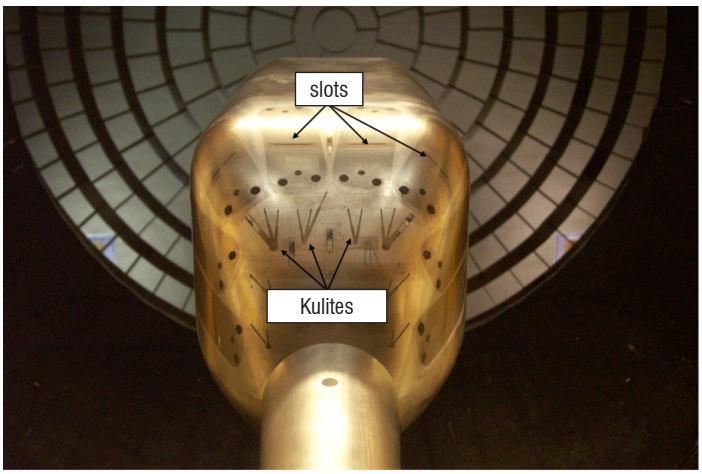


Figure 5 – Removable actuators in the backdoor region of the fuselage.

the command signal of these devices. This synchronization is provided by a programmable sequencer. Under these conditions, all PIV frames of one run are recorded at the same phase chosen in the periodic cycle of actuators. The analysis of all PIV images was carried out by the DaapPIV software developed by the department DAAP. For this evaluation, a local iterative multi-grid cross-correlation algorithm has been applied to all recordings. At the end of this analysis process, the final size of the interrogation windows is 32x32 pixels at a 16-pixel (overlap of 50%) step width in two spatial directions. In the transverse (PIV 3C) and longitudinal (PIV 2C) planes, it corresponds to a spatial resolution, respectively, of 2.4 mm and 1.9 mm.

Figure 6, Figure 7 and Figure 8 present some examples of PIV results obtained on the reference configuration with and without blowing. The streamlines clearly show the topology of flows, the presence of vortices and the effects of the jets on the latter (more information in [18] [19]).

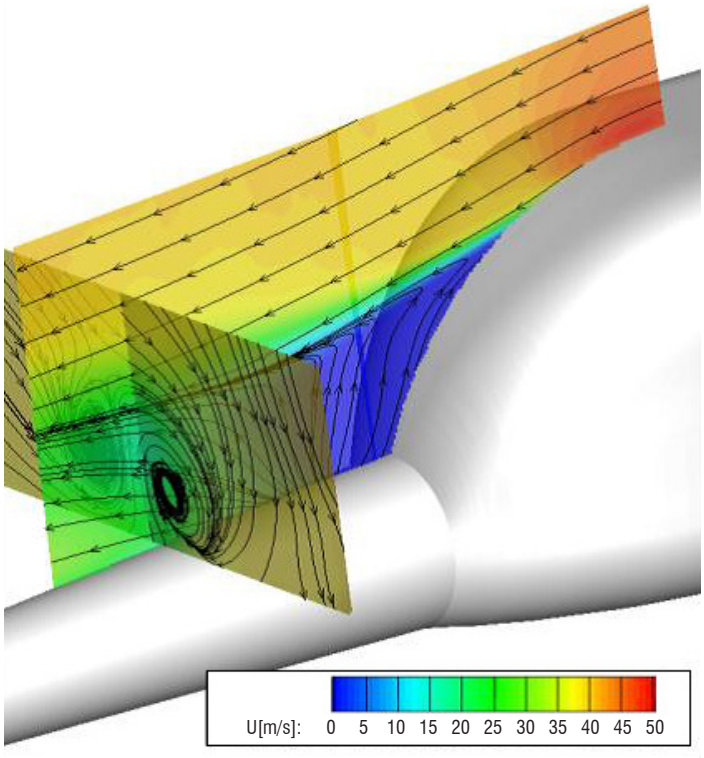


Figure 6 – PIV measurements for $\alpha = +5°$ and baseline (without blowing)

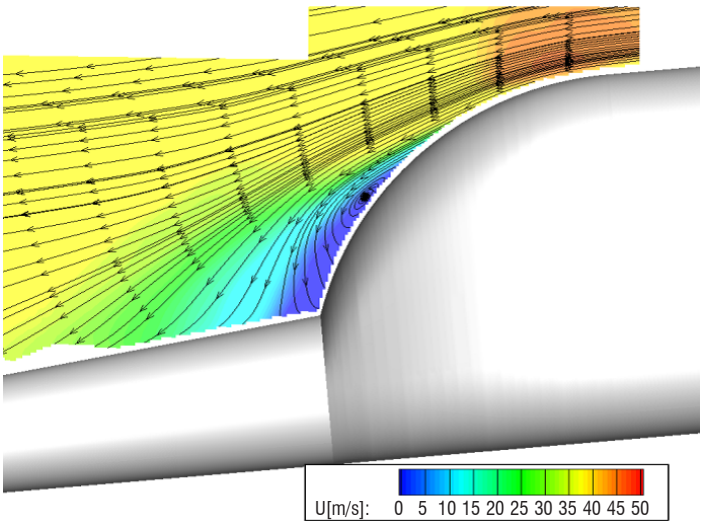


Figure 7 – PIV measurements for $\alpha = -5°$ and baseline (without blowing)

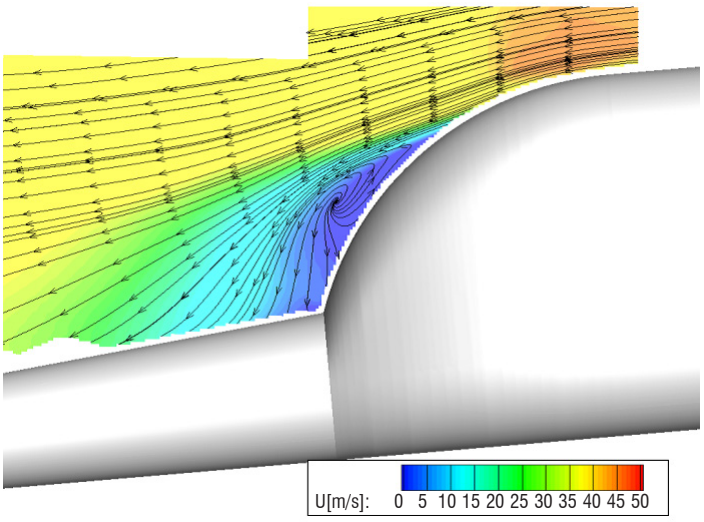


Figure 8 – PIV measurements for $\alpha = -5°$ and blowing

The double topological and unsteady approach accessible by PIV has been a great help in the understanding of the phenomena existing in this type of flow, as well as to evaluate the effectiveness of various fluid actuators studied during the experimental tests [18] [19]. At the present time, major efforts are underway to facilitate the generalization of this measuring technique in the large research facilities of the DAAP and to better meet the demands of its customers.

PIV for the characterization of reacting flows

In combustion experimental studies, the knowledge of the velocity field and its structure is of prime interest, since it plays a key role in the flame stabilization process. Therefore, the development of PIV has allowed significant insight to be gained in reacting flows, with respect to the previously used Laser Doppler Velocimetry (LDV) technique, which provided only a local measurement. PIV has been extensively used in combustion research or development laboratories worldwide. At ONERA, it has also been continuously implemented in air-breathing test facilities located in Palaiseau.

The time development of the research ramjet combustor program conducted at ONERA since the mid-90's illustrates the benefits brought first by the introduction of the PIV technique, and then by the improvement of its technical performance. This research program, supported by the French DGA (*Direction Générale de l'Armement*), was initiated at ONERA with the aim of improving ramjet combustion chamber design and tuning methods by using validated Computational Fluid Dynamics (CFD) codes. A three-dimensional ramjet combustor geometry (combustor model) was specifically defined and built for investigations in hot flow conditions, with combustor operating conditions (pressure, velocity, temperature) comparable to

those of real engines. This combustor, a schematic view of which is represented in Figure 9, was designed to be used either as solid-ducted rocket or liquid-fueled ramjet engines.

Experimental investigations of the high-speed flow inside the duct section of the combustor started in 2001, using a 2D-LDV system [20]. However, this required a large number of test runs, in order to be able to reconstruct the mean and fluctuating flow structure through several transverse profiles of the obtained local mean velocities and turbulent correlations, measured at different axial locations along the length of the duct section. Later on, the introduction of the PIV technique allowed, in addition, instantaneous planar flow patterns to be obtained for the first time [21]. An example of the mean flowfield obtained for reacting case conditions is provided in Figure 10a. PIV and LDV data were found to be in good agreement, as shown in Figure 10b. The PIV measurement results then allowed the impact of chemistry description in Large Eddy Simulation (LES) based computations [22] to be studied. However, these first PIV data were obtained with a low repetition rate (3.75 Hz), which did not allow unsteady phenomena to be investigated. This could be done more recently thanks to the development of high-speed or time-resolved PIV systems. PIV was performed at 1,000 Hz and synchronized with visualizations of the OH* chemiluminescence – indicating the location of the reacting zones – at the same repetition rate [23].

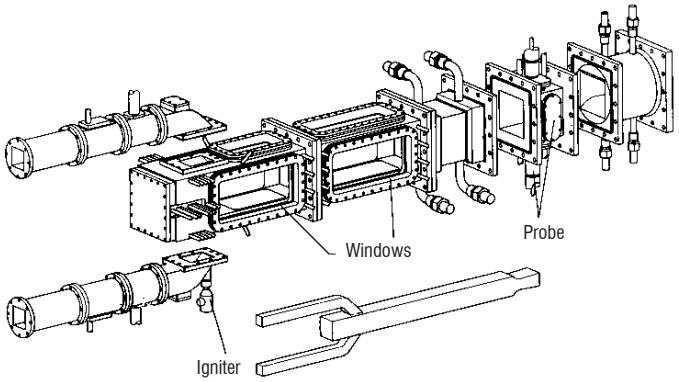


Figure 9 – Schematic view of the research ramjet model [20-21].

Compared to non-reacting flows, the implementation of PIV in reacting flows requires additional technical constraints to be taken into account. The seeding of the flow has to be performed with small (in the micron range) solid particles (usually metal oxides, such as MgO or TiO_2). However, dispersing systems used to generate such

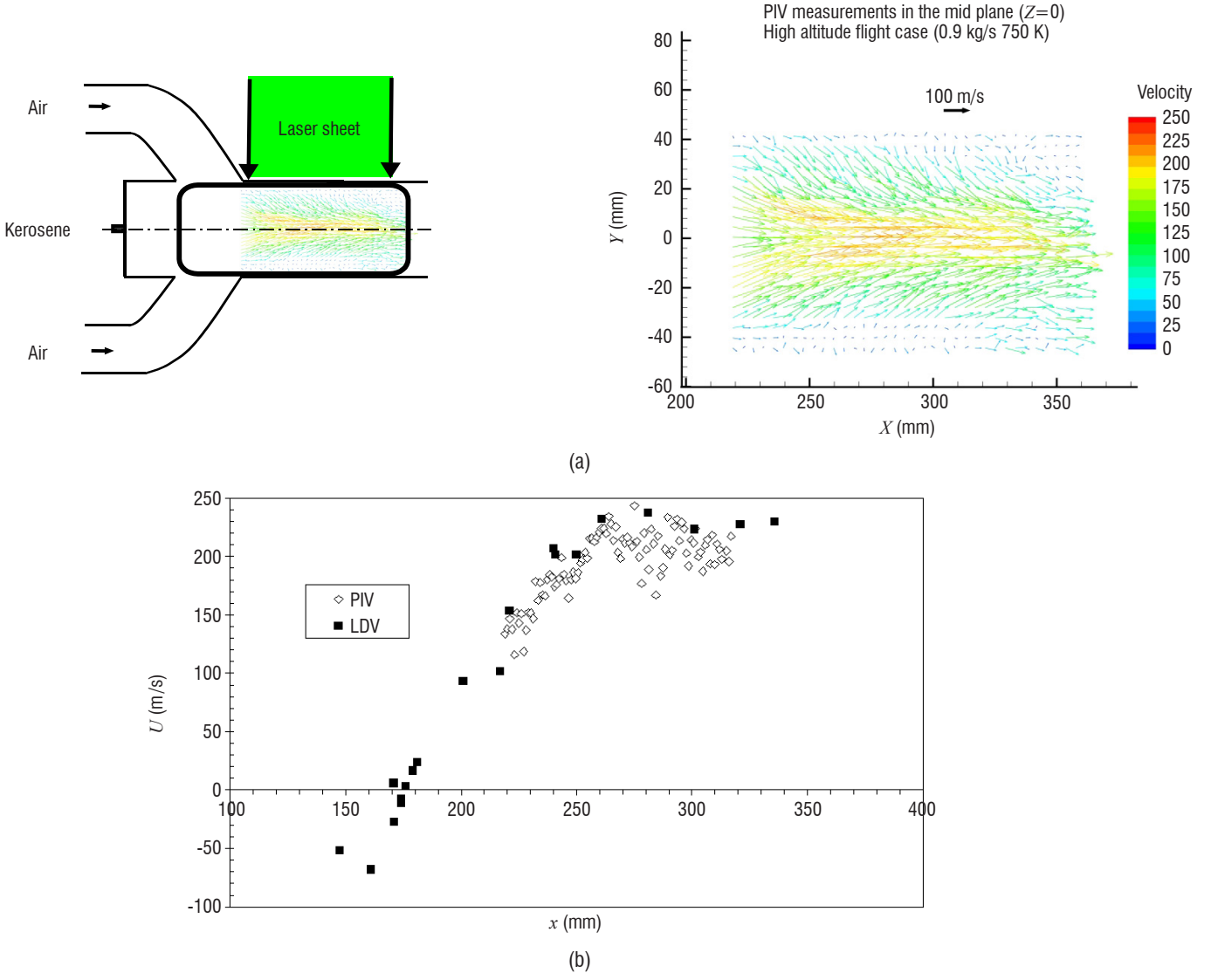


Figure 10 – Examples of PIV measurements obtained in the mid-plane of the research ramjet combustor duct section, for a reacting flow case [21]: a) mean flowfield obtained by PIV; b) Comparison with LDV measurements in the axial profile (mean longitudinal velocity component).

particles are prone to unsteady behavior, and precautions must be taken in the operating procedure, in order to limit particle deposits on the windows. Window access on combustion chambers is usually limited due to mechanical and thermal constraints, thus limiting the possibility of using stereoscopic PIV systems. So far, PIV has been conducted in tests with chamber pressure levels limited to 0.2 MPa. The development of a new test rig, named "MICADO", is in progress at the ONERA Palaiseau center (MICADO is a French acronym that stands for "Investigation Means for Air-breathing Combustion using Optical Diagnostics"). This test rig will be dedicated to the study of combustion under test conditions representative of single or multi-sector aeronautical engine combustors, and to the development of the associated advanced diagnostics. The maximum target chamber pressure is 4.0 MPa. One of the main challenges will be to adapt the PIV technique so that it can be used at chamber pressures as high as possible. In particular, beam steering effects are one of the issues that must be addressed. These effects, generated by optical index gradients in the combustion chamber, result in image blurring, and can be very significant in highly turbulent flows at high chamber pressures.

PIV and Stereo PIV in industrial wind-tunnels

Particularities of PIV in industrial wind-tunnels

Over the past years, the PIV technique has started to be used for commercial purposes in large industrial wind-tunnels, at ONERA and elsewhere. This has posed a number of challenges in the progress towards an offer compatible with industry needs. In terms of hardware, the tunnel test-section has a larger size than that needed in most laboratory experiments and offers limited optical access, which requires careful optical adjustments, powerful lasers and suitable seeding techniques. Access to the PIV hardware during the experiment is often difficult and time-consuming due to its out-of-reach position, or pressurization. In addition, the flow often needs to be examined at numerous positions in space. These limitations are progressively overcome by means, for example, of enlarged windows or remote-controlled hardware.

Most of the time, both large and complex flow fields are investigated, which necessitates a compromise to be made between field-of-view and accuracy. Also, depending on the tunnel environment, it may be impossible to position the cameras in the perfect place, to eliminate vibrations and background on the images, to achieve and maintain proper seeding. As a consequence, the flow frequently presents large sub-window movements, zones with small values of the particle displacement, poor particle illumination for at least one camera, and other flaws that can make images challenging to process.

In preparation and during the test campaign, time and cost constraints are significant and a high level of productivity is expected for both the setup and the image acquisition and processing. Last but not least, staff training is crucial to transfer the know-how from highly skilled PIV researchers to people in the wind tunnel performing the measurements.

Examples of 2 recent uses

High-accuracy Clinometric PIV in the F1 pressurized wind-tunnel

The F1 wind tunnel is a low-speed, continuous-flow, pressurized wind tunnel, with a test section 3.5 m high, 4.5 m wide, and 11 m long. It is used by a large number of customers from all over the world for testing aircraft in landing and take-off configurations, as well as for research activities.

Several measurement campaigns have been carried out with PIV, in order to refine our knowledge of the flow angularity (upwash and sidewash) in the empty test section of the F1 wind-tunnel. An original approach, reported in [24], was used to achieve an angle accuracy estimated at $\pm 0.05°$ and a speed accuracy estimated at $\pm 0.1\%$. Results have been obtained with various different PIV setups (both 2-component and 3-component) and compared to the techniques historically used to assess angularity and flow speed, i.e., 5-hole, Pitot probing and the model inversion method.

As shown in Figure 11, measurements have confirmed the existence of a moderate upwash of about 0.16° in the test section. For this application, PIV demonstrated important advantages over historical methods: non-intrusiveness, high spatial density of measurements and applicability close to the floor. Measurement provided an unprecedentedly detailed description of non-uniformities in the wind tunnel flow, especially near the floor, therefore securing and improving corrections brought to wind tunnel measurements.

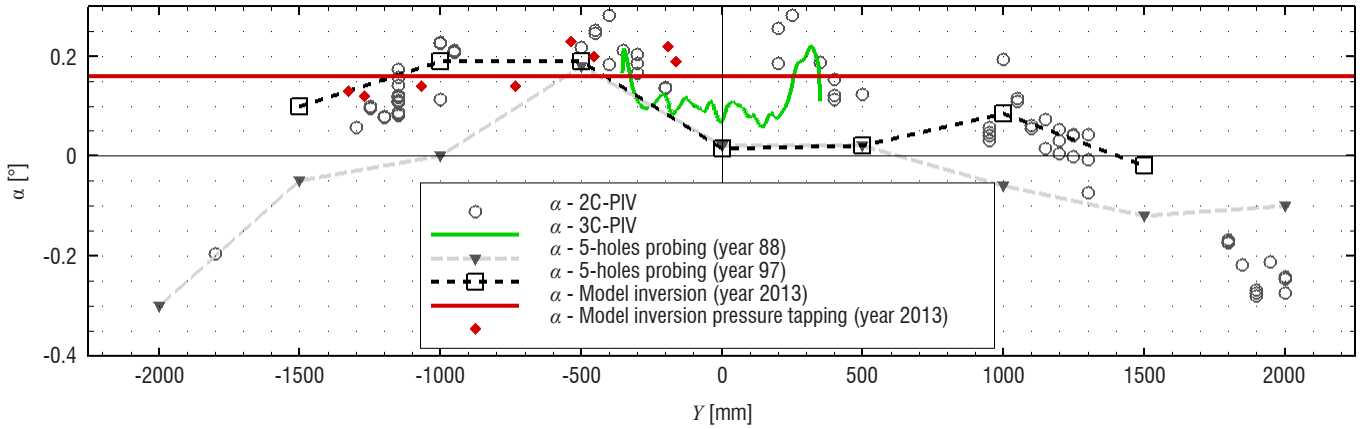


Figure 11 – Distribution of upwash along the spanwise direction at the center of the test section

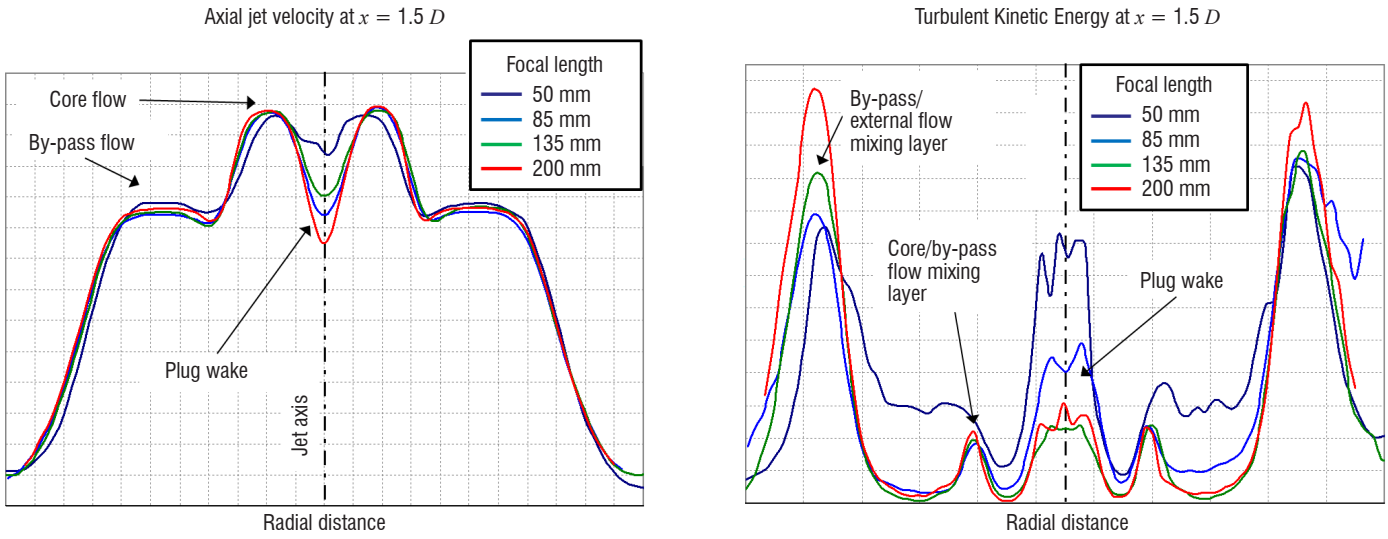


Figure 12 – Radial distribution of mean velocity (left) and TKE (right) in the jet flow. The size of the interrogation windows in pixels is kept the same between different focal lengths

Jet flow characterization in the CEPRA 19 anechoic wind tunnel

CEPRA19 is a large-scale, open-jet, anechoic wind tunnel particularly suitable for aeroacoustic studies of modern dual-stream, high-by-pass-ratio engine jets. Besides classical aerodynamic and acoustic measurement, PIV measurements have been carried out several times to characterize jet flows [25]. Beyond the distribution of the mean flow, it is interesting in acoustics to know the intensity of velocity fluctuations because they are strongly related to the generation of noise. To serve that purpose, a recent campaign was performed to improve the measurement accuracy of the TKE (Turbulent Kinetic Energy) [26].

As an example of the results of these investigations, a comparison of measurements with increasing focal length for camera objectives (hence increasing the spatial resolution) is presented in Figure 12. It can be seen that the mean axial jet velocity is easily captured in almost all cases. On the contrary, properly measuring the TKE is more challenging and only the larger focal lengths provide valuable results. This kind of measurement is highly interesting to validate unsteady flow simulations and to understand the physical mechanisms behind the noise reduction offered by nozzle devices such as chevrons or micro-jets.

Future needs

Industrial applications of PIV will naturally benefit from the probable continuation of the increase in computing power and camera resolution, but there are also some more specific needs. An immediate need is to better quantify the uncertainty of the PIV measurement. Considering the discrepancy between and across images (due both to flow features and to the imaging technique), a local uncertainty assessment is desirable. Generalization of remote-controlled hardware, possibly automated, is also desirable to reduce the time to set up the hardware and to be able to make adjustments at the tunnel startup or during the run (e.g., to correct deviations of the laser beam trajectory). Adaptation of wind-tunnels should be continued, in order to foster easier preparation of the tests. Interesting flow features can be found close to the model surface, but it is often difficult to perform reliable measurements in this area, because of laser reflection (despite the care taken to paint the model), model movements, the extreme close-ups demanded due to cameras being located far away, and optical access in a concealed area. Any progress in this direction would be welcome. Finally, in order to gain attractivity, a more elaborate outcome of the PIV measurements than velocity mapping should be sought. Examples include drag and lift analysis performed from wake surveys, or aeroacoustic applications.

TR-PIV for the measurement of acoustic velocities in a turbulent duct flow

Context

The development of non-intrusive measurement techniques for in-flow acoustic pressure and velocity estimations is necessary to better characterize and analyze the response of aeroacoustic surface treatments, such as acoustic liners. Over the past decades, several optical methods have been proposed [27]. Particularly, Minotti *et al.* [28] successfully developed at ONERA an approach based on Laser Doppler Velocimetry (LDV) allowing the extraction of both acoustic and turbulent flow velocities. Assuming no correlation between flow turbulence and acoustic perturbations, their approach was based on an estimation of the acoustic velocity components through spectral cross-correlations between the measured velocity signals and an unsteady pressure measurement performed on the wall of the duct in which multi-sine acoustic waves are generated. Other methods based on Doppler Global Velocimetry were later proposed [29], but with very limited spatial resolutions. In the literature, the use of PIV for similar purposes mainly relies on using phase-averaging approaches [30] [31], which is a non-trivial task when dealing with multi-sine acoustic waves. The present work demonstrates the applicability of the method developed by Minotti *et al.* [28] to the measurement of acoustic velocities in highly turbulent duct flows using TR-PIV.

Experimental Setup

In order to validate the TR-PIV acoustic velocity measurement approach proposed, a simple test case was selected. Measurements were performed with the ONERA aero-thermo-acoustics bench (B2A) in Toulouse. A schematic view of the bench is provided in Figure 13, the test cell having a square section of 50 mm x 50 mm. The bulk Mach number was set to Mb = 0.1 and the flow temperature was regulated at T = 20 °C. A simple acoustic plane wave was prescribed upstream at a frequency fac = 800 Hz and at various sound pressure levels above 130 dB. TR-PIV measurements were performed at a frequency of 3 kHz in the horizontal plane (xy) located in the middle of the test section. Turbulent intensities of about 5% were observed in the measurement window, leading to a signal-to-noise ratio of -12 dB for an acoustic wave at a sound pressure level of 138dB (results presented in Figure 14). For validation and comparison purposes, LDV measurements were performed under the same conditions on the test cell centerline.

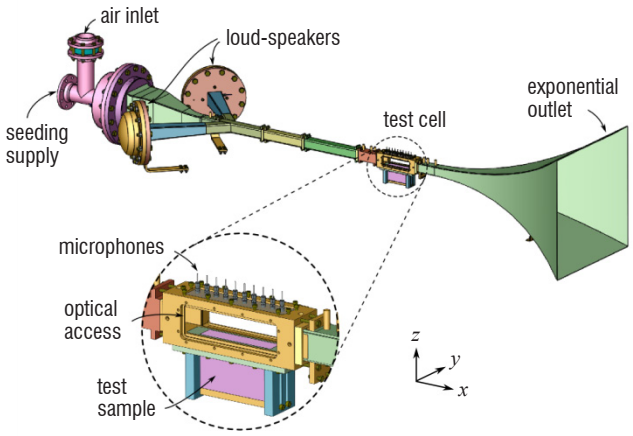


Figure 13 – Diagram of the aero-thermo-acoustic bench (B2A) used for experimental liner investigations.

Results

The results of both the TR-PIV and LDV campaigns were post-processed using the turbulence rejection method previously mentioned [28]: acoustic amplitudes and phases with respect to a reference signal, here the loud-speaker input, are evaluated using cross-power spectra estimates. For TR-PIV measurements, 30,000 pairs of images were acquired for every test condition, with a 1 Mpx camera. Such a large number of samples is necessary to provide accurate estimates, given the signal-to-noise ratios at play. Post-processing of TR-PIV measurements was performed on GPU using the FOLKI-SPIV software [16]. A comparison between the acoustic amplitude and phase estimations obtained using the two measurement techniques on the test case L = 138 dB is provided in Figure 14. As observed, an excellent agreement is obtained for both quantities, with satisfactory statistical confidence intervals.

From these results, it appears that TR-PIV measurements can provide valuable estimates of the acoustic velocity components, which can then be complementary to LDV measurements. Indeed, the latter are not as restricted in terms of acquisition frequencies, but are time-consuming when a fine spatial resolution is required. Further developments on the TR-PIV approach will address near-wall measurements and sound pressure levels lower than 130 dB.

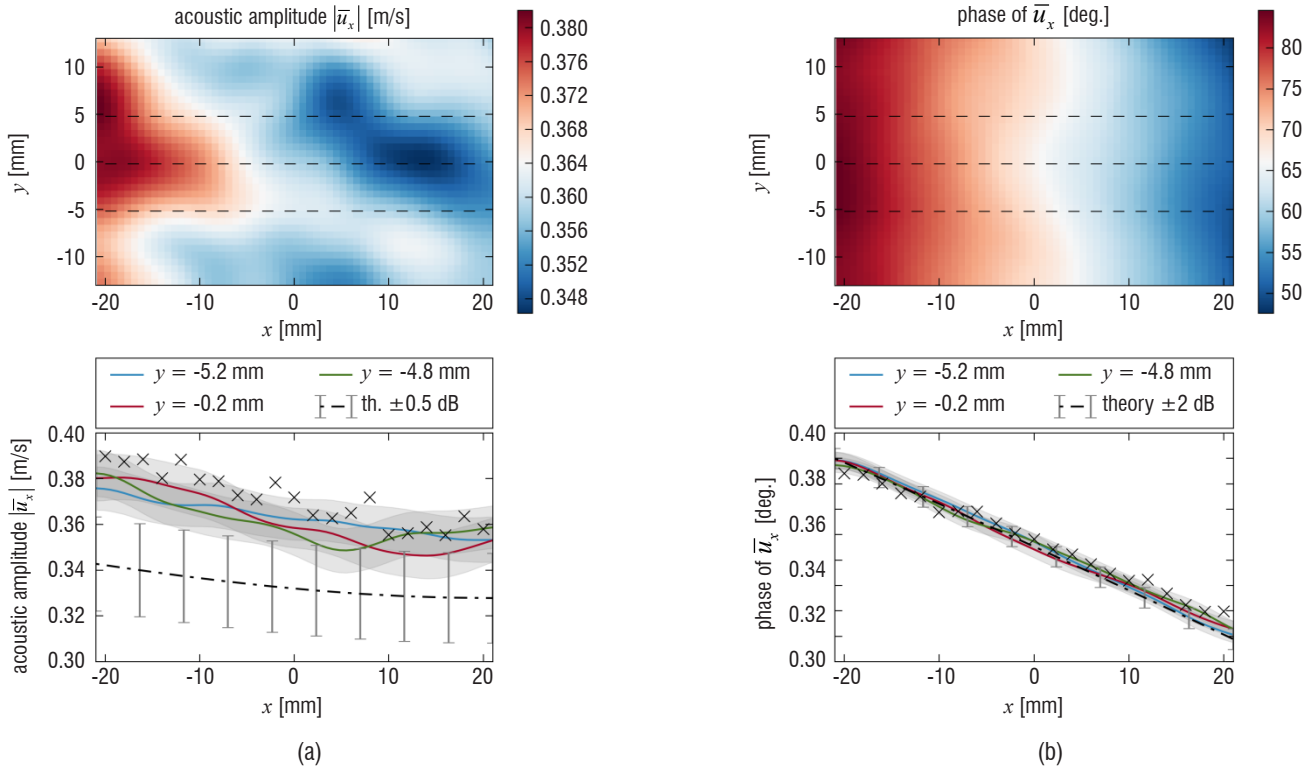


Figure 14 – Maps of acoustic (a) amplitude and (b) phase estimations yielded by TR-PIV measurements for an acoustic wave of 138 dB; comparisons with the corresponding estimations yielded by LDV (cross symbols) and the theoretical results obtained using a simple model considering plane waves in a uniform flow (dash-dot lines with 1 dB error bars) are provided in the bottom plots.

Conclusion

In this paper, examples selected from various applications studied at ONERA have been presented, in order to highlight the role played by planar PIV in aerospace research. In these studies, two-component, Stereo PIV and Time-Resolved PIV have been applied to gain better understanding of jet mixing physics, assess the efficiency of actuators to reduce helicopter fuselage drag, characterize the flow structure in a ramjet combustor, provide a precise characterization of flows in industrial wind tunnels and, finally, extract acoustic velocities from turbulent flows. In all of these applications, it is the wealth of spatial and, in the case of TR-PIV, temporal information that has been determinant in obtaining the results and contributing supplementary knowledge to the field.

Most of the results above have been obtained using well-validated approaches, which can be considered as traditional nowadays. PIV, TR-PIV and 3D PIV processing algorithms are still a very active field of research, with substantial effort being directed at estimating the measurement uncertainty, and extracting indirect and/or difficult quantities from the measurements, such as pressure fields and forces exerted on bodies, for instance. Since these constant developments contribute to enriching the trust placed in the measurement, as well as the spectrum of its capabilities, there is no doubt that the use of PIV in wind tunnels will continue to develop strongly in the next years, confirming its place as a major contributor to aerospace research ■

References

[1] M. RAFFEL, C. WILLERT, S. WERELEY and J. KOMPENHANS - *Particle Image Velocimetry, a Practical Guide*. Springer, 2007.

[2] J. WESTERWEEL, G. E. ELSINGA and R. J. ADRIAN - *Particle Image Velocimetry for Complex and Turbulent Flows.* Annual Review of Fluid Mechanics, vol. 45, p. 409-436, 2013.

[3] F. SCARANO - *Tomographic PIV: Principles and Practice*. Measurement Science and Technology, vol. 24, 2013.

[4] K. LYNCH and F. SCARANO - *A High-Order Time-Accurate Interrogation Method for Time-Resolved Piv*. Measurement Science and Technology, vol. 24, no. 3, 2013.

[5] Y. J. JEON, L. CHATELLIER and L. DAVID - *Fluid Trajectory Evaluation Based on an Ensemble-Averaged Cross-Correlation in Time-Resolved PIV*. Experiments in Fluids, vol. 55, no. 7, pp. 1-16, 2014.

[6] R. YEGAVIAN, B. LECLAIRE and F. CHAMPAGNAT - *Lucas-Kanade Fluid Trajectories for Time-Resolved PIV*. Measurement science and Technology, vol. 27, 2016.

[7] D. SCHANZ, S. GESEMANN and A. SCHROEDER - *Shake-The-Box: Lagrangian Particle Tracking at High Particle Image Densities*. Experiments in Fluids, vol. 57, 2016.

[8] B. VAN OUDHEUSDEN - *PIV-Based Pressure Measurement*. Experiments in Fluids, vol. 24, 2013.

[9] S. GESEMANN, F. HUHN, D. SCHANZ and A. SCHRÖDER - *From Noisy Particle Tracks to Velocity, Acceleration and Pressure Fields Using B-Splines and Penalties*. 18th Int. Symp. on Appl. of Laser Techniques to Fluid Mechanics, Lisbon, Portugal, 2016.

[10] J. SCHNEIDERS, I. AZIJLI, F. SCARANO and R. DWIGHT - *Pouring Time Into Space*. 11th Int. Symp. on Particle Image Velocimetry - PIV15, Santa Barbara, USA, 2015.

[11] J. F. SCHNEIDERS and F. SCARANO - *Dense Velocity Reconstruction from Tomographic PTV with Material Derivatives.* Experiments in Fluids, vol. 57, 2016.

[12] R. YEGAVIAN, B. LECLAIRE, F. CHAMPAGNAT and O. MARQUET - *Performance Assessment of PIV Super-Resolution with Adjoint-Based Data Assimilation*. 11th International Symposium on Particle Image Velocimetry - PIV15, Santa Barbara, USA, 2015.

[13] G. LE BESNERAIS - *Experimental Fluid Mechanics Goes 3D*. Aerospace Lab, 2016.

[14] S. DAVOUST, L. JACQUIN and B. LECLAIRE - *Dynamics of m=0 and m=1 Modes and of Streamwise Vortices in a Turbulent Axisymmetric Mixing Layer*. Journal of Fluid Mechanics, vol. 709, pp. 408-444, 2012.

[15] R. COURTIER - *Influence des conditions initiales et dynamique des grandes échelles dans les jets turbulents*. 2015.

[16] F. CHAMPAGNAT, A. PLYER, G. LE BESNERAIS, B. LECLAIRE, S. DAVOUST and Y. LE SANT - *Fast and Accurate PIV Computation Using Highly Parallel Iterative Correlation Maximization*. Experiments in Fluids, vol. 50, pp. 1169-1182, 2011.

[17] S. DAVOUST and L. JACQUIN - *Taylor's Hypothesis Convection Velocities from Mass Conservation Equation*. Physics of Fluids, vol. 23, 2011.

[18] C. LIENARD and A. LE PAPE - *Advancement Report on Active Flow Control Technologies for Helicopter Blunt Fuselage Drag Reduction*. CleanSky European Program: Green Rotorcraft, Report N° CS JU/ITD GRC/RP/2.2.2-5 & 2.2.4-2/32059, 2013.

[19] A. LE PAPE, C. LIENARD, C. VERBEKE, M. PRUVOST and J.-L. DE CONINCK - *Helicopter Fuselage Drag Reduction Using Active Flow Control: a Comprehensive Experimental Investigation*. Journal of the American Helicopter Society, vol. 60, no. 3, pp. 1-12, 2015.

[20] C. BROSSARD, P. GICQUEL, M. BARAT and A. RISTORI - *Experimental Study of a High Speed Flow Inside a Dual Research Ducted Rocket Combustor using laser Doppler velocimetry.* 11th International Symposium on Applications of Laser Techniques to Fluid Mechanics, Lisbon, Portugal, 2002.

[21] A. RISTORI, G. HEID, C. BROSSARD and S. REICHSTADT - *Detailed Characterization of the Reacting One-Phase and Two-Phase Flow Inside a Research Ramjet Combustor*. 17th ISABE Symposium, Munich, Germany, 2005.

[22] A. ROUX, L. GICQUEL, S. REICHSTADT, N. BERTIER, G. STAFFELBACH, F. VUILLOT and T. POINSOT - *Analysis of Unsteady Reacting Flows and Impact of Chemistry Description in Large Eddy Simulations of Side-Dump Ramjet Combustors*. Combustion and Flame, vol. 157, pp. 176-191, 2010.

[23] A. RISTORI, C. BROSSARD, T. LE PICHON - *Caractérisation de l'écoulement instationnaire dans un statoréacteur de recherche par PIV et visualisations OH* à haute cadence simultanées.* 14e Congrès Francophone de Techniques Laser (CFTL2014), Marseille, France, 2014.

[24] S. MASSEBOEUF, S. MOUTON, R. MARIANI and B. LECLAIRE - *Clinometric Measurements by Means of High-Accuracy PIV in the ONERA F1 Low-Speed Pressurized Wind Tunnel*. 13th Asian Symposium on Visualization, Novosibirsk, Russia, 2015.

[25] F. DAVID, J.-M. JOURDAN, F. CLÉRO, M. KOENIG and J. HUBER - *Large-Scale Jet Noise Testing, Reduction and Methods Validation "EXEJET": 4 - Flows Characterization with PIV in the CEPRA19 Anechoic Wind Tunnel of ONERA*. 20th AIAA/CEAS Aeroacoustic Conference and Exhibit, Atlanta, USA, 2014.

[26] Y. MICHOU, S. MASSEBOEUF, S. MOUTON, F. DAVID and J.-M. JOURDAN - *Applications of PIV Measurements in the ONERA Large Wind Tunnels*. In 31st AIAA Aerodynamic Measurement Technology and Ground Testing Conference, Dallas, USA, 2015.

[27] J. VALIÈRE - *Acoustic Particle Velocity Measurements Using Laser: Principes, Signal Processing and Applications*. Wiley, 2014.

[28] A. MINOTTI, F. SIMON and F. GANTI - *Characterization of an Acoustic Liner by Means of Laser Doppler Velocimetry in a Subsonic Flow*. Aerospace Science and Technology, vol. 112, no. 5, pp. 398-407, 2008.

[29] D. HAUFE, A. FISCHER, J. CZARSKE, A. SCHULZ, F. BAKE and L. ENGHARDT - *Multi-Scale Measurement of Acoustic Particle Velocity and Flow Velocity for Liner Investigations*. Experiments in Fluids, vol. 54, no. 17, 2013.

[30] C. HEUWINKEL, A. FISCHER, I. ROHLE, L. ENGHARDT, F. BAKE, E. PIOT and F. MICHELI - *Characterization of a Perforated Liner by Acoustic and Optical Measurements*. 16th AIAA/CEAS Aeroacoustics Conference, Stockholm, Sweden, 2010.

[31] A. FISCHER, E. SAUVAGE and I. ROEHLE - *Acoustic PIV: Measurement of the Acoustic Particle Velocity Using Synchronized PIV Technique*. The Journal of the Acoustical Society of America, vol. 123, no. 15, p. 3130, 2015.

Acronyms

PIV	(Particle Image Velocimetry)
IW	(Interrogation Window)
LDV	(Laser Doppler Velocimetry)
TR-PIV	(Time-Resolved Particle Image Velocimetry)
TKE	(Turbulent Kinetic Energy)
CFD	(Computational Fluid Dynamics)

AUTHOR

Benjamin Leclaire graduated from the *École Polytechnique*, France, in 2001. He obtained his PhD degree in Fluid Mechanics in 2006. He is now both a researcher at ONERA and an assistant professor at the *Ecole Polytechnique*. His research interests include optical metrology for fluid dynamics, and turbulent flows.

Christophe Brossard received his Engineering Degree in Energetics from the "*Institut National des Sciences Appliquées*" in Rouen, France, in 1991. He received his Doctoral Degree in Energetics from the University of Rouen in 1995. From 1996 to 2001, he worked at the Propulsion Engineering Research Center at Pennsylvania State University, USA. Since 2001, he has worked as a research scientist in the Fundamental and Applied Energetics Department at ONERA, focusing on flow field characterization, in different non-reacting or reacting environments, using optical laser-based diagnostic techniques (PIV, LDV, PDA, PLIF).

Romain Courtier graduated from the University Paris Sud 11 in 2011. He obtained his PhD degree from the *École Polytechnique* in 2015. He is now a research engineer in a Parisian startup developing wind turbines. His interests range from fluid mechanics to data science.

Frédéric David joined ONERA in 1983 and was engaged in research on experimental aeroelasticity, for which he performed numerous wind tunnel tests on launchers, shuttles, suspension bridges, fan blades, wind machine and helicopter blades up until 1991. He currently carries out aero-acoustic tests in the large CEPRA19 wind tunnel of ONERA, for airframe and jet noise acoustic characterization with far-field and PIV measurements.

Samuel Davoust graduated from the *École Polytechnique*, France, and obtained his PhD degree in Fluid Mechanics in 2011 while at ONERA. After a 4-year experience dedicated to developing new Lidar applications for Leosphere in Orsay, Samuel became a research engineer for General Electric in Munich. His interests include industrial innovation, turbulent jets, atmospheric boundary layers and wind turbine systems.

Anne Gilliot received the VKI Diploma in 1990 and her doctorate in Mechanics in 1997 from the University of Lille. Since 1991, she has been employed as a scientist by ONERA and has taught Fluid Mechanics and Turbomachinery in engineering schools. She started with experimental works on Fundamental Fluid Mechanics problems. She is now working mainly on the development of Particle Image Velocimetry (PIV) and its use in aerodynamic applications and in large industrial scale wind tunnels, as a member of the PIV team of the Applied Aerodynamics Department.

Laurent Jacquin received his Master's Degree in Mechanical Engineering (1980) and his PhD (1983) from Aix-Marseille University. He obtained his Research Habilitation Degree (*thèse d'état*) from the University of Lyon in 1987 and joined ONERA as a research engineer. He led the Fundamental and Experimental Aerodynamics Department (DAFE) from 2003 to 2015. He is currently the Scientific Director of the Fluid Mechanics and Energetic Branch of ONERA and a Professor in the Mechanics Department of the *École Polytechnique*. His research topics include: turbulence, hydrodynamic stability, vortex dynamics, compressible flows, aerodynamics and experimental methods.

Jean-marc Jourdan joined ONERA in 1990 and has been working in the large CEPRA19 wind tunnel over the last 26 years. He has taken part in many aero-acoustic tests for airframe and jet noise acoustic characterization, and recently worked on the setup of the PIV measurement technique in this wind tunnel.

Olivier Léon graduated from the *École Centrale de Lyon*, France, and simultaneously obtained his M.Sc. degree in Aerospace Engineering from Pennsylvania State University, USA, in 2009. He obtained his PhD degree in Fluid Mechanics in 2012 from Toulouse University. His research interests include fluid dynamics metrology and turbulence.

Sylvain Masseboeuf joined ONERA in 2011 as a test engineer in the F1 Wind Tunnel, which is part of the Wind Tunnel Division. He has taken part in several major low-speed wind tunnel test campaigns over the last five years, and has worked on the industrialization of some optical measurements means, such as PIV or infrared thermography.

Jean-Claude Monnier (DAAP) received his doctoral degree in Mechanical Engineering from the University of Lille. Since 1990, he has worked as a research scientist in the Applied Aerodynamics Department at ONERA. He started with experimental work on Fundamental Fluid Mechanics problems. For 25 years, he worked mainly on the development of Particle Image Velocimetry (PIV) and its use for aerodynamic applications and in large industrial wind tunnels. Since 1998, he has been in charge of the PIV team of the Applied Aerodynamics Department. He also teaches Fluid Mechanics in the Engineering School.

Sylvain Mouton joined ONERA in 2005 after graduating from ISAE-SUPAERO and he first worked in the Applied Aerodynamic Department, in which he performed numerous flow simulations and shape optimization, mostly on aircraft configurations. He was also very involved in several wind-tunnel tests of light aircrafts and UAVs at the Lille center. In 2013, he joined the Wind Tunnel Division as a scientific and technical expert, helping to continuously improve the testing methods and providing assistance and expertise to the Director and the engineering staff of the Fauga-Mauzac test center.

Estelle Piot graduated in 2005 from *École Polytechnique* in Paris, and received her PhD in Fluid Mechanics from the University of Toulouse/Supaéro in 2008. Since then, she has been a research engineer at ONERA, in charge of aeroacoustic characterization and optical measurements in flow ducts. In 2014, she became the head of the Instability, Transition and Acoustics research unit.

Arnaud Ristori received his Engineering Degree in Thermics-Energetics from the Polytech "*Institut des Sciences de l'Ingénieur en Thermique-Energétique et Matériaux*" in Nantes, France, in 1989. He received his Master's degree in Fluid Mechanics from the "*École Centrale de Nantes*" in 1989. Since 1989, he has worked as a research scientist in the Fundamental and Applied Energetics Department at ONERA, focusing on blade cooling studies and air-breathing propulsion for aero-engines and rockets, performing both experimental studies on test rigs and numerical studies). In 2006, he joined the team in charge of optical diagnostics in the same Department, for flow field characterization in various different non-reacting or reacting environments, using optical laser-based diagnostic techniques (PIV, LDV, PDA, PLIF).

Delphine Sebbane joined ONERA in 2010 as a technician in the Aerodynamics and Energetics Modeling Department in Toulouse. She is in charge of acoustics measurements in the Instability, Transition and Acoustics research unit.

Frank Simon, graduated with a degree in Aeronautical Engineering (1989), a PhD in Mechanical Engineering (1997) and a Habilitation for Supervising PhD students (2007) from the University of Toulouse. He is a Research Master at ONERA, with the Qualification of University Professor, specialized in the development of aerooacoustic measurement techniques, and has participated in various different EC projects in the aeroacoustic domains (i.e., TIMPAN, and JTI GRA), and has collaborated with NASA in the characterization of liners for engine nacelles.

C. Besson, A. Dolfi-Bouteyre, G. Canat, N. Cézard, B. Augère, A. Durécu, L. Lombard, M. Valla, A. Hallermeyer
(ONERA)

E-mail: claudine.besson@onera.fr

DOI: 10.12762/2016.AL12-08

Doppler LIDAR Developments for Aeronautics

Detecting atmospheric hazards such as wake vortices, turbulences, wind shear or wind gusts would improve flight and airport traffic safety. Moreover, to face the current and future challenges of air traffic growth in terms of environment, safety and costs, one of the major challenges is to increase runway throughput by optimizing the distance separation between aircraft. These are some of the issues, among others, addressed by Doppler LIDARs. This powerful technique can provide remote air speed data and is being deployed on the ground or in flight.

The paper presents some of the recent Doppler LIDAR developments at ONERA. Some insight into state-of-the-art technology is provided, as well as perspectives for future applications in aeronautics.

Introduction

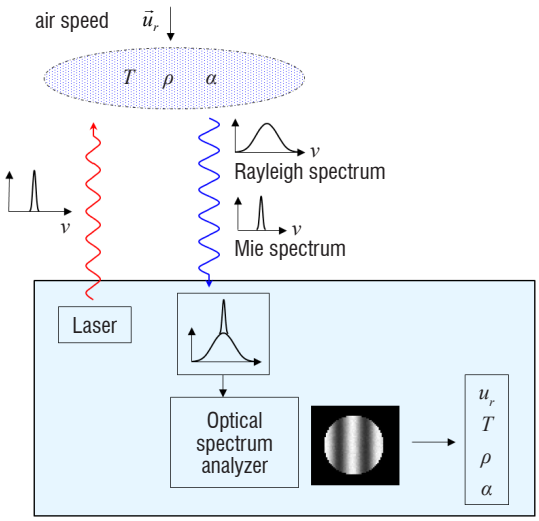
Aeronautics being the science of navigation in the air, measuring air motion remotely from a ground station or from an aircraft is therefore a major issue. In the vicinity of an airport, for instance, wind mapping may allow dangerous gusts and wind shears to be detected. Above airport runways, detecting and tracing wake vortices generated by aircraft landing and taking off can increase flight safety and optimize airport operation. From an aircraft, remote wind measurement can provide information about the dynamic state of the atmosphere that the aircraft is about to fly into. This can enable the cross-checking of classical Pitot tube measurements and improve flight security. Long-range wind measurement ahead of an aircraft could also allow severe clear-air turbulence events to be mitigated or avoided.

LIDAR (Light Detection and Ranging) are ideal tools for remote wind measurements. LIDARs emit beams of monochromatic light toward the desired measurement area in the atmosphere. Photons are scattered along the beam by dust particles (Mie scattering) and molecules (Rayleigh scattering) contained in the air. A very small fraction of light is 180°-backscattered toward the laser emitter, beside which a receiver is placed. Particles (and molecules) are so light that their movement perfectly follows air motion and they can be used as air speed tracers. Due to the air mass motion relative to the LIDAR emitter/receiver, backscattered light collected by the receiver is frequency-shifted by the Doppler effect by a quantity $\Delta f = 2V_R / \lambda$, where V_R is the so-called radial wind speed (speed projected on the laser line of sight) and λ is the laser wavelength. Doppler LIDARs analyze the frequency of the backscattered signal to obtain the air speed.

This paper deals with wind LIDAR developments at ONERA for aeronautics applications. Most developments imply Doppler LIDARs, except for long-range clear air turbulence detection. In a first part, we focus briefly on high-altitude wind speed measurements. Such applications have a high potential for in-flight aeronautics, but technological issues make these developments very challenging. In a second part, we present Coherent Doppler LIDAR, which is the mainstream technology for wind speed measurement from a ground station or from low and medium altitude aircraft.

Direct Detection Doppler LIDAR

At high altitude, since particles are sparsely distributed above the atmosphere boundary layer, a LIDAR detects the signal backscattered by air molecules. The analysis is performed in the optical domain and the LIDAR is called "direct-detection" Doppler LIDAR. The spectrally broad Rayleigh-Brillouin molecular backscatter involves a receiving channel based on an optical spectral filter and a detector, which obtains the amplitude or the shift of the spectrum by flux or



Detected spectrum properties

- Central frequency (Doppler effect) → air speed u_r
- Rayleigh broadening → temperature T
- Rayleigh signal energy → air density ρ
- Mie signal energy → mixing ratio α

Figure 1 – Direct detection Rayleigh-Mie LIDAR principle

fringe pattern analysis (see Figure 1). Such systems generally require medium linewidth (<1 GHz) monomode lasers emitting at small wavelengths (UV or visible), since the backscattered signal decreases as λ^4 (λ being the LIDAR wavelength) [13].

For atmospheric science applications, Direct Detection Doppler LIDARs are used for studies in connection with climate issues, for example, aerosol transport characterization and stratosphere/troposphere mixing layer observations. Such LIDARs have long been developed by atmosphere science laboratories such as LATMOS, which was a pioneer of the technology in France (see for example [6][16][22]). Ground-based direct Detection Doppler LIDARs for high-altitude observations require high-power laser (>100 mJ/pulse), large telescopes (~1m). In France these systems are installed at the *Observatoire de Haute-Provence* and the *Observatoire du Maido* (Reunion Island). These high-performance tools will soon be supplemented by a space-based Direct Detection Doppler LIDAR called ALADIN (Atmospheric Laser Doppler Instrument). ALADIN is the LIDAR payload of the ADM-Aeolus mission, a program launched 15 years ago by ESA. It will provide direct measurements of global wind fields from space. The purpose of the mission is to demonstrate improvements in numerical weather prediction and climate models. The LIDAR operates in the UV (355 nm) [45]. The optical receiver combines a "Mie" channel (Fizeau interferometer analyzing winds from low-altitude aerosols and clouds) and a "Rayleigh" channel (double edge Fabry-Perot interferometer analyzing air molecules). The system is currently being qualified and should be launched in 2017.

In the field of aeronautics, airborne Direct Detection Doppler LIDARs are tested to evaluate the possibility of short-range turbulence detection and mitigation. ONERA has built a short-range ground prototype demonstrating the feasibility of robust and daytime measurements with a Michelson fringe-imaging technique for Doppler analysis [14]. Similar developments have been made and promoted until their demonstration in flight throughout the EU project AWIATOR [48] [56], which used the Fabry-Perot fringe-imaging technique.

Another strong concern calling for Direct Detection LIDARs in aeronautics is the long-range detection of clear air turbulences (CAT). Aircraft experience turbulence when they encounter a vertical airflow that varies on the horizontal length scales in the range of 30 m to 2 km for large commercial aircraft. Vertical airflow on these scales is associated with two distinct physical mechanisms: wave breaking caused by shear instabilities in clear air (Kelvin-Helmholtz instability) and convective updraughts and downdraughts in and around clouds and thunder storms. Clear-air turbulences are difficult to observe because of their relatively low occurrence. The physical phenomenon has been, up to now, poorly understood and they are not well predicted. ONERA has been involved in two projects aimed at collecting and analyzing turbulence data: the MMEDTAC project [27] aimed at acquiring turbulence data with the LATMOS ground-based Rayleigh LIDAR located at the *Observatoire de Haute-Provence*. The DELICAT project involved a Rayleigh LIDAR assembled by the DLR/IPA, installed on board the NLR Citation aircraft for in-flight data acquisition [36]. Both of these projects were based on measuring the density fluctuations that are linked to strong vertical winds in atmospheric layers with high stability, rather than measuring the wind speed itself. It must be noted that, with this method, the atmosphere stability characterized by the Brunt-Väisälä frequency is assumed to be constant at the scale of measurement. The principle is to measure the amplitude fluctuations of the LIDAR signal due to a fluctuation in the molecule contribution to the backscattered signal. Particle contribution must thus be rejected, to avoid interference and spurious detections. Flying the system onboard adds extra complexity, since an agile filter is required to remove Mie scattering that is Doppler shifted by the varying aircraft speed. Particle signal rejection was performed by an optical filter based on polarization separation, which assumes that at the flight test altitude (30,000ft) particles are essentially polarizing ice crystals and ashes. CAT detection was considered to be positive only in areas with a low-level particle signal. The estimation of relative fluctuations of the backscatter signal is assumed to not be modified in the presence of low-level particle scattering due to a well-mixed aerosol layer, the ratio between aerosol and molecular backscattering remaining constant inside this layer. Although it was not possible to ensure that the atmosphere was totally free from aerosols, both projects enabled likely CAT data to be collected. New algorithms performing CAT retrieval have been developed for these two systems and will be used for future work on CAT characterization [28].

Rayleigh LIDAR involves cumbersome lasers and optics and complex key technology (e.g., agile filters), which explains why the technique is still at the level of demonstration for airborne sensors. Technology maturation is necessary to comply with aeronautics size and consumption constrains. However, since it relies on molecule scattering, Rayleigh LIDAR would be a reliable concept for operational sensors to detect atmospheric hazards at any altitude with reasonably easy certification procedures.

High power coherent Doppler LIDAR

Unlike direct detection LIDARs, coherent detection LIDAR rely solely on the presence of particles. Particles are dense in the lower atmosphere, *i.e.*, from the ground up to a few kilometers, especially in the atmospheric boundary layer. At low altitudes, remote sensing of the air speed is also an important issue. Indeed, during landing and take-off, a minimal distance separation between aircraft is

necessary, in order to avoid the risk of encountering the wake vortex from a preceding aircraft. Wake vortices are two coherent counter-rotating flows created behind the aircraft wings and they induce a potentially dangerous rolling moment to the following aircraft. ICAO wake turbulence separation minima defined forty years ago are being redefined, in order to take into account airport traffic increase. Indeed, the strength and lifetime of wake vortices vary a lot with weather conditions. Their dissipation rate varies depending on the atmospheric turbulence level. They can also be transported out of the way towards incoming traffic by cross-winds. Other air disturbances, such as wind gusts or rapid change of the incoming wind direction, are also detrimental to the airport traffic flow. Thus, anticipating these phenomena in the vicinity of airports is key information for air traffic optimization and safety.

Wake vortex locations and trajectories, wind turbulence level or wind maps can be used during certification or flight tests. They can be provided by long range-range resolved Doppler LIDAR or Radar, which measure the wind speed with a high spatial resolution [31][54]. Such sensors have been evaluated for airport safety and re-categorization purposes within the framework of various projects, such as CREDOS, FIDELIO, SESAR or UFO [19][20][55].

Not only airport safety, but also flight tests of new air vehicles, manned or unmanned, could benefit from accurate knowledge of air dynamics disturbances in the vicinity of airports. Due to recent progress in fiber lasers and amplifiers, 1.5 μm fiber LIDAR technology is becoming a serious candidate for lightweight, compact, eye-safe airborne systems. Demonstrators have been developed and flight-tested for new avionics or certification aids [3][51].

In this chapter, we review recent LIDAR achievements at ONERA and report on performance and field test results for two types of coherent Doppler LIDARs: ground based range resolved LIDAR and an airborne true air speed sensor.

In coherent detection LIDAR, the spectrally narrow Mie backscatter from aerosols and cloud particles is analyzed via the mixing of the backscattered signal and a reference wave (local oscillator) to produce a radio frequency beat note, which can be temporally analyzed with a standard photodiode. The mixing also enhances the detection sensitivity, thanks to the optical product of the signal beam and the reference beam, and enables very low signal amplification. Compared to the direct detection LIDAR, the receiving channel of the coherent detection LIDAR is much simpler.

Although the technology has been used for decades, over the last ten years a renewed interest has arisen thanks to the availability of fiber components developed for the telecommunication market. Indeed, LIDAR based on fiber technology are low-cost and resilient to vibrating environments. When emitting around 1.5 μm, they can benefit from telecom industry components with a large market at competitive costs and increased reliability. They offer simplified maintenance procedures compared to free-space technology and enable compact system designs. It is thus well suited to on-the-field operation for other domains besides aeronautics, such as defense, security or space [5][9][18].

Range-resolved wind LIDAR

Range resolution can be achieved with frequency modulation (FMCW) or pulsed emission. Whatever the technique, the fiber laser architecture based on a Master Oscillator Fiber Power Amplifier (MOFPA) offers the advantage of versatility: FMCW emission, pulse shape, duration, repetition rate and frequency characteristics can be varied from pulse to pulse at 100s of kHz rates. This unique "smart pulse" feature provides multifunction capability and perspectives for applications, such as pilot aid, with simultaneous ranging and ground velocity measurement or anemometry and velocity, gas detection with multi-wavelength DIAL LIDAR (DIfferential Absorption LIDAR), simultaneous DIAL-Doppler LIDARs, etc.

A typical coherent fiber LIDAR set up is depicted in Figure 2. The master oscillator is a continuous laser diode. Its output is split using a 50/50 fiber coupler. The transmitter channel goes through an acousto-optic modulator (AOM), which shapes the pulse and eventually sets an intermediate frequency (IF). It is then amplified through a fiber pre-amplification stage and a filter (BPF) and a booster stage. At the output of the booster stage, the beam is polarized, quasi-single mode, and its temporal shape is optimized for efficient coherent detection. A passive fiber pigtail (1) is used to connect to the polarization beam splitter (PBS). The polarization beam splitter enables the transmitter channel and the receiver channel to circulate thanks to a quarter wave plate. Up to the beam splitter, all components are fibered. At that point, for high-peak-power emissions, a Brillouin effect may occur in fibers and free space optics are used (Paragraph 2 elaborates on the Brillouin effect). At the telescope output, the signal beam is emitted into the atmosphere boundary layer. The signal backscattered from natural aerosols is coupled to the fiber-beam splitter pigtail (2) and mixed with the local oscillator using a fiber coupler. The electric signal of the photodetector (DET)

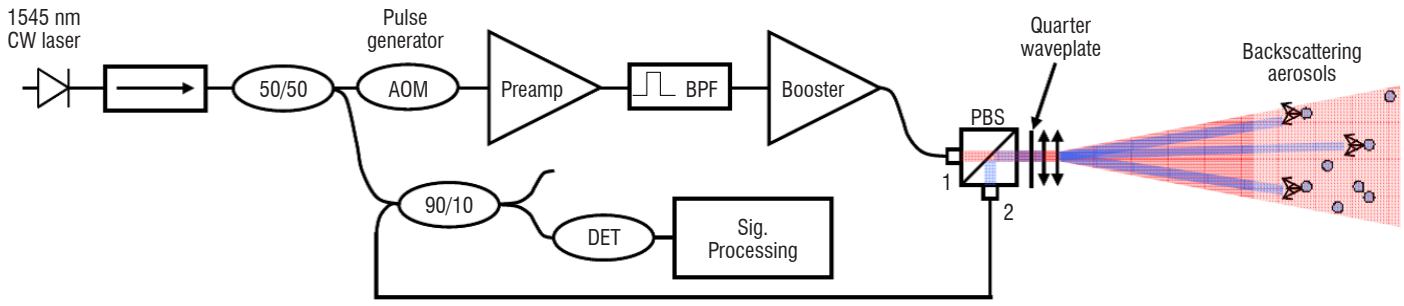


Figure 2 – MOFPA coherent fiber LIDAR setup

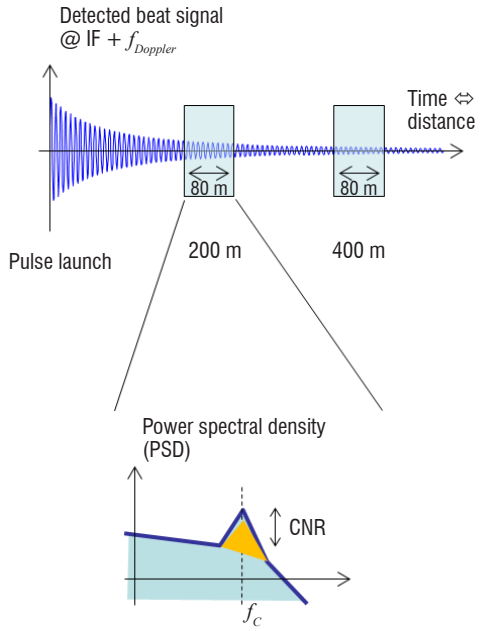


Figure 3 – signal detection

is digitized and processed in real time. The detected signal is split into temporal (or spatial) bins, which are Fourier transformed (see Figure 3). The maximum frequency of the spectrum fc is estimated with a suitable algorithm (maximum, centroid, etc.). Each bin is thus defined by its spatial coordinates and its central frequency enabling, for example, wind map reconstructions.

Most systems emit at around 1545 nm, which corresponds to high gain in erbium doped fibers and a good atmospheric transmission. The monostatic set-up (single pupil for transmitter and receiver channels) sketched here is convenient for small-size scanning systems.

Wake vortex monitoring

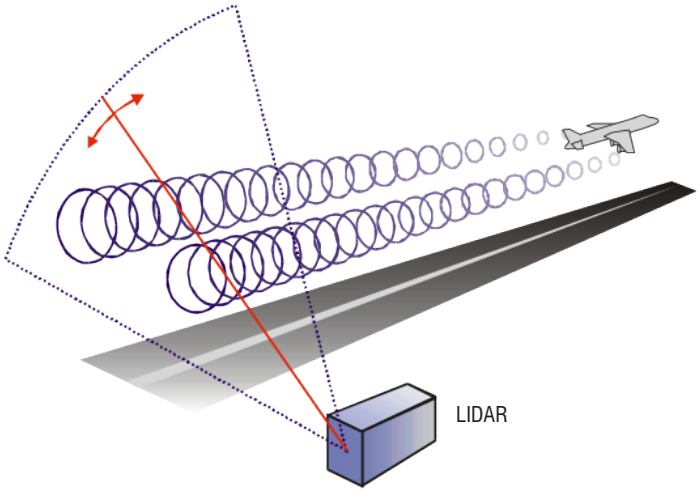
For wake vortex monitoring from the ground, e.g., above runways, measured tangential velocities are high when the LIDAR scanning pattern is in a plane perpendicular to the aircraft path (Figure 4 left). The wake vortex LIDAR signature shown in Figure 4, right, is typical with four patches of positive negative velocities around the vortex cores. Wake vortex positions can be traced from scan to scan and their trajectory retrieved.

The detailed principle and results achieved at ONERA can be found in [55]. Since 2009, operational use of ONERA technology has occurred thanks to technology transfer to a SME, Leosphere. Commercial LIDARs are deployed for windmill prospection and wind turbine monitoring. They are also operated in a number of airports to analyze wake vortex parameters and their dependency on weather and aircraft characteristics. Improvement of the technology is ongoing jointly at ONERA and Leosphere and progress has been made on the hardware and software. Recent achievements in signal processing are described in the following paragraph.

Signal processing for wake vortex characterization

The main sought features of a wake vortex are its strength, characterized by its circulation, and its position. In the literature, two notable families of algorithms aimed at obtaining these features from the LIDAR velocity spectra can be found. The first one [4] is a non-parametric method, based on extracting the positive and negative velocity envelopes using a spectrum threshold that depends on the Signal-to-Noise ratio (SNR) and the aircraft category, and thus needs to be adapted by the user. In the velocity spectra, every velocity value encountered by the laser pulse in the measurement volume is represented. The maximum (and minimum) velocity is the velocity of the wake vortex's field line tangent to the laser beam. The extrema of each velocity spectrum gives the wake vortex envelope (positive and negative). These envelopes provide the vortex positions (located where the positive envelope has its maximum, and the negative envelope has its minimum). To compute the circulation, various methods are exploitable, such as integrating these envelopes at the range of the vortex core between 5 and 15 meters. In this case, the circulation is called Γ_{5-15}.

The second family gathers the algorithms, making use of estimators such as the maximum likelihood estimation on the spectra from a Wake Vortex analytical model [24], or as the maximum of a correlation function between the radial velocities and a vortex model. A new method based on a hybrid algorithm has been developed during PhD work jointly supervised by ONERA, Leosphere and Supelec [26]. The purpose is to accurately and precisely estimate the circulations



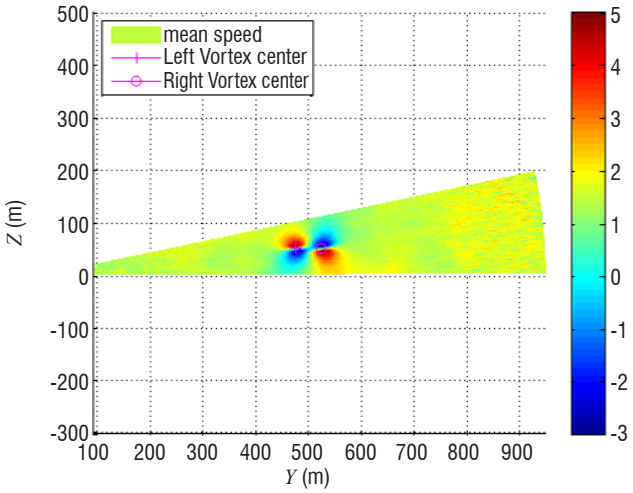


Figure 4 – LIDAR scanning pattern for lateral wake vortex detection, (left) scan pattern for CREDOS and SESAR campaigns; (right) LIDAR measurement and display of average wind field, perturbed by the wake vortex, projected on LIDAR line-of-sight

with a processing time as short as possible, so as to make this algorithm exploitable for operational projects. This is why a method has been set up to evaluate its precision and its robustness. The results of tests on simulated scenarios of different aircraft vortices and different weather conditions show that this algorithm is able to precisely locate wake vortices and to accurately estimate their circulation: the root mean square error is below 5%. In order to model the complex old vortex in the ground effect, large eddy simulations performed by UCL have been used. In this case, the error reaches 20%.

High spectral brilliance all fibered sources and SBS mitigation

The assets of fiber lasers are their high efficiency enabling a wall plug efficiency of 10% to 15%; their size – the volume of a coherent 2 W fiber laser is the size of a thick book – and their reliability due to spliced connectors.

ONERA has developed specific know-how in high-peak power, narrow linewidth pulsed fiber lasers and amplifiers. The additional laser power provides increased coherent LIDAR capability in the range and scanning of large areas, as well as better system resistance to adverse weather conditions.

In order to reach long ranges within a short acquisition time, coherent wind LIDARs require high-power (~kW), narrow-linewidth (few MHz) pulsed laser sources with nearly Fourier transform limited pulse duration (~1 μs). Eyesafe, all-fiber laser sources based on MOPFA (master oscillator, power fiber amplifier) architecture offer many advantages over bulk sources, such as low sensitivity to vibrations, efficiency and versatility. In the MOFPA coherent fiber LIDAR set up shown in Figure 2, the pulses are amplified in a sequence of fiber amplifier stages with increasing pump power and core size, which are separated by isolators and band-pass filters. MOFPA designers seek fibers with low numerical aperture (NA), which is favorable to single-mode guiding and low amplified spontaneous emission (ASE), in order to achieve high system efficiency.

Narrow linewidth pulsed MOPFA are usually limited in peak power by nonlinear effects arising in the fiber, such as Stimulated Brillouin Scattering (SBS) to a few 10s-100s Watts in standard fibers [7]. In standard single-mode fibers, the SBS threshold is reached when the peak power length product reaches about 80 W.m. The high-power density generates acoustic waves in the fiber core, which acts as a grating. The light transmitted by the fiber saturates, and part of this light is backscattered by the grating. In fiber amplifiers, the backscattered light may be amplified in the active fiber generating high-peak-power backscattered pulses potentially harmful for upstream components [33].

The strength of the Brillouin nonlinearities in a fiber can be quantified by the intensity length product:

$$B = g_B \frac{\int_0^L P(z)\,dz}{A_{eff}} = \frac{g_B P(L) L_{eff}}{A_{eff}} \quad (1)$$

where g_B is a Brillouin gain, L is the fiber length, A_{eff} is the effective area of the fundamental mode, $P(z)$ is the signal power distribution along the fiber, and L_{eff} is the effective length. For a passive fiber the effective length is equal to the fiber length and for an active fiber the effective length depends on the fiber gain along the fiber [30]. From (1) it appears that there are three ways to act on the nonlinearities.

Increase in the mode area

The first way is to increase the fundamental mode effective area. Large-Mode-Area (LMA) fibers, with a large core diameter and low numerical aperture (NA) may be used to increase the mode area, but the fiber tends to guide higher-order transverse modes degrading the beam quality. A good spatial quality is crucial for many applications. For instance, in a coherent detection LIDAR when the beam quality factor M^2 equals 1.7 we expect the Carrier-to-Noise Ratio (CNR) to decrease by 3 dB [12]. The most efficient doped fibers today are based on the Erbium-Ytterbium codoping of phosphosilicate glasses. This composition has two important limitations. First, the core refractive index is quite large and special pedestal structures are needed to reduce the NA to 0.09 and limit the proportion of higher-order transverse modes. A second limitation is that the core itself is usually inhomogeneous with a central dip in the refractive index profile resulting in a poor beam spatial quality.

To go beyond this core diameter limit, IPHT and ONERA have built 40 μm core diameter multifilament-core fibers (MFC) to amplify high-peak-power pulses [8]. In this technology, the fiber core is replaced by a microstructured core composed of 37 Erbium-Ytterbium doped filaments surrounded by fluorine doped silica (Figure 5). It was shown that this structure behaves as an effective singlemode fiber, since the light only perceives an average index value for the structured core.

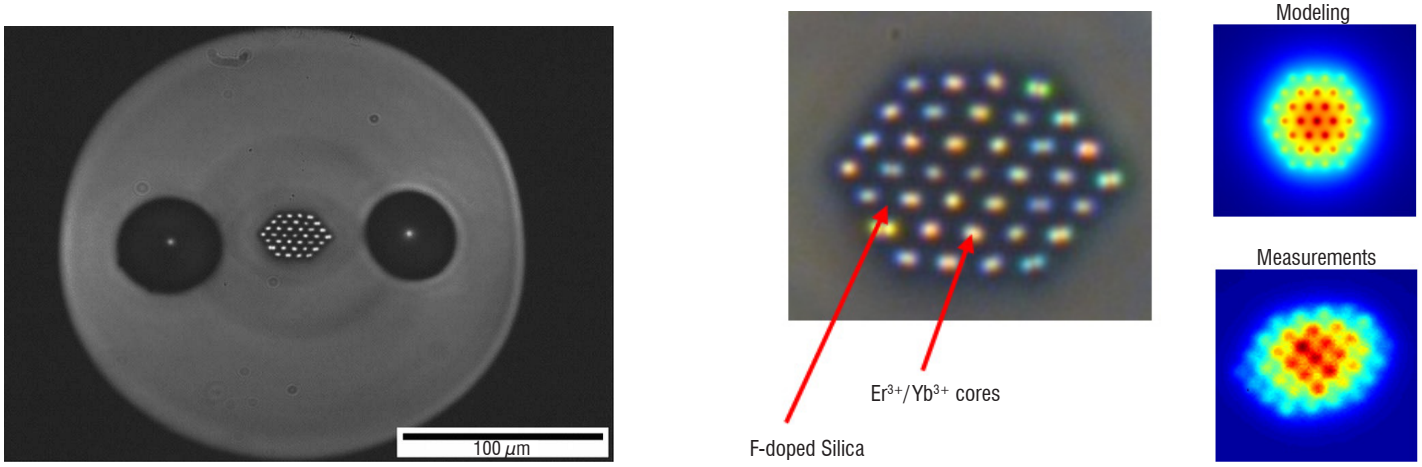


Figure 5 – Left, middle: View of the multifilament core fiber cleave. Right: mode field distribution

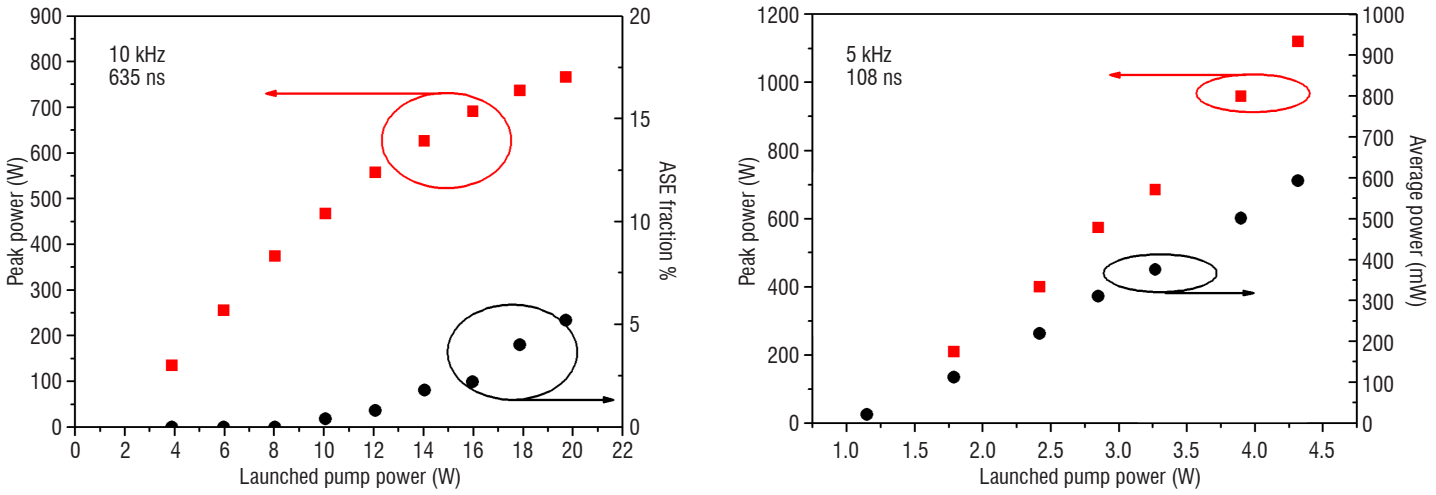


Figure 6 – Pulse peak power with respect to launched pump power in the pulsed setup. Left: pulse duration = 635 ns; PRF = 10 kHz. The ASE fraction in total power is shown on the right axis. Right: pulse duration =108 ns; PRF = 5 kHz. The average power is shown on the right axis

Pulses with 940 W peak power, 1 μs duration and 1 MHz laser linewidth were achieved. The beam quality is good with $M^2 \sim 1.3$. However, this laser setup was not all-fiber and MFC fiber fabrication and integration in all-fiber lasers systems are challenging.

Another method to increase the core size while preserving a good beam quality is to change the core composition. We have tested phosphoaluminosilicate glass fibers (PAS). It has been previously demonstrated that the simultaneous codoping with alumina and phosphorous results in the reduction of the refractive index, leading to lower NA cores [2][53]. We have tested a double cladding PAS Erbium-Ytterbium doped fiber with a 30 μm diameter core with a pedestal and 300 μm inner cladding. The core NA is 0.09 with respect to the pedestal. This fiber can be highly doped and we managed to make an efficient power amplifier with a short piece of fiber only 2 m long. When the pump power is increased, the pulse peak power increases steadily up to a 770 W peak power corresponding to a 650 ns pulse energy of 450 μJ (Figure 6 left). The ASE fraction in the total output power is 6% at that point, and the slope efficiency is 23%. For 108 ns pulses, the peak power is increased up to the SBS threshold at 1120 W (Figure 6, right). The ASE fraction is then below 1%.

Strain distribution

A second way to decrease the Brillouin effect is to minimize the integral of the power over the fiber length. By definition, this term is also equal to the product of the fiber output power and the effective length Leff. In a standard fiber, the entire fiber length contributes to the Brillouin reflectivity at the same characteristic frequency. A strain gradient applied along the fiber can be used to make each fiber segment contribute to different frequencies and spread the reflected light spectrum [17], thus allowing a higher optical peak power in the fiber before reaching the Brillouin threshold.

We have recently developed a high-power single-frequency all-fiber MOPFA using high reliability passive components and this strain technique. This MOPFA was used in the first experimental demonstration of a fiber laser based wind LIDAR with a range greater than 10 km [46]. We have been able to spread the light reflected by a 25 μm core fiber resulting in an increase by more than 3 dB in the SBS threshold. Without the strain gradient, SBS limits the pulse peak power to 300 W (200 μJ per 600 ns pulses). The M^2 is measured to be less than 1.3 in both axes. With the strain gradient applied to the doped fiber, peak power up to 630W was achieved (350 μJ per 550 ns pulses) (see Figure 7).

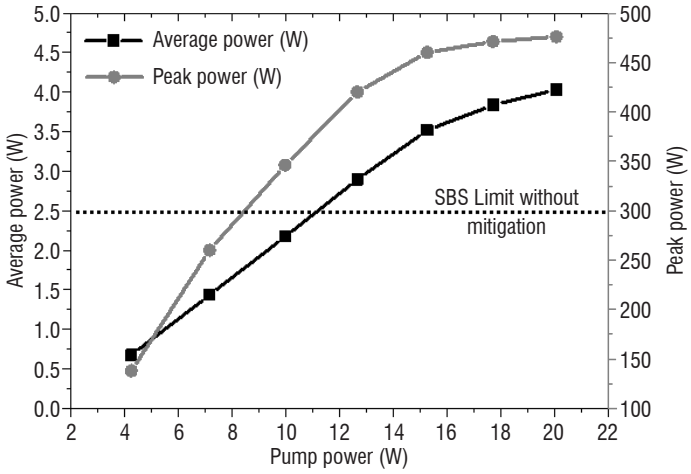


Figure 7 – Changes in the average power and peak power out of the high-power amplifier implementing the fiber straining technique. The dotted line corresponds to the amplifier maximum peak power without strain implementation

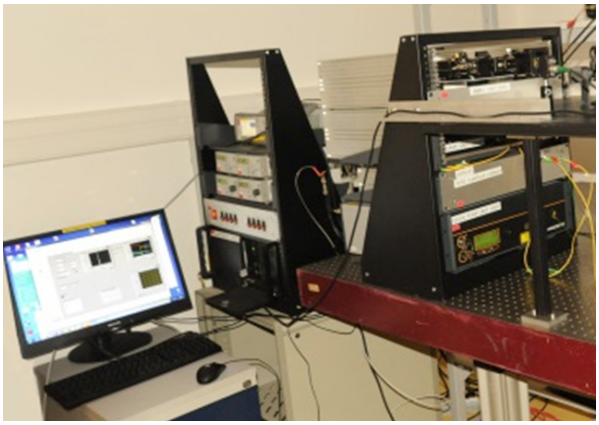


Figure 8 – The fiber laser integrated engineering model of HEPILAS for CO_2 monitoring from space

Since then, this technique has been applied to design HEPILAS, a high power laser emitting at 1579 nm, for space applications (see Figure 8). The aim of this ESA project was to study a proof of concept for monitoring CO2 from space with a Differential Absorption LIDAR (DIAL) based on an all-fiber laser. 1.7 kW peak power was obtained for 150 ns pulses with excellent beam pointing stability.

A factor of 3 gain was obtained with a modest value of strain, thereby ensuring that the fiber lifetime was compatible with a 5 year duration space mission. The laser was also designed to withstand 10 kRad radiation, thanks to iXfiber radhard active fiber [10].

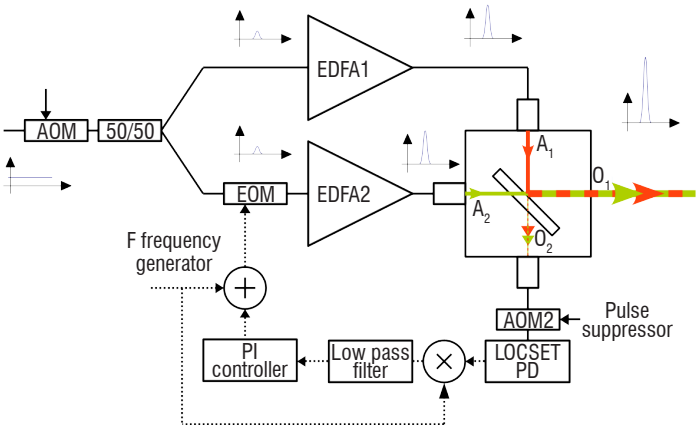


Figure 9 – Coherently combined pulsed amplifier setup. AOM, acousto-optic modulators; EDFA, erbium doped fiber amplifiers; PD, photodiode. The input signal is split into two beams, amplified separately in two EDFAs, combined into output O1, and sent into the atmosphere. A coherent pulse combination at output O1 is maintained through the minimization of output O2 by the LOCSET phase control system [49] modified for pulse operation. A signal leak between the pulses is used for phase measurement. The pulse suppressor before LOCSET PD is necessary to avoid the PD saturation by the high pulse power [35]

Coherent combination

The third strategy to tune equation (1) is to coherently add up the energy emitted by several fibers limited by SBS, while preserving the good beam quality. Coherent Beam Combination (CBC) allows the output power of MOPFA sources to be improved by adding the outputs of single amplifiers. To achieve this, two or more amplifiers seeded by the same pulsed oscillator are coherently combined into a singlemode beam. To achieve high CBC efficiency, the phase differences are compensated using a controller. We have recently demonstrated the CBC of two amplifiers in 100 ns-pulse regime, limited to 95 and 123 W respectively, resulting in a peak power of 208 W [35]. Beam quality and spectral linewidth were maintained. This source has been successfully tested in a LIDAR configuration [37]. The setup of the CBC of two pulse amplifiers is shown in Figure 9.

In particular, we demonstrated that the pulse operation CBC phase controller does not impair the LIDAR performances. Indeed, the source of Figure 9 has been used as a Booster in a LIDAR Setup with 240 ns pulse duration. We have compared the LIDAR performances when running (A) one amplifier at full power, (B) two amplifiers at half power and (C) two amplifiers at full power. The time recorded estimated wind speed is shown in Figure 10. The Power spectral densities (PSD) of each recorded time evolution are reported in Figure 11.

The extracted noise floors (0.21 m/s for Mode A, 0.19 m/s for Mode B and 0.08 m/s for Mode C) show that no performance degradation is

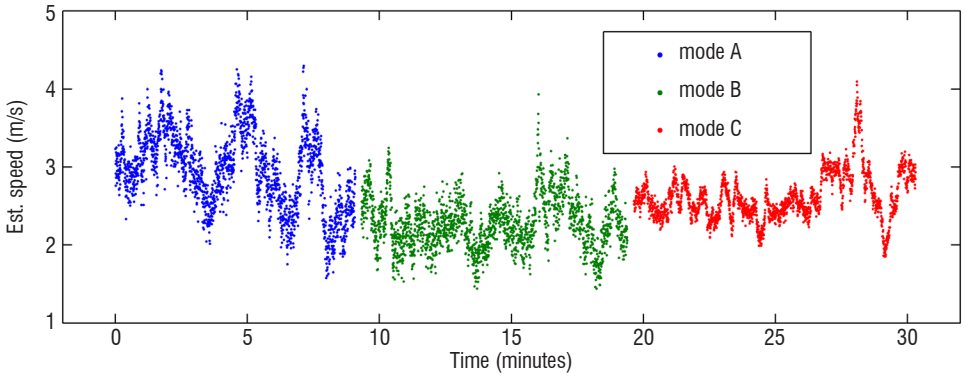


Figure 10 – Estimated wind speed, as recorded by the wind LIDAR in Configurations A, B and C

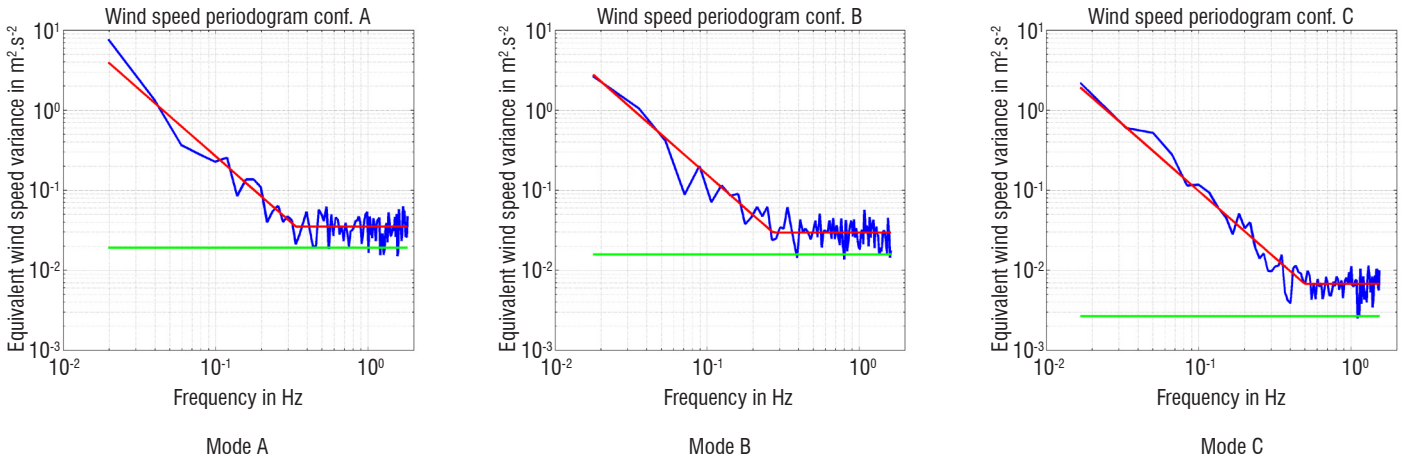


Figure 11 – LIDAR performance comparison at 150m – Instrument noise floor comparison; blue: PSD of estimated wind speed from Figure 8; red (low freq.): $\alpha f^{-5/3}$ Kolmogorov law (α fitted to data); red (high freq.): LIDAR noise floor fitted to data; green: LIDAR theoretical noise floor from estimated CNR (Cramer-Rao bound)

measured when using CBC. When both amplifiers are run at full power, the expected LIDAR performance improvement is reached. In all cases, the LIDAR performance is very close to its optimum value, given by the Cramer-Rao bound. These results show the compatibility of coherent beam combination with coherent wind LIDAR. Potentially, tens to hundreds of those amplifiers can be combined by this technique.

Long-range LIDAR measurements

These novel, high spectral brightness, all-fibered sources have been used in various LIDAR configurations to achieve record ranges and EDR measurements [38]. The ONERA wind LIDAR system LICORNE (shown in Figure 13) is a mobile platform dedicated to field tests of new lasers, components, architecture or signal processing for coherent LIDAR. Sample results are illustrated in Figure 12 in a fixed line-of-sight configuration: the wind-speed spectrogram shows that ranges beyond 15 km are achieved with 0.1 s averaging time.

The CNR and the estimated wind-speed are obtained after averaging frequency spectra over 1024 pulses, corresponding to 0.1 s acquisition time. The CNR map for vertical scan shown in Figure 13 (top left) illustrates the measurement capability inside the boundary layer up to 1200 m, and inside a cloud layer between 1500 and 2000 m. A horizontal scan of Figure 13 (top right) shows the range limit of 11 km.

High-power lasers enable the design of fast scanning LIDARs that are required for long-range wind mapping at airports. During the UFO European project, a 1.5 μm high-power pulsed fiber amplifier was integrated into a Leosphere Windcube 200S® (see Figure 14 left). A two-month campaign was performed at Toulouse-Blagnac airport in April to May 2014, in order to assess the LIDAR performances. A variety of other sensors were also deployed (Thales: 2D Electronic Scanning X-Band RADAR, Vertical X-Band RADAR, Leosphere: vertical 1.5 micron Doppler wind profiler) to study a Wake-Vortex Advisory System for weather-dependent distance separation concepts [52].

The scanning LIDAR was configured in order to perform wind and EDR measurements every km along the glide path in less than one minute. A sequence of several horizontal scanning patterns, as illustrated in Figure 14, ensures the theoretical accuracy of wind and EDR retrievals and the fastest update rate possible along the glide path and around the airport. Typical accumulation times of 0.16 s were used per line of sight, in order to allow the measurement range of the Windcube scanning LIDAR up to 10 km.

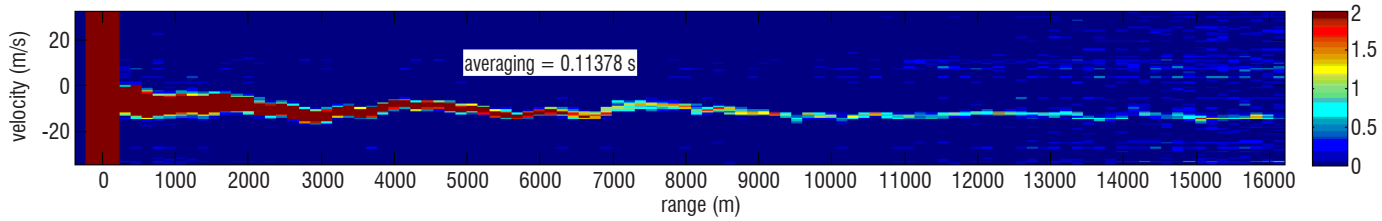


Figure 12 – Wind-speed spectrograms versus range. Color scale is the power spectral density in 0.1 s average time

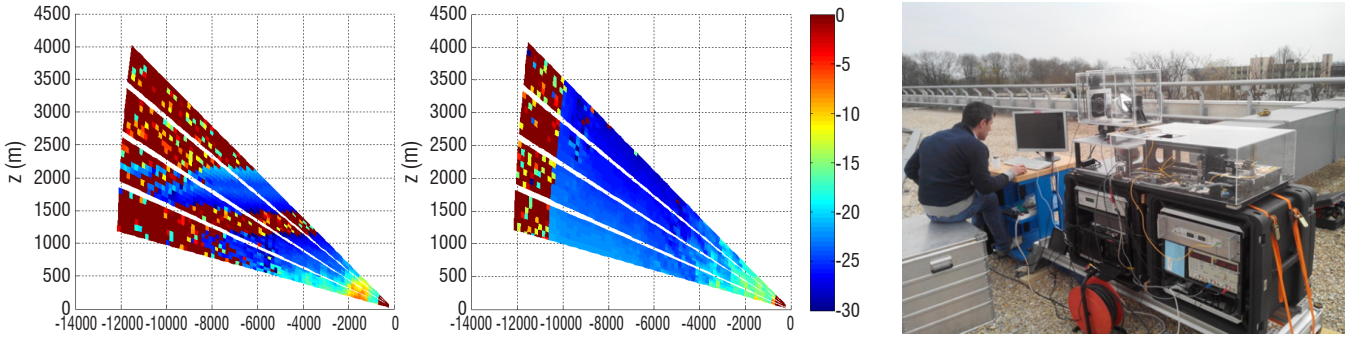


Figure 13 – Measurement in the scanning configuration: left: CNR map for a vertical scan; middle: CNR map for an azimutal scan; right: ONERA LICORNE LIDAR

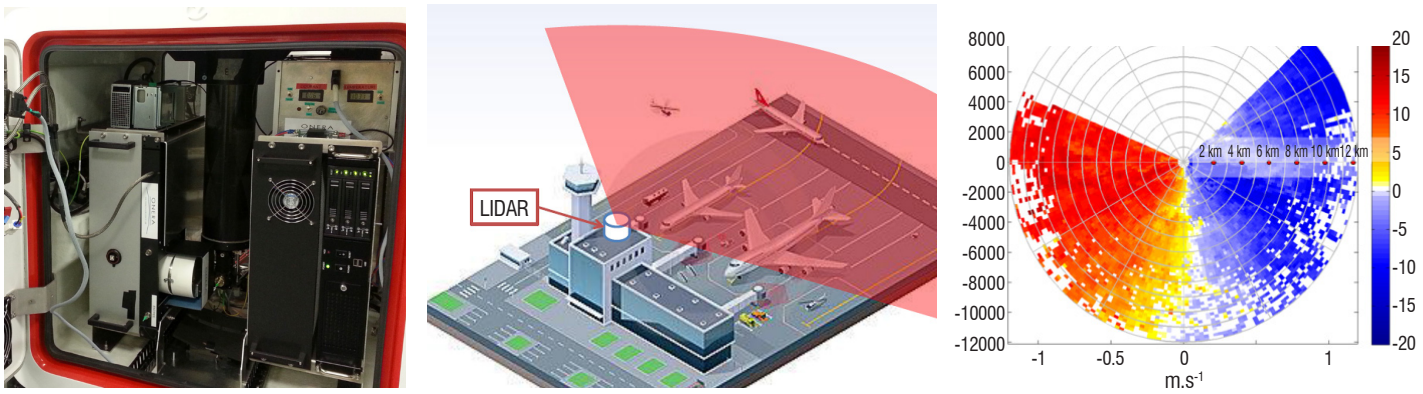


Figure 14 – left: ONERA high-power laser installed in the modified Windcube 200S®; middle: PPI scan for airport measurements; right: raw-wind data (radial wind speed) collected at Toulouse Blagnac airport. Data range up to 10 to 12 km

Wind mapping

Coherent LIDAR measures the radial velocity of the wind, i.e., its projection along the instantaneous line-of-sight of the LIDAR. For wind monitoring, it is desirable to know the local magnitude and direction of the wind. The Velocity Azimuth Display (VAD) technique is a method that has been used for decades for vertical wind profiling with Doppler radars [34].

The principle is described in Figure 15, where the remote wind sensor performs a vertical conical scan of the atmosphere, assuming a locally homogenous wind. Due to the angular dependence of the wind projection along the measurement axis, the Doppler velocity shift of a homogenous wind follows a sine curve.

The principle of VAD can also be extended for the wind field reconstruction of horizontal scans obtained by a long range wind LIDAR. Under the same assumption of a locally homogenous wind, the regression of a truncated sinusoid can be processed from radial wind data, corresponding to a small area of the LIDAR scan. Components of the wind vector are obtained through an ordinary least-square algorithm, as illustrated in Figure 16.

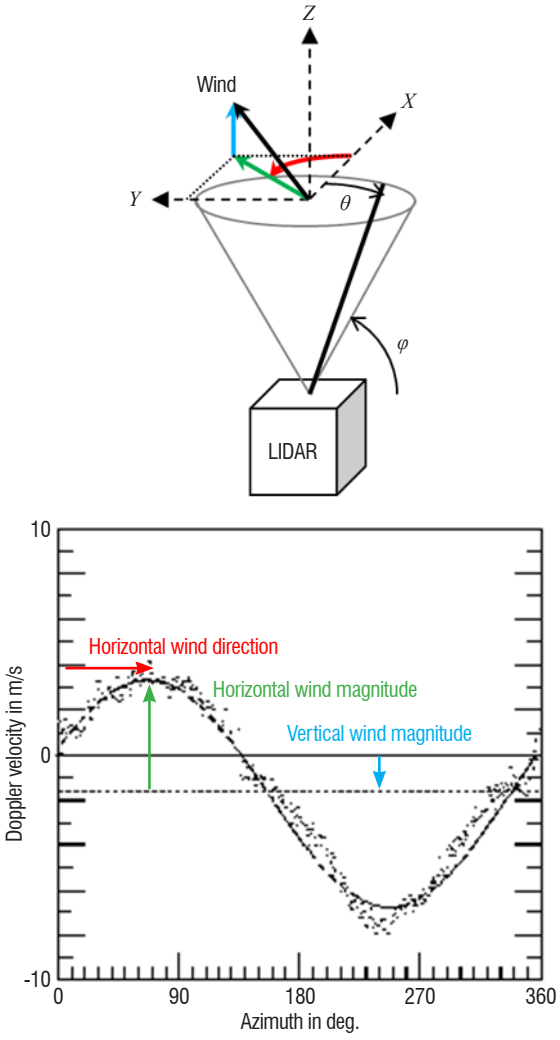
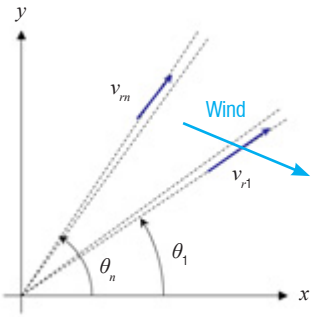


Figure 15 – Velocity Azimuth Display (VAD) principle: The horizontal wind direction is given by the maximum of the sinusoid (red), the horizontal wind magnitude is proportional to the amplitude of the sinusoid (green), and the vertical wind magnitude is proportional to the sinusoid offset (blue). The local vertical component is independent of the scanning angle.



$$\begin{bmatrix} v_{r1} \\ \vdots \\ v_{rn} \end{bmatrix} = M \begin{bmatrix} v_x \\ v_y \end{bmatrix} \quad M = \begin{bmatrix} \cos\theta_1 & \sin\theta_1 \\ \vdots & \vdots \\ \cos\theta_n & \sin\theta_n \end{bmatrix} \cos\phi \quad \begin{bmatrix} \hat{v}_x \\ \hat{v}_y \end{bmatrix} = \left(M^t M\right)^{-1} M^t \begin{bmatrix} v_{r1} \\ \vdots \\ v_{rn} \end{bmatrix}$$

Figure 16 – VAD principle applied to wind mapping. φ is the elevation angle of the LIDAR axis and θi are the azimuth angles

By applying the above to various spatial bins of the LIDAR scan, one can obtain a wind vector map in the x-y plane from radial velocity data, such as that shown in Figure 17. The same principle can be applied with various elevation angles, in order to obtain the 3D wind vector.

In order to evaluate the quality of the wind vector reconstruction, dual LIDAR field tests have been performed. Long-range LICORNE LIDAR has been measured in horizontal scans, while a commercial Leosphere Windcube® profiler measured the wind profile on a vertical directly above the LIDAR, providing data for comparison. Figure 18 shows the experimental set up. The range between the two LIDARs

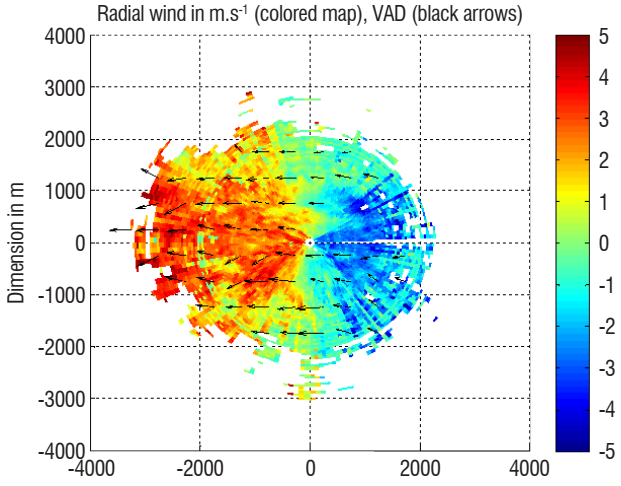


Figure 17 – 2D VAD applied to LICORNE LIDAR data. Black arrows indicate the reconstructed wind vector direction and amplitude

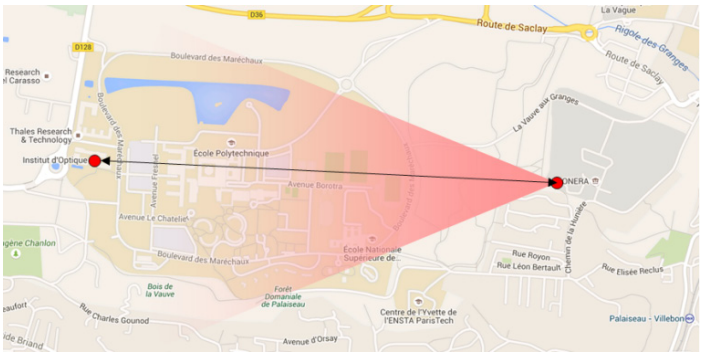


Figure 18 – Experimental set up: location of the profiler (red spot on the left), location of LICORNE (red spot on the right. The pink area is the horizontal scan pattern

was 2100 meters. The LICORNE horizontal scan pattern covered +/-15°. Given the necessary elevation angle of 15° (in order to avoid trees around ONERA), the horizontal scan pattern was 540 meters above the wind profiler location.

The transverse component of the wind located above the Windcube has been reconstructed and both vector data have been compared in magnitude and direction. When the wind is laminar enough, good agreement between both LIDARs is found, as illustrated by Figure 19.

Under turbulent conditions, the wind vector reconstruction is more difficult. However, an interesting feature of the LIDAR is that it can also remotely obtain the state of turbulence of the atmosphere as described in Paragraph "Obtaining the Eddy dissipation rate (EDR)". Under operational conditions, when the scanning LIDAR is used as a stand-alone system, obtaining the EDR is a technique that can thus assess the quality of the wind vector reconstruction.

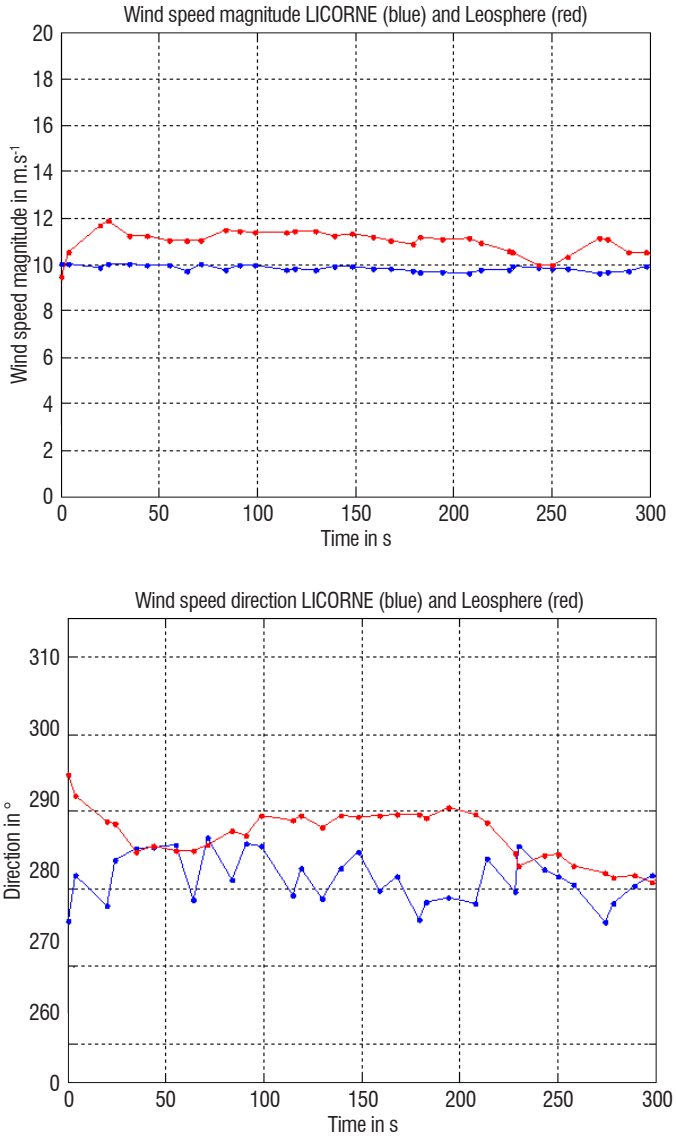


Figure 19 – Example of obtaining the transverse wind component. Top: wind speed magnitude; down: wind speed direction. Blue line: reconstructed from scanning LIDAR LICORNE; red line: measured by the Windcube® profiler

Obtaining the eddy dissipation rate (EDR)

Wake vortex decay will be quicker with high atmospheric turbulence. Theoretical analyses show that the decay time of wake vortices is proportional to the cubic root of the eddy dissipation rate (EDR).

Coherent Doppler LIDAR are commonly used to measure wind fields. They can also provide information about wind field spatial statistics and then yield an estimation of the turbulence Eddy Dissipation Rate. EDR retrieval from LIDAR data remains a relative new topic, especially for addressing operational purposes, such as air traffic applications.

Several authors have performed a review of atmospheric turbulence estimates from Pulsed LIDAR wind measurements [42][47]. Doppler LIDARs can provide information about wind field spatial statistics and then yield an estimation of the turbulence or Eddy Dissipation Rate [4][23][25]. The estimation can be made from the:

- Doppler Spectrum width,
- Velocity Variance or
- Velocity Structure function.

In the spatial domain, the EDR ε appears, for example, in the second order radial velocity structure function $D_v(s)$ defined by:

$$D_v(\vec{s}) = \left(V(\vec{r}) - V(\vec{r}+\vec{s})\right)^2 \quad (2)$$

where $V(\vec{r})$ is the radial velocity measured at range $\vec{r}$ from the LIDAR, $\vec{s}$ is a distance and the average is done on a spatial area. The structure function $D_V(\vec{s})$ can be estimated either along the LIDAR axis (the wind longitudinal structure function $D_{LL}(s)$) or along a transverse direction (the wind azimuthal structure function $D_{NN}(s)$). For a scanning LIDAR, the azimuthal structure function $D_{NN}(s)$ method is preferred according to Chan's analysis [15]. This procedure enables a much better resolution than the effective LIDAR range resolution.

Furthermore, there is a direct relationship with the velocity energy spectrum E, leading to the following expression:

$$D_{NN}(s) = C_{NN} \cdot \varepsilon^{2/3} \cdot s^{2/3} \quad (3)$$

where $C_{NN} = \frac{27\Gamma\left(\frac{1}{3}\right)}{55}\frac{4}{3}C \approx 2.67$ and C is the universal Kolmogorov constant [40][41]. The EDR value expressed as the cubic root of ε in $m^{2/3}.s^{-1}$ is then obtained by fitting the 2/3 slope of Equation (3) for the structure function D_{NN} obtained from wind data. The estimation of D_{NN} from LIDAR data requires further processing as proposed by Frehlich [25], in order to take into account LIDAR filtering. LIDAR filtering includes the effects of pulse length and signal analysis length.

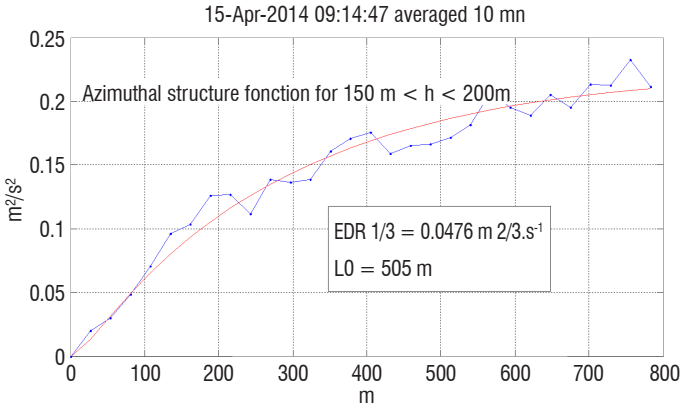


Fig. 20 – EDR measurement: example of a structure function fit for measurement points between 150 and 200 m averaged over 10 min

Figure 20 shows an example of the velocity structure function fit (red) for measurement points (blue) every 50 m in altitude and averaged over 10 min. This EDR value corresponds to quiet air.

Figure 21 shows EDR estimates for 2014, April 24th. The EDR values range from 0.0002 to 0.04 m^{-2}/s^3, yielding $EDR^{1/3}$ values between 0.06 and 0.35 $m^{-2/3}/s$, such values corresponding to moderate to medium turbulence levels.

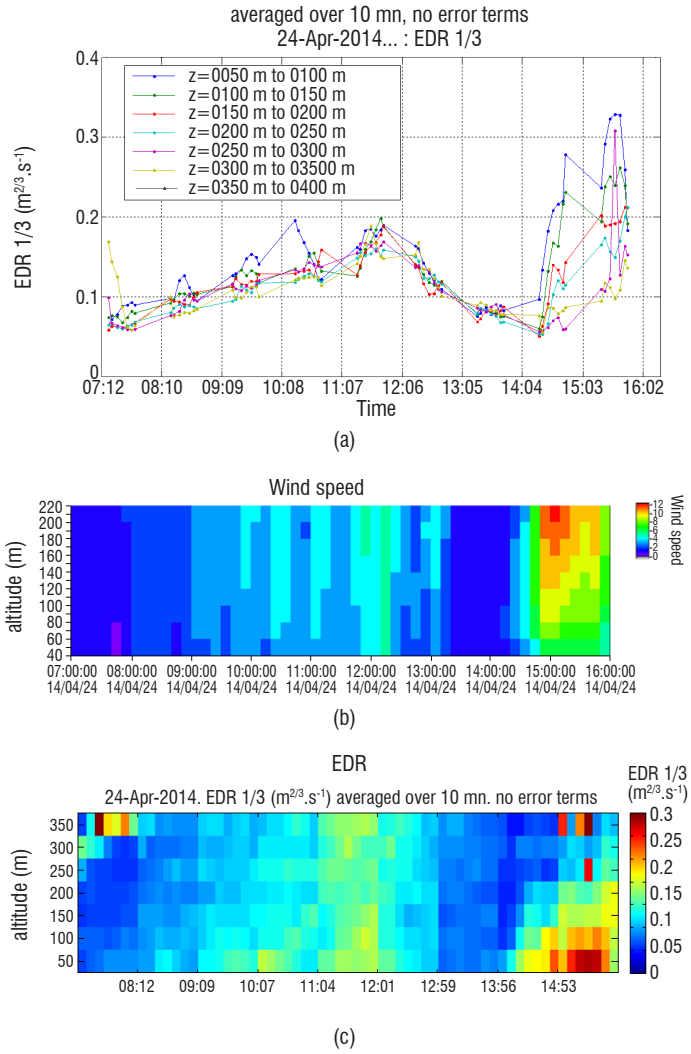


Figure 21 – April 24th, 2014 measurements.(a) EDR as a function of time for different altitudes; evolution of wind speed (b) and EDR (c) in a color scale displayed versus time (horizontal axis) and altitude (vertical axis)

Measurements show a slow increase in the EDR values from 8 to 12 am, then return to calm air at 1:50 pm before the event of a more turbulent period between 2:00 to 4:00 pm. The increase in the EDR in the morning is mainly due to the growth of the convective planetary boundary layer forced by solar radiation, whereas the increase in the afternoon is due to the increase in the horizontal wind speeds. The displays show that high wind speed in the upper layer creates a turbulent layer close to the ground.

These tests demonstrate that wind turbulence characterization is possible over large areas using a range resolved wind LIDAR.

Airborne coherent LIDAR demonstrator

Airborne remote airspeed sensing with coherent LIDAR has a number of applications, such as navigation & pilot aid (e.g., for helicopters), air data system calibration or flight test instrumentation. In addition to size and consumption constraints, reliability in order to reduce maintenance cost is one of the main design drivers for airborne systems. The maturity level for operational systems based on direct detection LIDARs (Paragraph 4) is around TRL2-3, whereas coherent detection LIDAR maturity is around TRL4-5 or higher. For airborne sensors, good laser efficiency is an important feature and is an asset of the fiber technology. Plug efficiencies around 10% are obtained with a continuous fiber laser emitting 5 W at 1.5 μm. Multi-watt fiber airspeed LIDAR demonstrators are being designed and tested throughout the world [1][3][29][51].

Vibration

One of the main assets of fiber technology is its resilience in a shaking environment, which makes the fiber LIDAR an excellent candidate for operation in flight. Demonstrators are tested on the ground in a vibrating environment at ONERA before flight tests. The LIDAR performances are tested on reference Doppler targets and on wind.

In this example, the LIDAR is a new high-power demonstrator built at ONERA/DOTA in 2016, which will be flight tested this year. The LIDAR sensor head is installed on an electromagnetic shaker and aimed at the Doppler target (rotating disk, see Figure 22). The excitation corresponds to the main vibration lines recorded during a strong shake of an aircraft.

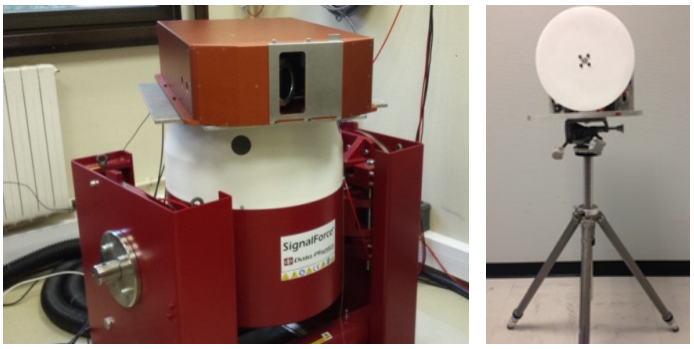


Figure 22 – left: LIDAR sensor head installed on the shaker; right: rotating disk used as the reference Doppler target

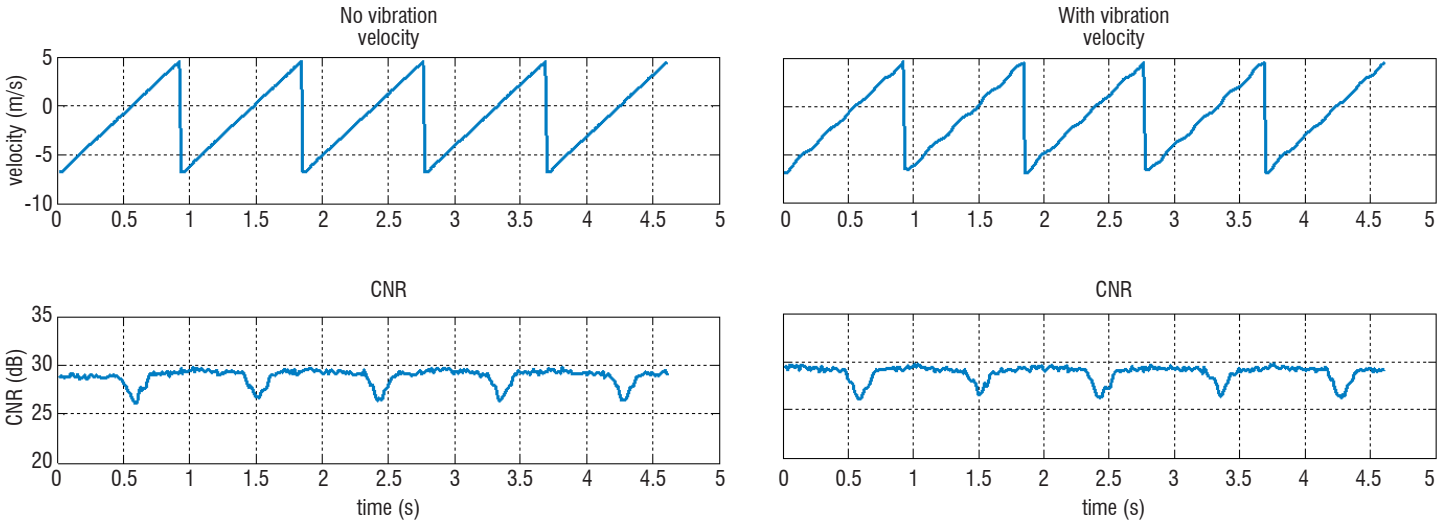


Figure 23 – left: without vibration; right: with vibrations; top: target radial velocity as a function of time (the modulated signal in the figure on the right is due to the LIDAR's own velocity component); bottom: CNR as a function of time (dips occur when the laser crosses the center of the target with a lower albedo)

A scan of the laser beam along a diameter of the rotating target (0.6° in 1 s) enables the reflectance inhomogeneity to be smoothed. Figure 23 displays the LIDAR output signal: target radial velocity (top) and CNR profiles (bottom).

The average CNR value, 29 dB remains unchanged with and without vibrations. This indicates that the vibrations do not degrade the mixing efficiency and that the system design is resilient to vibrations even under turbulent conditions.

True air speed

Measuring the true air speed (TAS) non-intrusively in the free stream region in front of an aircraft is useful to complement the air data system. A short-range LIDAR providing local air speed at a few tens of meters from the aircraft allows a precise and remote measurement of the air speed outside the range of the flow disturbance from the aircraft: it is able to give the velocity in real time with no in-flight calibration, using autonomous onboard equipment and without *a priori* assumptions regarding the atmosphere. The LIDAR is aimed through a dedicated window, which avoids protruding equipment on the aircraft skin.

For rotary wing aircraft or UAV, even close-to-zero air speed values are measured, securing the vehicle also in hover flight. Such anemometer probes for true air speed measurements with high velocity precision have already been flight tested. In 2003, DALHEC, a 3-axis wind LIDAR, was built and flown onboard a Dauphin helicopter [11]. It was dedicated to True Air Speed measurements at medium range (30-100 m), away from rotor flow perturbations. The purpose was to evaluate the technology with regard to air speed accuracy, essentially at low speed.

Doppler LIDAR is also useful during certification procedures. Indeed, the calibration of an aircraft air data sensor requires cumbersome procedures, including specific equipment and costly dedicated flight tests. Calibrations of the pitot static system and vanes using a laser anemometer have increased accuracy compared with those obtained with conventional techniques, such as a towed cone, tower-fly-by, or a pacer aircraft.

An airborne proof-of-concept measuring the air speed radial component (single axis LIDAR) ahead of an aircraft has been fly tested even at high altitude using a high-power fiber laser [51]. The LIDAR developed within the framework of the AIM2 European project, aimed at developing a 4-axis fiber anemometer to measure the full air speed vector. The purpose of the demonstrator was to assess the in-flight airspeed calibration of True Air Speed (TAS), Angle-Of-Sideslip (AOS) and Angle-Of-Attack (AOA). Its performances have been tested during aircraft maneuvers, a functionality that is not possible with the towed cone.

It should meet the following requirements:

- TAS dynamic: 50 m/s ≤ Vx ≤ 200 m/s
- TAS accuracy: ≤1 m/s
- Angle dynamics: -20° ≤ AOS ≤ +20° ; -2° ≤ AOA ≤ +15°
- Angle accuracy: 0.5°
- Measurement frequency: 16 Hz.

The LIDAR measures the air speed along the laser line-of-sight and the time-multiplexed 4 axes are used for speed reconstruction, assuming a homogeneous flow at the short focus distance of 50 m (see Figure 24).

The system is fully fibered ensuring good reliability against vibrations and the latest commercial fiber components.

A power spectral density is computed every 2.6 μs and, in order to improve the signal-to-noise ratio, each spectrum is averaged using 13 ms of signal acquired along each LIDAR axis. The system computes the 4 radial speeds every 100 ms and converts them into a 3D air speed vector in the aircraft reference frame.

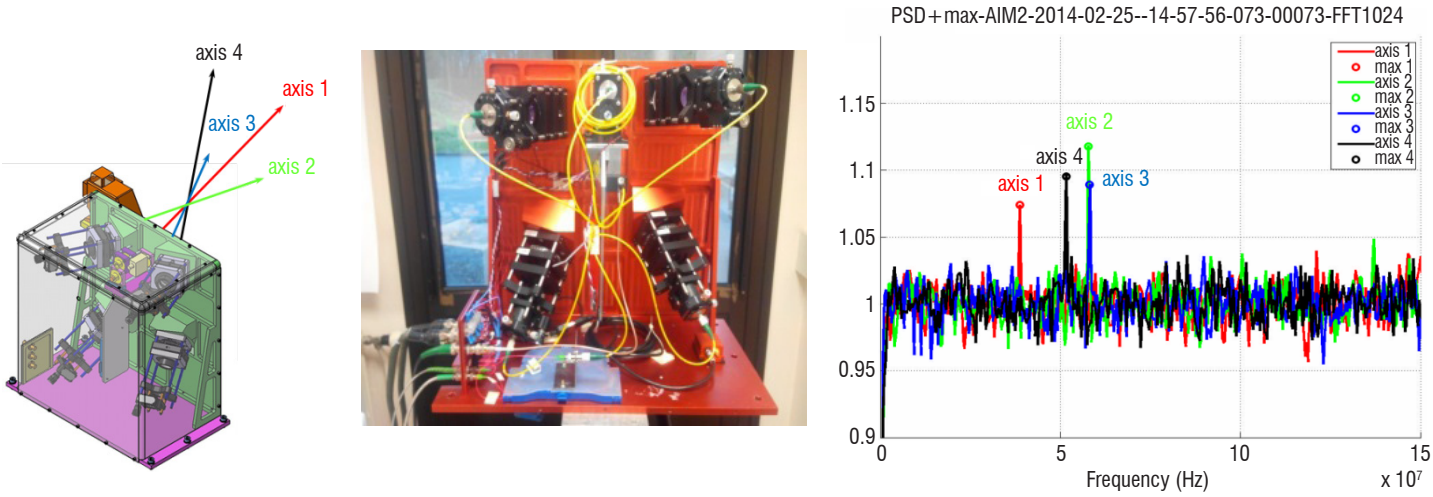


Figure 24 – right & middle: Sensor head with 4 Doppler measurement axes; left: Power Spectral Density for each axis and frequency maximum detection

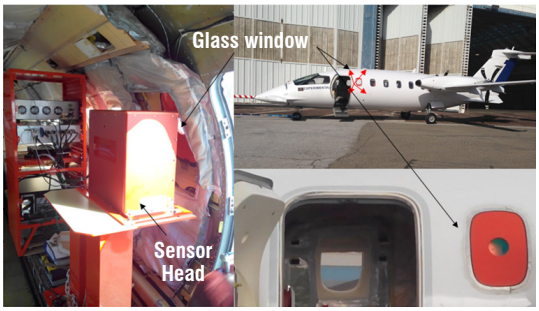


Figure 25 – AIM2 installed onboard a P180

The LIDAR was installed onboard the Piaggio 180 and aimed through a passenger window fitted with an optical grade window (Figure 25). Having no protruding part facilitates flight procedures.

The system was flight tested at the Genova airfield in 2014 [3]. Reconstruction and graphical comparisons of True Air Speed, Angle-Of-Sideslip and Angle-Of-Attack with the aircraft Flight Test Instrumentation (FTI) have been performed in the LIDAR sensor head reference frame and are shown on Figure 26. Data have been analyzed during various maneuvers: Multiple ground tracks, rollercoaster pull up, level acceleration, and flat turn.

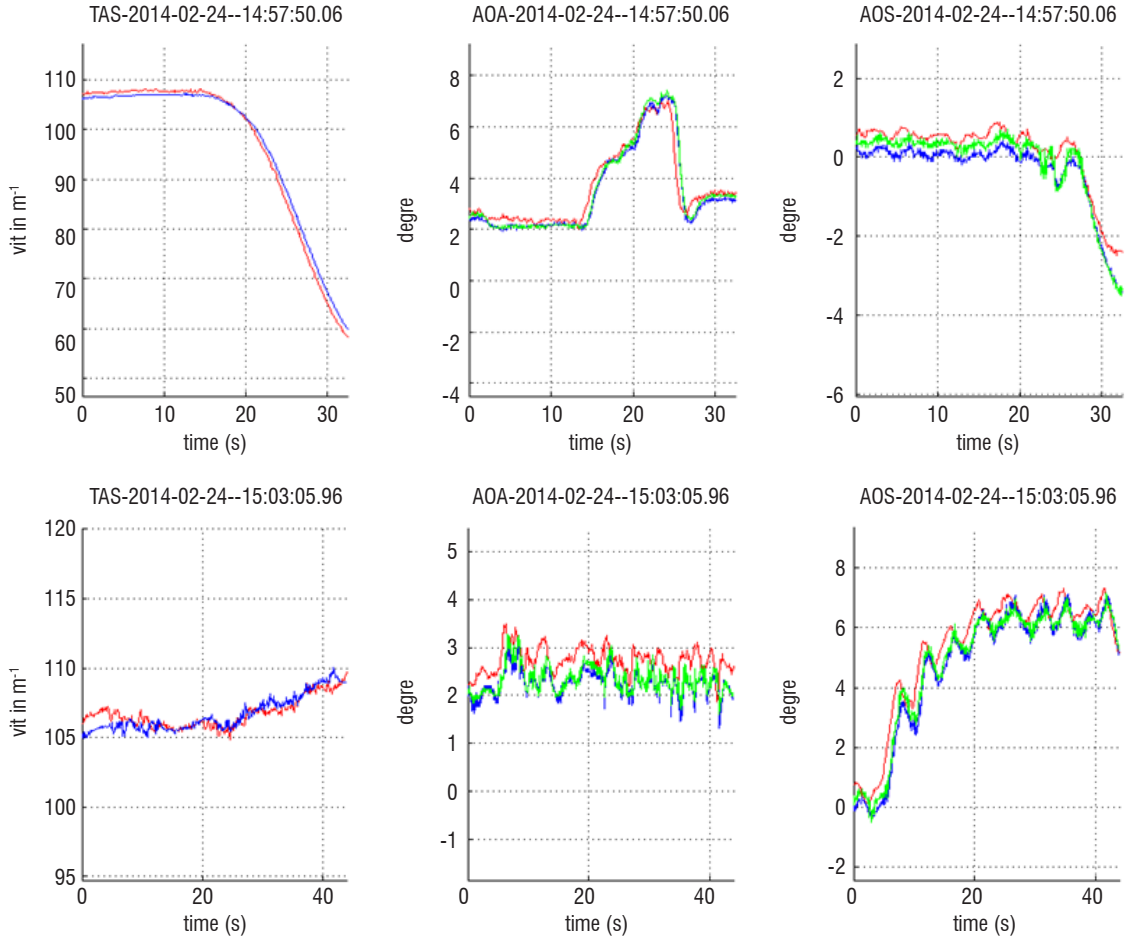


Figure 26 – LIDAR (red) and FTI data (green & blue) flight data – top: rollercoaster pull up – bottom: flat turn; left: TAS; middle: AOA; right: AOS

All bias between the LIDAR and the FTI data sets was within specifications demonstrating a very good agreement showing that fiber LIDAR technology has matured enough to provide a powerful technique for onboard air data calibration. The technology offers a significant capability in size reduction and could also be envisaged for UAV flight tests in a near future.

Conclusion

ONERA range resolved wind LIDARs based on new high-power, high-spectral-brightness, all-fibered sources have been used in various LIDAR configurations to demonstrate performance and evaluate technology maturity. We have investigated eyesafe high-peak-power sources using various SBS mitigation techniques to overcome nonlinear optics limitations. Using these sources, record ranges have been obtained in field tests. Improved algorithms for real-time wake vortex characterization, new EDR algorithms have been applied to wind measurement and validated with simulations.

High-power fiber lasers with optimized coherent LIDAR architectures offer new perspectives for applications related to airport safety and aircraft navigation, not to mention other fields of applications such as defense, security or space.

Operating an airborne LIDAR at cruise altitude would be of interest for clear air turbulence detection. However, the technology needs to mature in order to meet size and costs constrains. Nonetheless, at lower to medium altitude, coherent Doppler LIDARs based on fiber technology offer a rich panel of onboard functionalities. They can complement pitot vanes to increase air data system reliability and flight safety or to optimize the certification cycle, saving time and cost. Probing airspeed at longer ranges than the present air demonstrators using more powerful fiber lasers would enable new capabilities to be implemented, such as gust detection in a broad range of altitudes, providing anticipated warning against air dynamics hazards. They would also help in the design and certification of new air vehicles (fixed wing, rotary wing, UAV, airship, etc.) and ease their integration into air traffic. Their assets include increased reliability and resilience to vibration and low power consumption. Moreover, new designs taking advantage of the fiber architecture intrinsic multifunction capability could provide enhanced survivability of air vehicles in degraded atmospheric environments ■

Acknowledgments

The authors would like to thank the European Commission for funding the CREDOS and FIDELIO projects within the framework of the FP6 and the UFO, DELICAT and AIM2 projects within the framework of the FP7, and ESA for funding the HEPILAS project.

The authors would like to thank Leosphere for a fruitful and long-lasting collaboration.

Some of this work was also supported by the French *Direction Générale de l'Armement*.

References

[1] M. AKBULUT, *et al. - Pulsed Coherent Fiber LIDAR Transceiver for Aircraft In-Flight Turbulence and Wake-Vortex Hazard Detection.* Proc. of SPIE Vol. 8037, 2011, doi: 10.1117/12.883990.

[2] A. N. ABRAMOV, *et al. - Fabrication of Heavily Er2O3 Doped Aluminophosphosilicate Glass Fibers*. Inorganic Materials, 46(4), 2010, p. 439-444.

[3] B. AUGERE, *et al. - 1.5 µm LIDAR Anemometer for True Air Speed, Angle of Sideslip and Angle of Attack Measurements Onboard Piaggio P180 Aircraft*. Measurement Science and Technology Journal, 2015, MST-102092.R1.

[4] VA BANAKH, *et al. - Estimation of Turbulent Energy Dissipation Rate from Data of Pulse Doppler LIDAR*. Journal of Atmospheric and Oceanic Optics 10, 1997, 957-965.

[5] C. BESSON, *et al. - New Fiber Laser for LIDAR Developments in Disaster Management.* Proc. of SPIE, Vol. 9250, 2014, doi: 10.1117/12.2073241.

[6] D. Bruneau - *Mach-Zehnder Interferometer as a Spectral Analyser for Molecular Doppler Wind LIDAR.* Appl. Opt. Vol.40, n°3, 2001, p. 391.

[7] G. Canat, *et al. - Performance and Limitations of High Brightness Er3+/Yb3+ Fiber Sources*. C. R. Physique 7 (2), 2006 p. 177-187.

[8] G. Canat, *et al. - Multifilament-Core Fibers for High Energy Pulse Amplification at 1.5 µm with Excellent Beam Quality*. Opt. Lett. 33, (22), 2008, p. 2701-2703.

[9] G. CANAT, *et al. - High Power Pulsed Fiber Laser Development for CO_2 Space Based Dial System.* International Conference on Space Optics (ICSO 2014), Tenerife, 07-10/10/2014.

[10] G. CANAT, *et al. - High Peak Power Single Frequency Amplifiers Based on Efficient Erbium-Ytterbium Doped LMA Fibers.* European Conference on Lasers and Electro-Optics-European Quantum Electronics Conference (CLEO/Europe-EQEC 2015), Munich (Allemagne), 21-26/06/2015.

[11] J.-P. CARIOU, *et al. - All-Fiber 1.5 µm CW Coherent Laser Anemometer DALHEC. Helicopter Flight Test Analysis*. 13th Coherent Laser Radar conference, Kamakura, 2005.

[12] J.-P. CARIOU, *et al. - Laser Source Requirements for Coherent LIDARs Based on Fiber Technology*. Comptes Rendus Physique, Volume 7, Issue 2, 2006, p. 213-223.

[13] N. CÉZARD, *et al. - Airflow Characterization by Rayleigh-Mie LIDARs.* Journal of Aerospace Lab AL01-06, Issue 1, December 2009.

[14] N. CÉZARD - *Performance Evaluation of a Dual Fringe-Imaging Michelson Interferometer for air Parameter Measurements with a 355nm Rayleigh–Mie LIDAR.* Appl. Opt., Vol. 48, No. 12, 2009.

[15] P. W. CHAN - *LIDAR-Based Turbulence Intensity Calculation Using Glide-Path Scans of the Doppler LIght Detection And Ranging (LIDAR) Systems at the Hong Kong International Airport and Comparison with Flight Data and a Turbulence Alerting System*. Meteorologische Zeitschrift 19 (6), 2010, p. 549-563.

[16] M.-L. CHANIN, *et al. - A Doppler LIDAR for Measuring Winds in the Middle Atmosphere*. Geoph. Res. Let., Vol.16, No. 11, p. 1273, 1989.

[17] J. M. CHAVEZ BOGGIO, *et al.* - *Experimental and Numerical Investigation of the SBS-Threshold Increase in an Optical Fiber by Applying Strain Distributions.* J. Lightwave Technol., 23, 2005, 3808.

[18] E. Dakin, *et al.* - *LandSafe® Precision Flight Instrumentation System, the DVE Solution.* American Helicopter Society 68th Annual Forum, Fort Worth, Texas, (2012) Copyright © 2012 by the American Helicopter Society International, Inc.

[19] A. DOLFI-BOUTEYRE, *et al.* - *Pulsed 1.5- µm LIDAR for Axial Aircraft Wake Vortex Detection Based on High-Brightness Large-Core Fiber Amplifier*. IEEE Journal of Selected Topics in Quantum Electronics, vol.15, No. 2, p. 441,450, March-april 2009.

[20] A. DOLFI-BOUTEYRE, *et al.* - *Aircraft Wake Vortex Study and Characterization with 1.5 µm Fiber Doppler LIDAR.* Journal of Aerospace Lab AL01-07 Issue 1, December 2009.

[21] M. EMANUEL, *et al.* - *In Situ Performance Standard for Eddy Dissipation Rate.* 2013.

[22] C. FLESIA, *et al.* - *Theory of the Double Edge Molecular Technique for Doppler Wind LIDAR Measurement*. Appl. Opt. Vol. 38, No. 3, p. 432, 1999.

[23] R. FREHLICH, *et al.* - *Coherent Doppler LIDAR Measurements of Wind Field Statistics*. Boundary-Layer Meteorology 86, p. 233-256, 1998.

[24] R. FREHLICH, *et al.* - *Maximum Likelihood Estimates of Vortex Parameters from Simulated Coherent Doppler LIDAR Data*. Journal of Atmospheric and Oceanic Technology, Vol. 22, No. 2, p. 117-130, 2005.

[25] R. FREHLICH, *et al.* - *Measurements of Boundary Layer Profiles in an Urban Environment*. Journal of Applied Meteorology 45, p. 821-837, 2006.

[26] A. HALLERMEYER, *et al.* - *Development and Assessment of a Wake Vortex Characterization Algorithm Based on a Hybrid LIDAR Signal Processing.* AIAA, (accepted), 2016.

[27] A. HAUCHECORNE, *et al.* - *Set-Up of a Ground-Based Rayleigh LIDAR to Detect Clear Air Turbulence.* Proceedings ILRC 25, 2010.

[28] A. HAUCHECORNE, *et al.* - *Tentative Detection of Cear Air Turbulence Using a Ground-Based Rayleigh LIDAR*. Appl. Opt. (accepted), 2016.

[29] H. INOKUCHI - *Development of an Onboard Doppler LIDAR for Flight Safety.* Journal of Aircraft Vol. 46, No. 4, July-August 2009, DOI: 10.2514/1.41738.

[30] Y. JAOUEN, *et al.* - *Power Limitation Induced by Nonlinear Effects in Pulsed High-Power Fiber Amplifiers*. C. R. Physique 7, p. 163-169, 2006.

[31] S. KAMEYAMA, *et al.* - *Compact All-Fiber Pulsed Coherent Doppler LIDAR System for Wind Sensing*. Appl. Opt. 46, 1953-1962, 2007.

[32] A. KOBYAKOV, *et al.* - *Stimulated Brillouin Scattering in Optical Fibers*. Advances in optics and photonics, 2(1), p. 1-59, 2010.

[33] G. KULCSAR, *et al.* - *Multiple-Stokes Stimulated Brillouin Scattering Generation in Pulsed High-Power Double-Cladding Er3+-Yb3+ Codoped Fiber Amplifier*. Photonics Technology Letters, IEEE 15 (6), p. 801- 803, 2003.

[34] R. LHERMITTE, *et al.* - *Precipitation Motion by Pulse Doppler*. Preprints 9th Weather Radar Conf., Kansas City, Amer. Meteor. Soc., p. 218-223, 1961.

[35] L. LOMBARD, *et al.* - *Coherent Beam Combination of Narrow-Linewidth 1.5 µm Fiber Amplifiers in a Long-Pulse Regime*. Optics letters, 36(4), p. 523-525, 2011.

[36] L. LOMBARD, *et al.* - *Flight Testing Delicat – A Promise for Medium-Range Clear Air Turbulence Protection.* Conference on Lasers and Electro-Optics (CLEO 2014), San José (USA), 08-13/06/2014.

[37] L. LOMBARD, *et al.* - *Eyesafe Coherent Detection Wind LIDAR Based on a Beam-Combined Pulsed Laser Source*. Opt. Lett., vol. 40, No. 6, p. 1030-1033, 2015.

[38] L. LOMBARD, *et al.* - *Long Range Wind LIDARs Based on Novel High Spectral Brilliance All-Fibered Sources*. SPIE Remote Sensing 2015, Toulouse (France), 21-24/09/2015.

[39] P. MATTHEW, *et al.* - *Atmospheric Turbulence Estimates from a Pulsed LIDAR*. 51st AIAA Aerospace Sciences Meeting, 2013.

[40] A. S. MONIN & A. M. YAGLOM - *Statistical Fluid Mechanics : Mechanics of Turbulence – Vol 2.* The MIT Press, 1975.

[41] S. POPE - *Turbulent Flow.* Cambridge University Press, 2000.

[42] M. PRUIS, *et al.* - *Atmospheric Turbulence Estimates from a pulsed LIDAR*. 51st AIAA Aerospace Sciences Meeting, 2013.

[43] S. RAHM, *et al.* - *Characterization of Aircraft Wake Vortices by Airborne Coherent Doppler LIDAR.* Journal of Aircraft Vol. 44, No. 3, May-June 2007.

[44] S. RAHM, *et al.* - *Aircraft Wake Vortex Measurement with Airborne Coherent Doppler LIDAR*. Journal of Aircraft Vol. 45, No. 4, July-August 2008.

[45] O. REITEBUCH, *et al.* - *The Airborne Demonstrator for the Direct-Detection Doppler Wind LIDAR ALADIN on ADM-Aeolus*. Part I: Instrument Design and Comparison to Satellite Instrument DOI: 10.1175/2009JTECHA1309.1, J. Atm. Oc. Tech., Vol. 26, 2009.

[46] W. RENARD, *et al.* - *Beyond 10 km Range Wind-Speed Measurement with a 1.5 µm All-Fiber Laser Source*. CLEO: Applications and Technology (p. AW1P-5), Optical Society of America, June 2014.

[47] A. SATHE, *et al.* - *A Review of Turbulence Measurements Using Ground-Based Wind LIDARs*. Atmos. Meas. Tech., 6, 3147-3167, doi:10.5194/amt-6, 3147-2013, 2013.

[48] N. P. SCHMITT, *et al.* - *The AWIATOR Airborne LIDAR Turbulence Sensor.* Elsevier Aerospace Science and Technology 11, p. 546-552, 2007.

[49] T. M. SHAY, *et al.* - *First Experimental Demonstration of Self-Synchronous Phase Locking of an Optical Array.* Opt. Express 14, 12015, 2006.

[50] I. SMALIKHO, *et al.* - *Method of Radial Velocities for the Estimation of Aircraft Wake Vortex Parameters from Data Measured by Coherent Doppler LIDAR*. Optics Express, Vol. 23, No. 19, p. A1194-A1207, 2015.

[51] S. M. SPULER, *et al.* - *Optical Fiber-Based Laser Remote Sensor for Airborne Measurement of Wind Velocity and Turbulence*. Appl. Opt., Vol. 50, No. 6 / 20, February 2011.

[52] L. P. THOBOIS, *et al.* - *Wind and EDR Measurements with Scanning Doppler LIDARs for Preparing Future Weather Dependent Separation Concepts.* 7th AIAA Atmospheric and Space Environments Conference (AIAA AVIATION 2015), Dallas (USA), 22-26/06/2015.

[53] G. G. VIENNE, *et al.* - *Role of Aluminum in Ytterbium–Erbium Codoped Phosphoaluminosilicate Optical Fibers*. Optical Fiber Technology, 2(4), p. 387-393, 1996.

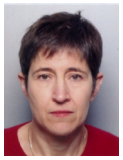
[54] X. ZHANG, *et al.* - *Single-Frequency Polarized Eye-Safe All-Fiber Laser with Peak Power Over Kilowatt*. Applied Physics B, p. 1-5, 2013.

[55] *The CREDOS Project D2-6 - WP2 Final Report, CREDOS European project.* August 2009. https://www.eurocontrol.int/sites/default/files/content/documents/sesar/credos-d2-6-wp2-final-report-v11.pdf.

[56] G. JENARO RABADAN, *et al.* - *Forward Looking Airborne LIDAR for Automatic Feed Forward Control of Turbulent In-Flight Phaenomena.* AIAA Journal of Aircraft, Vol. 47, No. 2, p. 392-403, March-April 2010.

AUTHORS

Claudine Besson received her PhD degree in physics from the Optics Graduate School, Paris, France in 1989. She is currently a senior scientist and research group leader at ONERA, the French Aerospace Lab. Her current research interests are fiber lasers and LIght Detection And Ranging systems (LIDAR).



Agnès Dolfi-Bouteyre graduated from the *Ecole Superieure d'Optique*, Orsay, (1986) and received a PhD degree in Physics from the University of Paris XI (1990). She joined ONERA in 1990, where she has been involved in laser and lidar system development for defense and aerospace. Her current areas of interest and specialization are: lidar, coherent lidar, atmospheric turbulence measurement, aircraft wake-vortex detection and lidar signal processing .

Guillaume Canat graduated from the *Ecole Polytechnique* (2000), Telecom Paristech (2002) and holds a PhD in physics (2006). He has been working at ONERA, the French aerospace lab, from 2005-2016. His research topics include fiber lasers, non-linear effects, and lidars. He is the author or coauthor of more than 25 papers and 80 international communications.

Nicolas Cézard was born in 1980. He obtained a MSc in Photonics and microwaves in 2004 and a PhD in Physics at *Ecole Polytechnique* in 2008. He works as a senior scientist in the Laser and Lidar group of the Optics Department at ONERA, the French Aerospace Lab. He is in charge of the development of a new lidar system for the remote measurement of winds and gas species.

Beatrice Augere is a senior research scientist at the *Office National d'Étude et Recherche Aérospatiale* (ONERA). She has a PhD in non-linear optics (1988). Since obtaining her degree, her research and development activities have been related to laser remote sensing (Lidar) for various applications in defense, security, aeronautics and space. She authored or co-authored more than 20 conference or journal papers. She has managed projects involving large budgets and various partners from industry or academic institutes. She has a wide experience in field testing and supervised Lidar campaigns, in numerous European projects, such as IWAKE, AWIATOR, CREDOS, FIDELIO, SESAR, AIM, and AIM2. She has been involved in measurements by Lidar for more than 20 years, with a special emphasis on coherent Lidar for various applications, such as true air speed measurements, vibration measurements, and vortex detection.

Anne Durécu is a researcher working in the DOTA department at ONERA. She graduated from the *Institut National des Télécommunications* in 2001 and received a PhD in Optics in 2005 from the *Université de Limoges*. After 3 years in the Research Department of Alcatel, where her research activities focused on non-linear effects in optical fibers, she joined ONERA. Since 2005, she has been working on fiber laser development for lidar applications and also on the applications of lasers on the battle field (laser weapons, laser-dazzling of imaging systems).

Laurent Lombard graduated from the *Institut d'Optique* in 2001 and obtained a PhD in optics from *Université Paris XI* in 2005 on high power fiber amplifier and wavefront techniques. He is now with ONERA and works on new lidar architectures for aircraft safety and wind measurement, as well as power scaling by beam combination and nonlinear effect mitigation of high-power laser sources.

Matthieu Valla was born in France in 1976, and received the engineering and PhD degree from Telecom Paristech (Paris, France) in 2001 and 2005. He is currently working as a research engineer at ONERA since 2002, in the field of coherent Doppler Lidar for anemometry, vibrometry and wind field retrieval.

Alexandre Hallermeyer received his Engineering degree from ENSEEIHT (*École Nationale Supérieure d'Électronique, Électrotechnique, Informatique, Hydraulique et Télécommunications*) INP-Toulouse in 2013 and is currently doing a PhD thesis on LIDAR signal processing for the characterization of Wake Vortices.

G. Le Besnerais, F. Champagnat, P. Cornic, A. Plyer, B. Leclaire, A. Cheminet, C. Illoul, G. Losfeld, Y. Le Sant, D. Donjat, F. Nicolas, F. Micheli
(ONERA)

E-mail: guy.le_besnerais@onera.fr

DOI: 10.12762/2016.AL12-09

Experimental Fluid Mechanics goes 3D: New Numerical Methods for Quantitative Instantaneous 3D Imagery of Fluids

We review recent developments made between three departments at ONERA (DTIM, DAFE, DMAE) with regard to 3D imagery used in fluid mechanics experiments. We first discuss 3D PIV (Particle Image Velocimetry) and present original contributions on the modeling of the imaging process, the reconstruction of the 3D volume of particles and the estimation of the 3D motion field between two time instants. These three contributions emphasize the pointwise character of particles in PIV, in contrast with classical Tomo-PIV approaches, and have been shown to outperform state-of-the-art methods, yielding more accurate 3D velocity estimations. Then, we consider 3DBOS (Background Oriented Schlieren), which is aimed at estimating the instantaneous 3D density field of a moving fluid. We have recently proposed an original one-step numerical approach for 3DBOS. This method has been successfully used in several experiments conducted in various ONERA facilities, in particular in the F2 and S1MA wind tunnels. Finally, we draw some important perspectives, especially for the study of compressible flows by combining both techniques.

Introduction

Over the last decade, new imaging set-ups and numerical methods aimed at 3D field measurement for fluid mechanics have been developed. In its most common form, tomographic PIV, introduced in [19], uses 4 cameras (or more) to reconstruct volumes of particles and derive the instantaneous 3D velocity field by 3D correlation. 3DBOS [1][32] is an extension of Background Oriented Schlieren, which allows the reconstruction of the instantaneous 3D density volume of flows, again by combining images recorded by several (usually more than ten) synchronized cameras.

Being instantaneous and volumetric, these 3D field measurement techniques yield unprecedented information on the flows under study. Indeed, before they were proposed, field characterization in fluid mechanics was only possible either in an instantaneous sense, but limited to 2D (for instance, in plane PIV, see our companion paper in this volume [27]), or could be obtained on 3D domains as well, but in a time-averaged sense (for instance, by scanning space with a pointwise sensor, e.g., by Hot-Wire Anemometry or Laser Doppler Velocimetry). Flows of industrial interest, in particular in the aerospace domain, are characterized by high Reynolds numbers and often also high Mach numbers. As such, they inherently exhibit a three-dimensional structure due to turbulence, not mentioning the increased complexity of the systems (three-dimensional model geometries, active flow control by mechanical or fluidic actuators, etc.). The characterization of complex unsteady three-dimensional flow structures thus appears to be essential to investigate problems of industrial relevance. Finally, from the theoretical point of view, 3D field measurement techniques offer the potential to take a major step towards the complete understanding of complex flows. As an example, they allow the measurement of the velocity gradient and, in particular, the vorticity, which provide information of major importance in wake vortex characterization and, more generally, regarding turbulence studies.

Nevertheless, 3D field measurement also brings several tedious issues from the experimental and numerical point of view. We focus here on the latter point and present original developments made between three departments at ONERA (DTIM, DAFE and DMAE) aimed at achieving more efficient and accurate data processing methods. In both the 3DPIV and 3DBOS contexts, our contributions build on a reformulation of the problem, a careful examination of its experimental conditions and limitations, and lead to new algorithm proposals.

The paper is organized as follows: we first consider 3DPIV and present several methodological contributions made since 2013. The second part of the paper describes the original 3DBOS reconstruction developed since 2012 and recently validated in several facilities at ONERA. Perspectives of these works and opportunities for joint studies between 3DPIV and 3DBOS are outlined in the last section.

3D PIV

Introduction

Recalling 2D PIV

PIV (Particle Image Velocimetry) is a measurement process where two images of a plane in a flow are recorded at two close time instants and correlated, in order to produce an estimated displacement field. This recording uses pulsed laser illumination and often high-speed cameras. The processing basically consists in seeking corresponding small regions (or interrogation windows, IW) in the two images, by optimizing some intensity-based criterion. Numerous algorithms have been proposed for this operation; see [38] for a review. In 2009, some of the authors of this paper proposed a fast parallel algorithm called FOLKI-SPIV [5], whose GPU implementation reached unprecedented performance: dense displacement fields on 4K images are computed in less than 0.2 s. FOLKI-SPIV runtime is faster than the time required for loading the images from the camera. Hence, when processing PIV image sequences, computation becomes actually transparent compared to storage.

Before turning to 3D PIV, some basic facts and constraints on 2D PIV should be briefly recalled. First, in most PIV settings related to aerodynamic studies, unless microscopic viewing conditions are used, the seeding particles appear as point sources for the cameras (in optics, they are said to be unresolved by the cameras). As a consequence, their actual shape in the image stems from the characteristics of the imaging process. This process is essentially controlled by the experimenter, through the tuning of the parameters of the illumination system and of the camera. Ideally, one aims at recording particles in the form of Gaussian-shaped images with diameters of 2-3 pixels, so as to minimize localization uncertainty and aliasing effects. In the end, the actual spatial resolution of PIV estimated displacement fields depends on two main parameters: the size of the IW and the density of particles. 10 particles per IW is recognized as a good rule-of-thumb to ensure the correct behavior of the correlation process, and, in this case, the spatial resolution is given by the size of the IW projected onto the illuminated plane [38].

Tomo-PIV

The main idea behind Tomo-PIV, which was first proposed in [19], is to extend the principle of PIV, *i.e.*, the cross-correlation of two 2D images of seeded flows, to 3D. Hence, one needs to acquire two "3D images" of the flow, in order to cross-correlate them to produce a 3D displacement field. In Tomo-PIV, each of these "3D images" is actually a volume representation obtained by tomographic reconstruction from several simultaneous 2D images of some illuminated 3D area of the flow. This operation is usually formulated as the iterative estimation of a volume, discretized over a large number of 3D grid cells (or voxels). It is important to recall that this tomographic process was originally designed to provide a representation of the actual volume of particles *suitable for correlation*. Hence, the same prescription as that concerning the 2D imaging process in PIV has been followed: the voxel size Δ is usually aligned with the pixel size ($v/p = 1$) and particles in the reconstructed volume appear ideally as 2-3 voxel-wide Gaussian-shaped blobs. This approach leads to the observation model outlined in the left part of Figure 1, where images result from the integration of a 3D intensity field along rays. The modeling of this integration process and the concatenation of the equations associated with all pixels of all cameras lead to a linear system, $I = WE$. In this equation, image intensities I_k on pixels k are related to volume intensities $E_n = E(n\Delta)$ through a weighting matrix W_{kn}, which is non-negative and sparse.

Given that most Tomo-PIV settings use 4 cameras (but systems with 6 or 8 cameras have also been described), the previous system is severely underdetermined – in a 4-cameras system there can be two orders of magnitude less of recorded data (I_k) than there are voxel values (E_n). Prior information on the reconstructed volume is essential here. All Tomo-PIV methods are based upon the hypothesis that the density of seeding particles is low, so that voxel intensities are most often zero and take on positive values only in the vicinity of the particles – and there are very few of these visible. In order to enforce these properties in the reconstructed volume, multiplicative algorithms, such as MART and SMART, originating from Computerized Tomography, are popular choices and their action is often restricted to "valid" voxels selected in an initial process such as MLOS [2].

The main and most studied factor affecting the performance of Tomo-PIV methods is the density of particles, which is usually evaluated in terms of a non-dimensional number, which is a projected representation of the seeding density: the average number of particles per pixel (ppp) recorded on the camera CCD sensor. The ppp is linked to the spatial resolution: the higher the ppp is, the better the resolution is. However, higher densities lead to a dramatic increase of "ghost particles", *i.e.*, false positives related to matching ambiguities of indiscernible particle images. A good trade-off between ghosts and spatial resolution is typically found at around 0.05 ppp [39].

However, we have demonstrated in simulation studies [7][8] that several other factors – often neglected in performance evaluation studies – affect the quality of the tomographic reconstruction. The "added background particles", *i.e.*, particles that are visible to the cameras but not accounted for in the reconstruction, have been identified as a source of ghost particles in [19]. The authors attribute these added particles to the Gaussian profile of the laser sheet or to uncontrolled light reflections, thus implying that a proper experiment could avoid them. However, we have demonstrated from simple geometrical considerations that added particles are unavoidable in Tomo-PIV. Indeed, while the reconstruction is done in the intersection volume of the camera field of view, all cameras also record particles lying in the union of the fields of view and not in the intersection. These added particles act as a strong source of noise in the reconstruction because multiplicative inversion algorithms such as MART or SMART strive to explain their images by ghost particles in the intersection volume.

Another important factor that has only recently been identified is the defocusing occurring when trying to image volumes with cameras having a limited depth of field [40]. By means of a thorough simulation study [8], we have shown that neglecting the defocus in the imaging model leads to a significant increase in the number of missed detections. It is a strong motivation for designing new reconstruction frameworks, where defocus effects could be modeled and accounted for more easily, such as the one presented in the next Section.

Particle Volume Reconstruction (PVR)

As already mentioned, the observation model behind usual Tomo-PIV methods does not correspond to a physical model of the actual imaging process. The latter is illustrated in the right part of Figure 1. A particle can be considered as a 3D point source located at some 3D position, whose image is a "point-spread function" (PSF) centered on a 2D point given by the geometrical model of the camera. Most often, PSF are modeled as truncated Gaussian functions, parameterized by their standard deviation σ_{psf}. As a result, the image appears as the sum of P PSF functions, as in Equation (1), where x denotes the 2D position in the image plane, X_p is the 3D position of the p-th particle, h is the PSF function (which, in the general case, depends on the 3D position of the particle because of defocus effects) and F is the projective transformation from the 3D world to a 2D image, identified by the calibration.

$$I(x) = \Sigma_{p=1}^{P} E_p h_{X_p}\left(x - F\left(X_p\right)\right) \qquad (1)$$

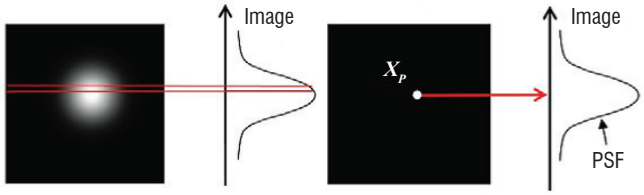


Figure 1 – Image formation models for 3DPIV. Traditional Tomo-PIV methods such as [19] are based on the "blob" model (left). In the proposed PVR model (right), particles are point sources for the camera, their image being the Point-Spread Function (PSF)

While this model underlines some previous works on Tomo-PIV, in particular the Iterative Particle Reconstruction (IPR) [44], it is only in [4] that it has been used to derive a linear problem suitable for standard inversion algorithms such as MART/SMART.

More precisely, we have demonstrated in [4] that the PVR model of Figure 1 can be approximated by a linear system $I = W\tilde{E}$, where I collects image intensities and the rows of the weighting matrix W are PSF samples $W_{kn} = h(k - n\Delta)$. The 3D field $\tilde{E}$ is a discrete approximate representation of the original Dirac-like particle field, where each particle is represented by a very restricted number of coefficients in neighboring voxels. As such, it is inherently a much sparser representation of the volume than that used in classical Tomo-PIV. Given $\tilde{E}$, the intensity and position of particles can be recovered unambiguously with subvoxel accuracy if they do not overlap. Particle overlapping in 3D space is extremely rare, given the typical values of the number of particles per volume. Thus, it is expected that the reconstructions $\tilde{E}$ will most often truthfully represent the particle distribution. In practice, one usually does not try to recover E and cross-correlate the discrete field $\tilde{E}$ to estimate velocity fields.

We have made comparative studies between two similar SMART algorithms (with MLOS initialization [2]), one, called Tomo-SMART, is based on the conventional Tomo-PIV model of [19] and the other is based on the proposed PVR model and is called PVR-SMART [4]. In all comparisons, PVR-SMART has been shown to increase the performance with respect to Tomo-SMART. For instance, in the simulation study illustrated in Figure 2, PVR-SMART consistently yields

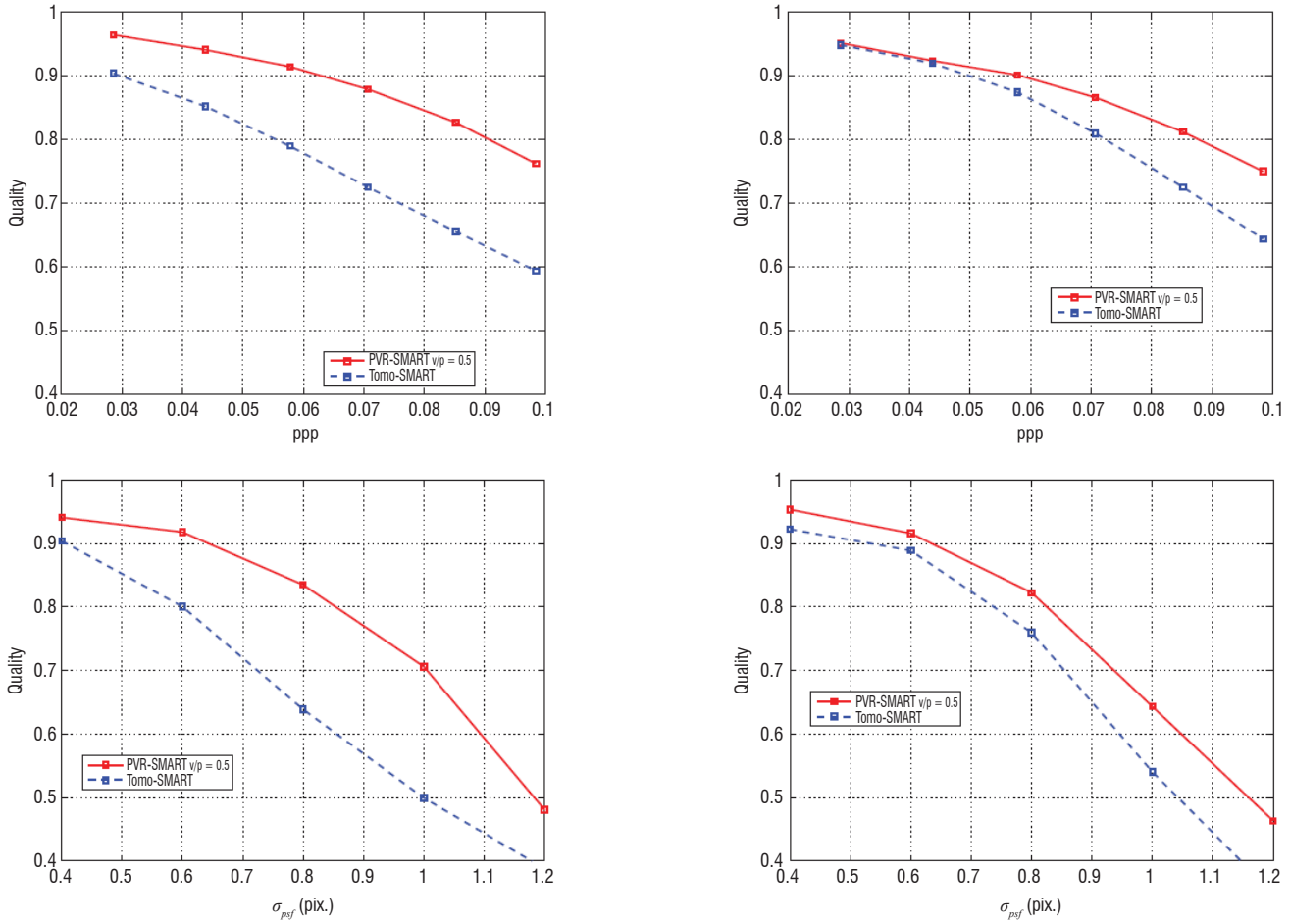


Figure 2 – Comparisons between conventional tomographic reconstruction and the proposed PVR approach, using the SMART algorithm for simulated images (see [4] for details). Top: curves for varying Nppp with $\sigma_{psf} = 0.6$, I/U = 0.47. Bottom: curves for varying σ_{psf} with I/U = 0.47, Nppp = 0.07,. Left: Q criterion of [19]. Right: fraction of detected true particles or Recall

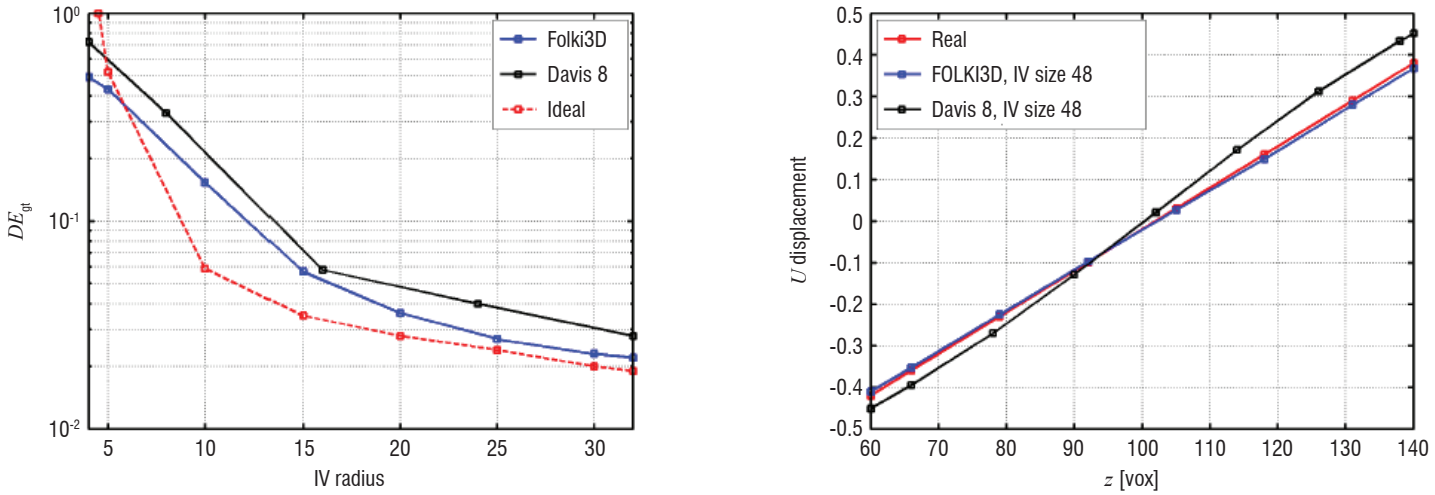


Figure 3 – Comparison of Davis 8.2 and FOLKI3D (IV radius of 24 voxel) on a synthetic shear layer type displacement field $(U,V,W)=(5+\alpha z,0,0)$ with $\alpha = 0.01$. Left: average displacement errors as a function of the IVs radius R – the red curve is FOLKI-3D applied on ground truth volumes. Right: displacement profiles along Z, averaged in the X, Y directions – red represents the true profile

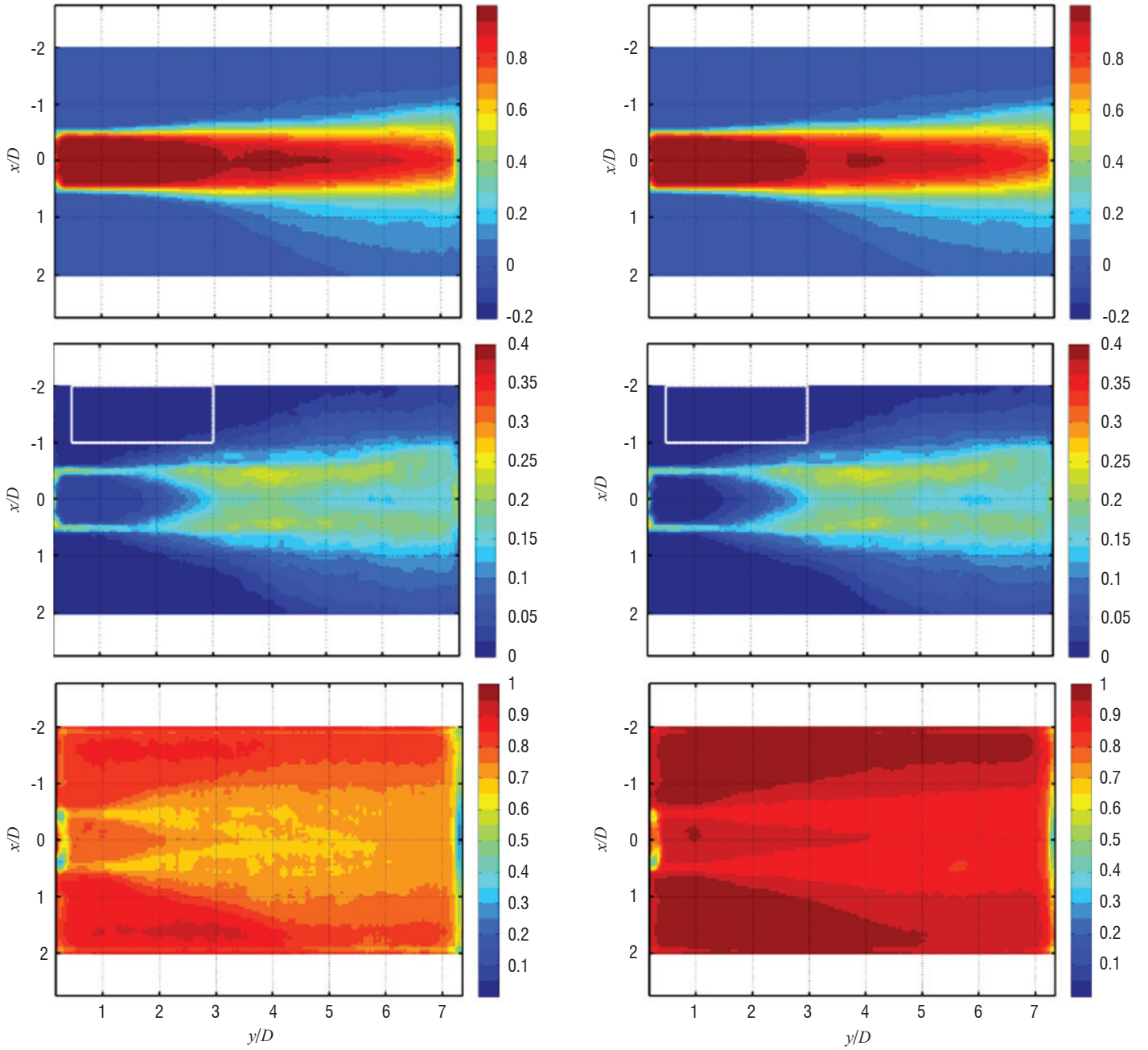


Figure 4 – Comparison of a 3D velocity field estimation by FOLKI3D from Tomo-SMART [19] (Left) and a PVR-SMART tomographic reconstruction (Right) – axial plane $z = 0$. Top to bottom: mean axial velocity, rms of axial velocity, correlation coefficient

better results than Tomo-SMART when varying the ppp (upper row) or the width of the PSF σ_{psf} (lower row). In the sequel, we describe other applications of the PVR model, first to estimate 3D displacement fields from volume correlation, then to derive a sparse approach.

Fast and dense 3D approach: PVR SMART and FOLKI3D

A 3D extension of FOLKI, FOLKI3D, has been proposed and characterized in [6][8]. It shares many characteristics of the 2D method. First, it is a dense calculation that provides the displacement field for all voxels. Second, with respect to the true displacement field, FOLKI3D exhibits a low-pass filter behavior closely related to the shape of the IW. Third, it has a highly parallel structure, allowing very high computational performance on GPUs. This algorithm has been evaluated on simulated volumes, in comparison with the commercial volume correlation software LaVision Davis 8.2. For the selected example, FOLKI3D outperforms Davis 8.2, both in terms of bias (Figure 3, right) and of rms error (Figure 3, left).

Experimental validations have also been conducted on a free round turbulent air jet at a Reynolds number Re = 4500 using a 3D-PIV setup made at ONERA together with a 2D-PIV camera providing reference measurements. The experimental set-up and measurement process, with advanced work concerning synchronization, illumination, seeding, calibration (including PSF calibration) and image preprocessing, are described at length in Ref. [8]. We present here only one comparison of 3D velocity fields, estimated using FOLKI3D from volumes reconstructed either by Tomo-SMART or by PVR-SMART.

Figure 4 first presents the mean velocity estimation (averaged over 300 snapshots). While results appear at first to be quite similar, comparisons of the noise level in regions with a low level of velocity fluctuation (the white rectangle in the images of the middle row) or of the correlation coefficient (last row of Figure 4) reveal that velocity fields estimated from PVR-SMART volumes are significantly less noisy than those estimated from Tomo-SMART. Other evaluations, presented in [8], confirm this difference. It could stem from a higher proportion of ghost particles and/or a higher peak-locking effect in the Tomo-SMART reconstruction method.

We then consider the reconstruction of instantaneous 3D flow structures of the jet, by visualizing selected iso-contours of axial velocity, and of axial and azimuthal vorticity, in Figure 5. Quantities and levels have been chosen so as to highlight the most important features of the jet near-field dynamics, involved in the first stages of mixing. As described for instance in [9][16], fluctuations are mostly dominated by axisymmetric structures of azimuthal vorticity (vortex rings, stemming from the Kelvin-Helmholtz instability), associated with axial velocity fluctuations in the core, and streamwise vortices contained in the mixing layer. When comparing the methods, the global view of the first row indicates that, again, both results are quite consistent. However, looking at the zoomed-in region (lower row of the figure), one can see that PVR-SMART estimates rounder and smoother vortex rings and that the counter-rotating streamwise vortices, which are the key ingredients of jet mixing, appear to be longer, smoother and bigger for PVR-SMART.

To conclude, we have developed an original tomographic PIV pipeline, revisiting the two stages of the classical approach, tomographic reconstruction and 3D correlation. This method is now routinely used at ONERA. Another approach, oriented towards particle localization and tracking, is described below.

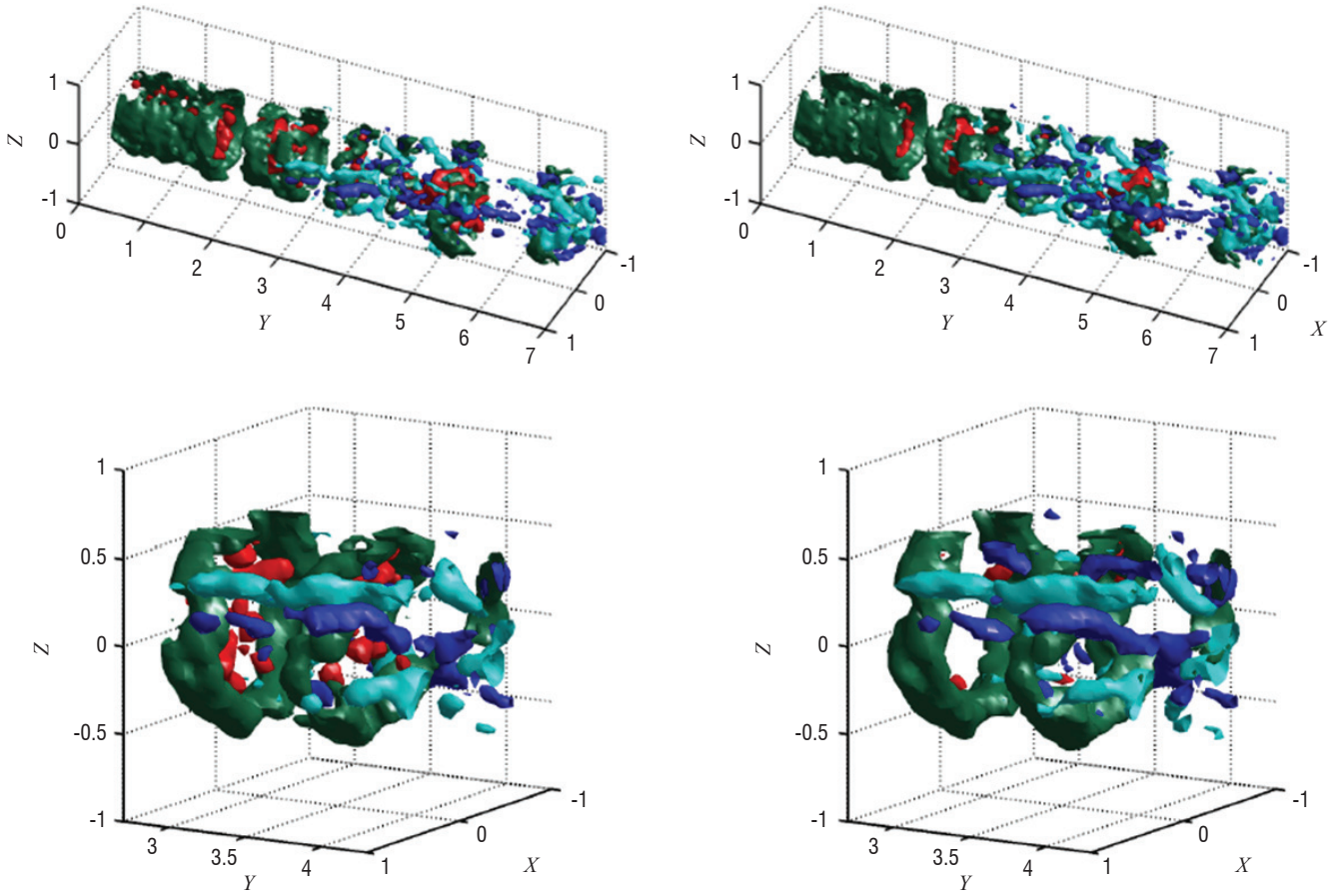


Figure 5 – Reconstruction of instantaneous 3D structures of the jet from Tomo-SMART (left) and PVR-SMART (right) reconstructed volumes. Isovalue of the axial velocity 1.05 V0 (red), Isovalue of the vorticity components $\omega_\theta = 2$ (green), $\omega_z = 1.2$ (cyan) and -1.2 (blue)

Back to particles: an efficient sparse Tomo-PIV on a discretized volume

As a consequence of experimental (seeding density) and numerical choices (spatial discretization), the sought volumes in 3DPIV are very sparse – typical choices lead to a mean density of 1^{-4} particle per voxel. Classical multiplicative algorithms, such as (S)MART, tend to concentrate the energy of the reconstruction on a limited number of "blobs", but they still cannot be considered as exploiting sparsity. Indeed, sparse algorithms are aimed at solving a linear system with a vector (*i.e.*, the discretized volume) having a fixed, limited number of non-zero components (the detected particles). Sparse algorithms have been the subject of important developments within the Signal Processing community, partly motivated by compressed sensing results published in the mid-2000s [18][3]. Few attempts were then made to apply such techniques to tomographic PIV [37][45], but they lead to computationally demanding methods and were applied to small, synthetic volumes. In 2013, based on the PVR model, we proposed the first computationally efficient sparse processing chain for tomographic PIV and demonstrated its performance in terms of the particle detection rate, compared to standard algorithms [10][11].

The initial step is a variation over the classical multiplicative line of sight (MLOS) of [2], where only the local maxima of the MLOS field are retained. Simulation studies have demonstrated that this "LocM" (Local Maximum) strategy allows a drastic reduction of the problem dimensionality, with a limited loss in terms of missed particles. Thus, we propose to solve the Tomo-PIV problem on a discretized 3D grid with a voxel-to-pixel ratio (v/p) equal to 0.5. Using the PVR formulation, the sparse problem concerned is written as

$$\min_E \|WE - I\|^2 \text{ subject to } \|E\|_0 \leq S \qquad (2)$$

where the L0 norm is the number of non-zero components in E, and S, the sparsity number, is the total number of particles in the reconstructed volume. The solution of the previous problem on a discretized 3D grid can be efficiently found with CoSaMP (Compressed Sampling Matching Pursuit) [31]. CoSaMP, a popular algorithm for sparse inversion, builds on the matching pursuit strategy aimed at identifying the support of the solution, *i.e.*, the subset of voxels that contains a particle. This is done by iteratively removing or adding voxels to the current support and using it to solve the problem, *i.e.*, to determine the intensities E_p of the particles that minimize the goodness-of-fit term of (2).

The main parameters of this algorithm, the discretization step and the sparsity number, have been studied in [11]. A discretization step corresponding to a voxel-to-pixel ratio (v/p) equal to 0.5 has been shown to provide a good trade-off between localization accuracy and computation time. The sparsity number can be chosen according to the size of the imaged volume and the seeding density. Comparison with the classical MLOS-SMART approach has also been provided on synthetic and real datasets, showing better performance of LocM-CoSaMP over a large range of seeding densities, as shown in Figure 6.

Two-time-step Tomo-PTV with sparse tomographic reconstruction

From the sparse reconstruction presented before, it is tempting to seek the individual matching of each reconstructed particle from one time instant to the next. This process can be related to PTV (Particle Tracking Velocimetry) [29]. Matching is very ambiguous, because particles are indistinguishable and PTV was traditionally restricted to low seeding density. Recent developments combining an iterative reconstruction algorithm and particle tracking over several time steps [41] have shown interesting results in terms of seeding density. Still, two-time-step PTV remained unsolved for a seeding density higher than 0.005 ppp. In this context, we have proposed a novel "tomoPTV" method, combining all previous results (*i.e.*, PVR model, FOLKI3D and LocM-CoSaMP) [12][13]. We start from the LocM-CoSaMP reconstruction of the particle volume at two time instants. The pointwise LocM-CoSaMP reconstructions are expanded with Gaussian filtering, and a low-resolution first estimation of the 3D displacement field is computed using FOLKI3D. Particle matching is then done by nearest-neighbor search within a limited

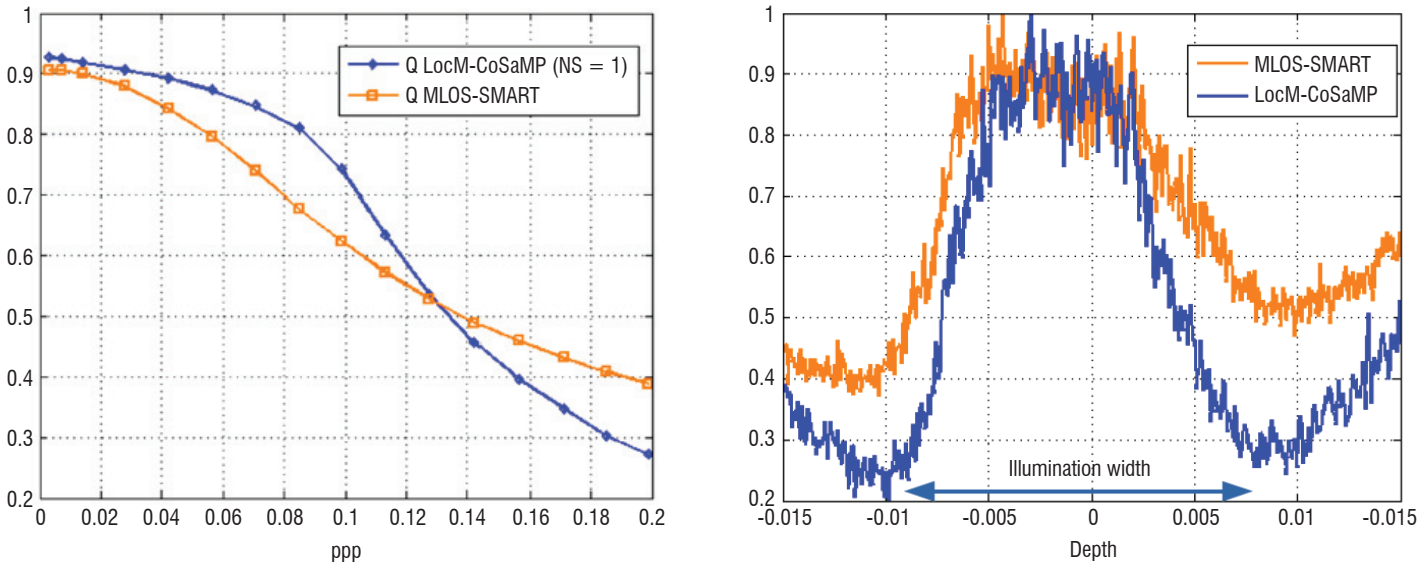


Figure 6 – Comparisons of tomographic reconstructions of volumes of particles using MLOS-SMART (orange colored curves) and LocM-CoSaMP (blue curves). Left: Q-factors [19] in synthetic experiments show that the LocMCoSaMP method outperforms MLOS-SMART on seeding densities up to ppp ≈ 0.12. Right: the Z intensity profiles of reconstructions on real data show a higher R-SNR (ratio between the plateau and edge values) for LocM-CoSaMP, indicating a higher efficiency to preserve real particles and remove ghosts [39]. Details on simulation and experimental conditions are given in [12]

region around the position predicted by FOLKI3D. Finally, subvoxel location of particles is identified by the iterative optimization of

$$\min_{E,X} \sum_j \sum_x \left\{ I_j(x) - \sum_p E_p h\left(x - F_j\left(X_p\right)\right) \right\}^2 \qquad (3)$$

where j refers to the camera index. Changes of location are restricted for each particle to a neighborhood around the discrete position given by the matching process. Note that, contrarily to [44][41], the optimization is done jointly for all particle positions, to account for overlapping particles.

Experimentation on the same low-speed round air jet experiment, as in the previous section, has been carried out to compare classical Tomo-SMART of [19] followed by FOLKI3D to the two-time-step Tomo-PTV technique. Figure 7 presents the U,V components of a slice of the 3D displacement field in the central plane of the jet. It can be seen that the jet structure and the development of its vortical structures (patches of alternate colors that identify Kelvin-Helmholtz azimuthal vortices) in the shear layer are nicely captured by both methods, which yield similar results. This demonstrates the ability of the proposed two-time-step Tomo-PTV technique to yield reliable results with experimental datasets.

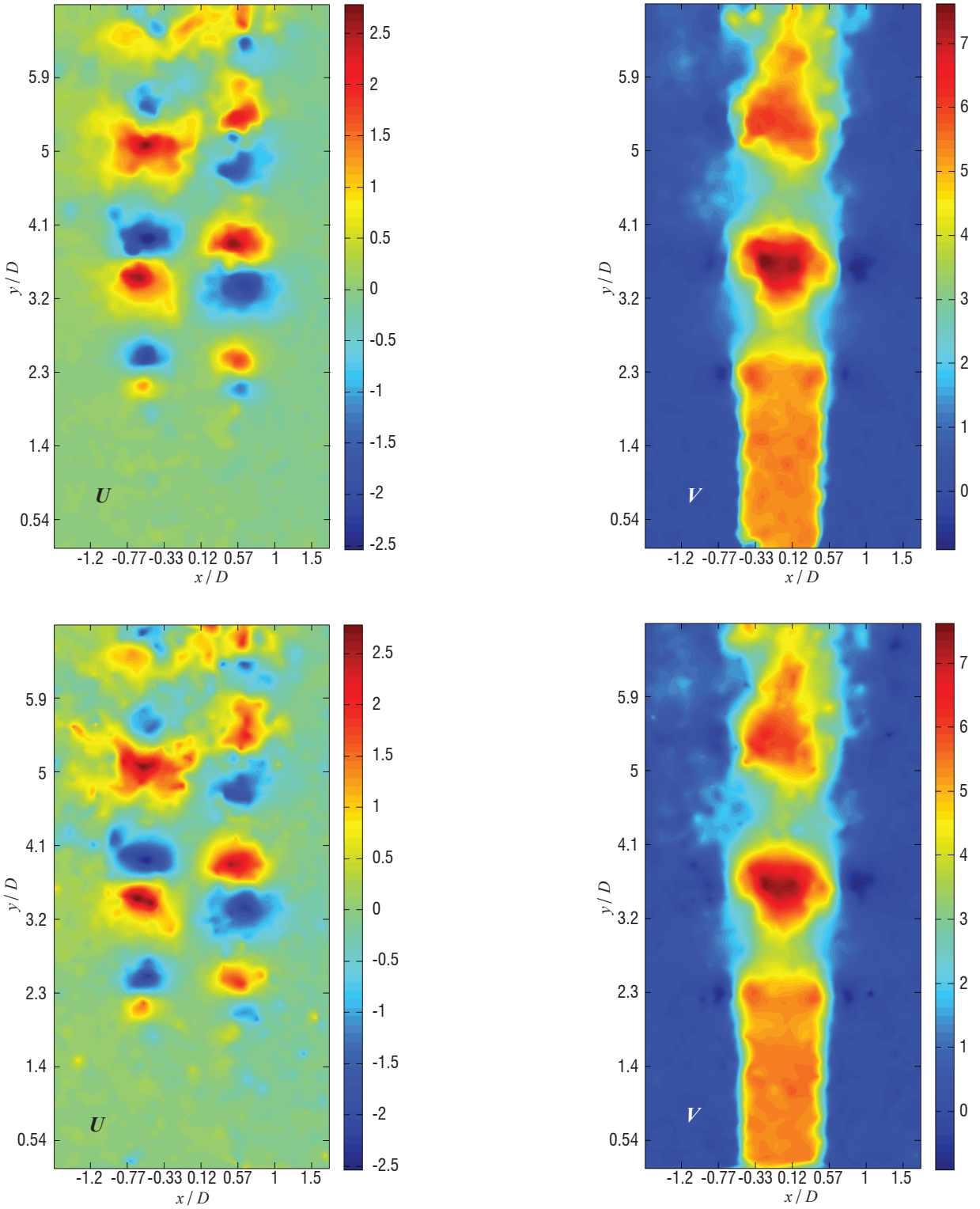


Figure 7 – Displacement fields for a round air jet snapshot. U is collinear to the light propagation and V is collinear to the jet axis.
Top: Tomo-SMART of [19] + FOLKI3D; Bottom: LocM-CoSaMP + PTV

A direct method for instantaneous 3DBOS

Introduction

Density field visualization, such as Schlieren techniques, has been used extensively to understand fluid mechanics phenomena such as shock waves or thermal effects. However, the mostly qualitative information resulting from this measurement does not yield enough information to refine computational fluid dynamics codes. Background-oriented Schlieren (BOS) is one way to obtain a quantitative measurement of density gradients [15][30]. It is based on the observation of ray deviations through a medium with inhomogeneous optical index. The BOS optical setting, whose principle is recalled in Figure 8, is remarkably simple: it only requires that the flow under study be placed between a camera and some textured background on which the camera is focused. Comparing the images of the background with and without the flow using digital image correlation techniques reveals displacements, which are the projections of the light ray deviations. Displacement fields are computed by digital image correlation algorithms, here using FOLKI-SPIV [5]. Conversion from displacement to deviation is usually done by approximating the curved ray by two straight lines intersecting at a "deviation point" located somewhere in the reconstructed volume. In the sequel, the deviation point is chosen on the unperturbed ray and equidistant from the entry and exit points.

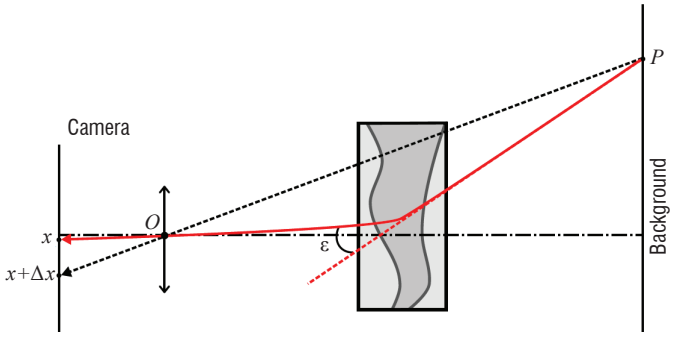


Figure 8 – Principle of BOS measurement: the camera is represented on the left hand side by its center of projection (O), input lens and image plane. It is focused on the background plane on the right. The flow under study, shown in the middle, induces a deviation of the light rays. A point P of the background is imaged at position x with the flow and at $x+\Delta x$ without. Note that the size of the region under study and the ray deviation angle are heavily exaggerated for the sake of clarity

The deviation angle is the main observable of BOS. Considering one ray to be going through the flow under study, the deviation angle ε is equal to the integral of the optical index gradient along the optical path:

$$\varepsilon = \frac{G}{n_0}\int_{ray\subset flow} \nabla\rho(s)\,ds \qquad (4)$$

where n_0 is the optical index outside the flow, and G is the Gladstone-Dale constant which relates the optical index and the density through $n - 1 = G\rho$ (for standard conditions $G = 0.23 * 10^{-3}$ m^3/kg). This problem is generally nonlinear, since the integration path depends on the unknown variable density of the flow. Similar to most other references, we assume a paraxial hypothesis and integration is done along the (straight) unperturbed ray.

A one-step method for 3DBOS reconstruction

If several synchronized cameras are placed all around the flow under study, each one facing a textured background, deviations along several rays passing through the medium with various orientations can be measured. Stacking these deviations in a single observation vector and discretizing Eq. (4), a linear system can be written as:

$$\varepsilon = \begin{bmatrix}\varepsilon_x\\ \varepsilon_y\\ \varepsilon_z\end{bmatrix} = A\rho = T\begin{bmatrix}D_x\\ D_y\\ D_z\end{bmatrix}\rho \qquad (5)$$

3DBOS amounts to inverting this linear system to estimate the discretized density ρ. The observation matrix $\boldsymbol{A}$ is often separated into spatial finite difference approximations of the gradient in each direction (e.g., $\boldsymbol{D}_x$, $\boldsymbol{D}_y$, $\boldsymbol{D}_z$) and a tomographic operator $\boldsymbol{T}$. For this reason, almost all previous references [1][27][36] adopt a two-step inversion scheme, where, first, components of the spatial gradients in all spatial directions are computed by three independent tomographic reconstructions, and, second, density is obtained by spatial integration. The main benefit is to be able to rely on known methods for each step, in particular algorithms originating from Computerized Tomography (CT) for the inversion of $\boldsymbol{T}$. In 2012, we have proposed a direct (or one-step) inversion method that considers the end-to-end observation matrix A [42][32]. In this approach it is easy to use the formalism of regularized inversion by penalization [23], where the density volume is sought as the minimizer of a compound criterion:

$$\mathcal{J} = \|\boldsymbol{A}\rho - \varepsilon\|^2 + \lambda\mathcal{R}(\rho) \qquad (6)$$

Here $\mathcal{R}$ is some regularization term, for instance the L2 norm [32] or L2L1 norm [42] of the spatial gradient of the density volume, and $\lambda > 0$ is a regularization parameter chosen according to an L-curve strategy [22]. Optimization is made with a classical conjugate gradient algorithm, with explicit step computation [42]. It should be mentioned that this problem is of very high dimensions, with typically 107 deviations and 108 reconstructed voxel values, even if we use a user-defined mask to reduce the support of the optimization. The most computationally demanding operations are the application of the observation operator and its adjoint, for instance to compute the gradient of the first term of the criterion $2\boldsymbol{A}^t(\boldsymbol{A}\rho - \varepsilon)$. The implementation is derived from the projection and back-projection studied in computerized tomography [21] and makes use of the parallel architecture of the GPU to limit computation time.

This one-step numerical inversion method has been studied on synthetic and real datasets using 12 cameras in Ref. [32]. Non-coplanar configurations of the camera setup were considered, using the Geode experimental bench dedicated to 3DBOS made at ONERA/DMAE (Toulouse) and illustrated in Figure 9 (left image). The experimental results

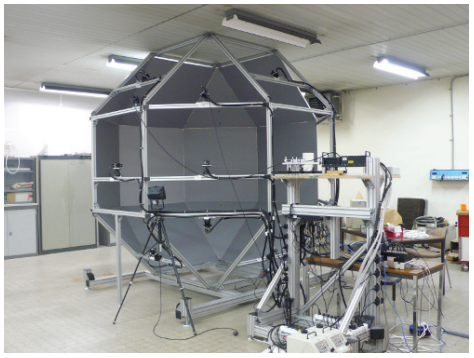


Figure 9 – Left: the "Geode" 3DBOS bench at DMAE allows simultaneous BOS acquisitions from several points of view by up to 12 cameras, each facing a background panel. Right: 3D reconstruction of the plume of a candle with 12 cameras in a 180° planar configuration [32] – Iso-density surfaces at densities 0.7 (red), 0.9 (green) and 1.1 (blue)

have demonstrated the ability of the method to capture complex non-symmetric volumes, such as a candle convective flow or the helicoidal plume of a rotating gas burner, as shown in Figure 9 (right image).

Experimental demonstration of 3DBOS in a research wind tunnel

The 3DBOS technique presented above was demonstrated in a large research wind tunnel in 2014 [34]. The experiment was conducted in ONERA's F2 wind tunnel located in the Fauga-Mauzac center. This subsonic facility can reach a speed of 100 m.s^{-1} and the test section is 1.4 m wide, 1.8 m high and 5 m long. Large window panels provide very good access for optical diagnostics. A combination of screen honeycombs, an acoustics dumper and a contraction ratio of 12 contribute to a turbulence level below 0.05%.

The tested flow is a co-flowing hot jet (with a total temperature of 100°C) generated at the wingtip of a simplified half-wing fixed on the floor of the test section. The upstream wind tunnel flow has a speed of 20 m.s-1 with total temperature and pressure equal to ambient conditions. The 3DBOD system consists of 12 cameras distributed on one side and the ceiling of the test section, the ground and the other side being dedicated to background panels. In parallel, the hot co-flowing jet is also investigated with the stereo PIV technique and thermocouple measurements. In particular, thermal measurements allow a complete comparison with 3DBOS. Figure 10 presents some illustrations of the experimental set-up.

Three-dimensional density fields of the unsteady jet flow have been successfully obtained by the one-step inversion method previously described, as shown in Figure 11 (upper row), and mean density

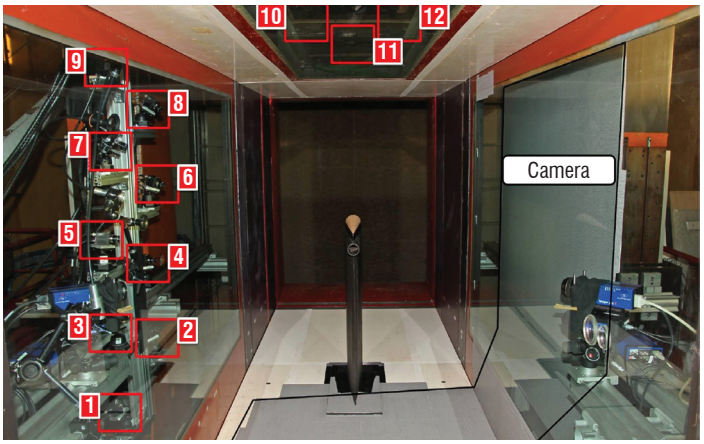


Figure 10 – 3DBOS experiment in ONERA's F2 research wind tunnel (Fauga-Mauzac, France). The twelve cameras are represented with red squares, while the background panels are shown as gray regions. The two cameras of the stereo PIV system are also visible (in blue)

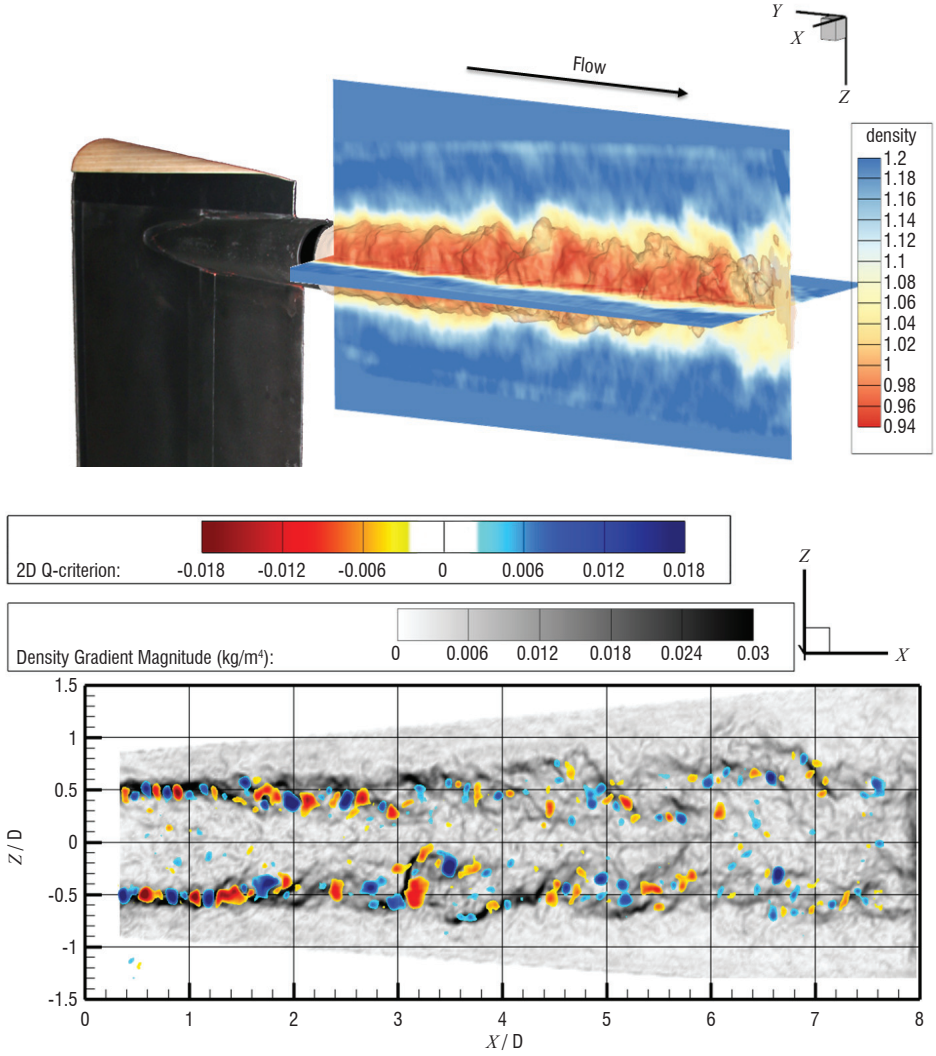


Figure 11 – Instantaneous measurement of a flowing hot jet in ONERA's F2 research wind tunnel. Top: 3DBOS reconstruction of the density field of the unsteady jet flow shown next to the wingtip. Bottom: Density gradient magnitude in gray scale superimposed with the most energetic Q-criterion structures computed on stereo-PIV data. Experimental conditions: Tjet = 100 °C; Qjet = 80 g/s; Vext = 20 m.s^{-1}

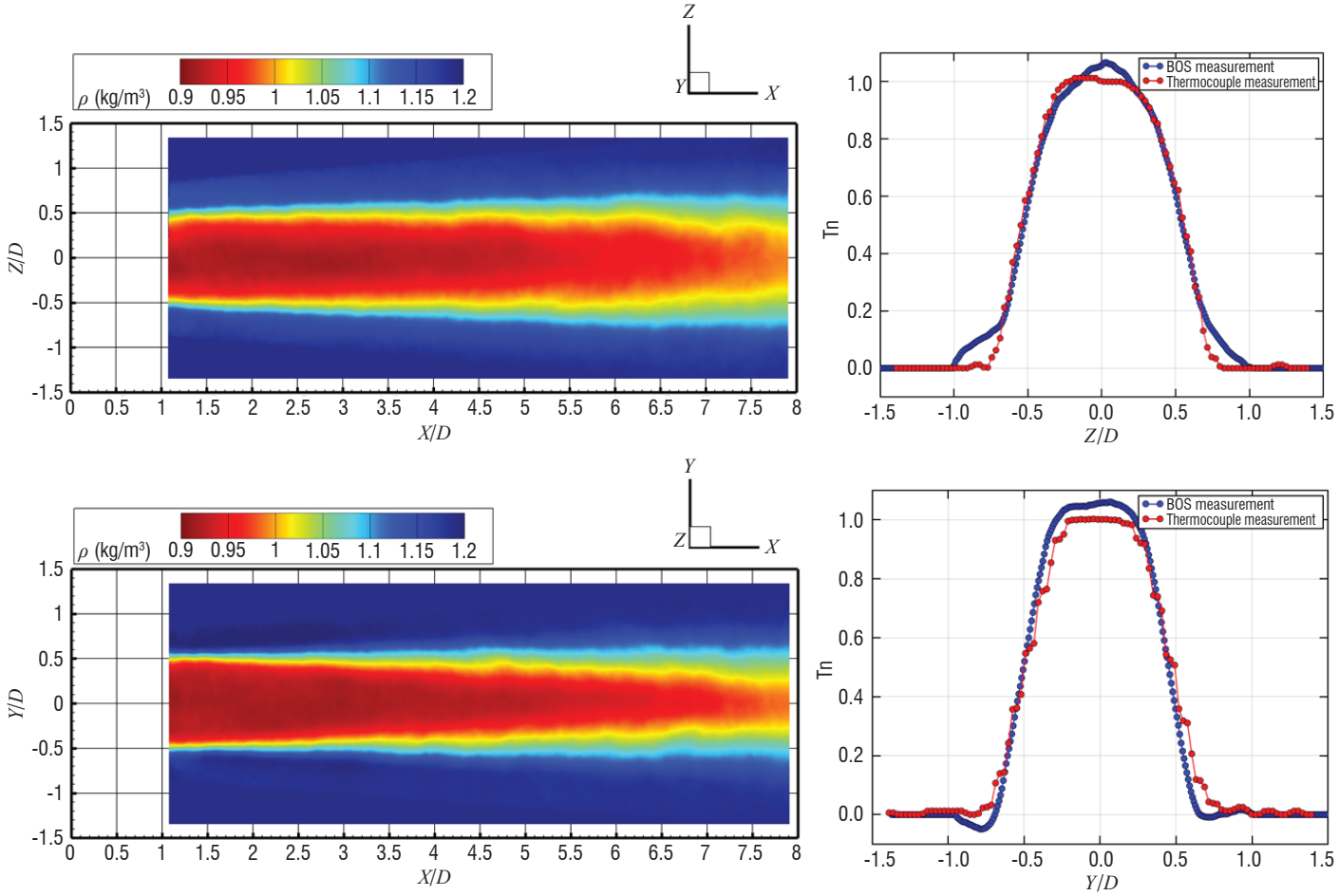


Figure 12 – Mean density reconstructed by 3DBOS and comparison with thermal measurement (Experimental conditions : $V = 20$ m.s^{-1}, $Q = 80$ g.s^{-1}, Tjet $= 100$°C). Left column: Longitudinal profiles of the averaged density field (top: $Y = 0$; bottom: $Z = 0$). Right column: normalized temperature profiles from 3DBOS and from thermocouples (top: vertical; bottom: horizontal) in the $X/D = 2.8$ slice.

fields reconstructed on a 116 Megavoxel volume with a spatial resolution of 0.5 mm are presented in Figure 12 (left column).

In this configuration, the influence of pressure on density can be neglected and temperature maps can then be extracted directly from the density field. Both measurements are compared on a mean flow field in Figure 2, using a normalized temperature to account for the changes in the wind tunnel flow external temperature (see details in Ref. [34]). This comparison shows a very good agreement between both techniques. The small overprediction in temperature noticed for the BOS measurement is below the thermocouple measurement uncertainty. This shows the potential of 3DBOS for providing the full 3D temperature field in non-compressible flows with known pressure distribution.

Moreover, Figure 11 (bottom row) presents a superimposition of the instantaneous stereo-PIV and 3DBOS results illustrated, respectively, by the Q-criterion and the density gradient magnitude maps in the same plane. Note that the 3DBOS results have been obtained with only 9 cameras, in order to avoid reciprocal illumination effects of the two measuring systems. However, the degradation in spatial resolution is limited within the visualization plane. This representation nicely emphasizes that the most energetic Q-criteria are located in the mixing layer around the Kelvin-Helmholtz windings revealed by BOS.

Achieving 3DBOS reconstruction of compressible flows

Our main concern is now the extension of the proposed 3DBOS method to flows with higher density gradients, which are responsible for strong and highly inhomogeneous blurring effects in recorded images. Nevertheless, we have been able, thanks to a refined model of the observation and careful choice of experimental conditions, to produce deviation maps with a spatial resolution comparable to that of Schlieren visualization, as illustrated in Figure 13.

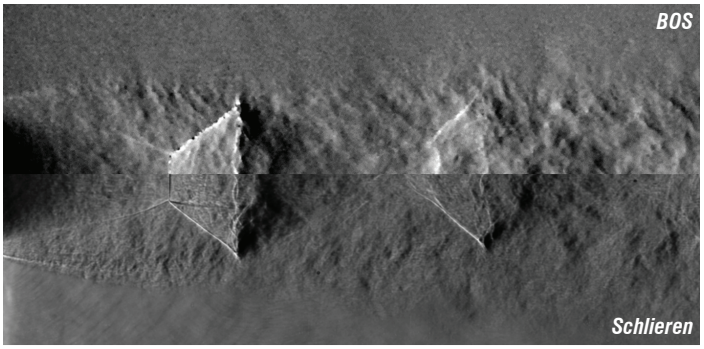


Figure 13 – Experimental images of an under-expanded jet with NPR = 5. With a careful choice of the experimental conditions, the BOS deviation image (top half-image) exhibits a spatial resolution comparable to that of traditional Schlieren (bottom half-image)

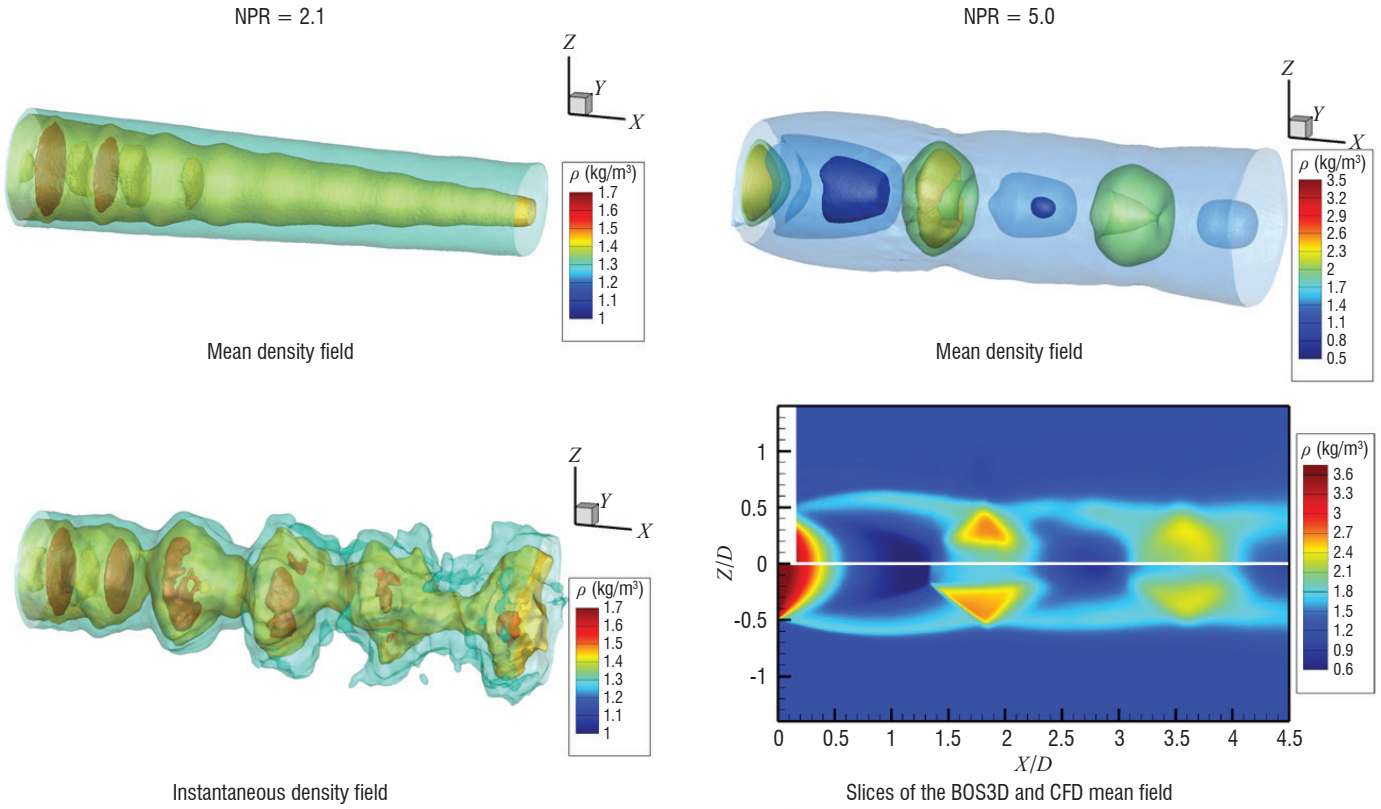


Figure 14 – 3DBOS reconstruction of the density field of an under-expanded flow with NPR = 2.1 (left column) and NPR = 5 (right column). Top row: mean density field from 900 recordings. Bottom right: instantaneous density field for NPR = 2.1. Bottom left: comparison of 3DBOS and CFD mean density field for NPR = 5.0. The calculations are done over volumes of 10 x 7 x 7 cm³ using between 40 and 100 million 0.2 mm-side voxels. More details can be found in [33].

However, contrarily to Schlieren techniques, BOS measurement opens the way to quantitative 3D reconstruction of the flow. In a very recent demonstration [33], the 3DBOS reconstruction technique presented above was applied to deviation data produced by a dedicated 12-camera bench on a underexpanded jet with variable nozzle pressure ratio (NPR). Figure 14 shows results for NPR = 2.1 and NPR=5.0, with 3D mean density field on the top row. In the NPR = 2.1 case, the train of shock waves is clearly visible in the potential core region. The two first shocks are found in the instantaneous flow (bottom left image), while the jet topology is much more complex further downstream as it becomes destabilized. For NPR = 5, a Mach disk is clearly visible. The 3DBOS reconstruction shows a very good agreement with a DNS simulation. Both the density levels and the flow topology are well recovered, with a consistent location of the Mach disk and shock-cells. To the best of our knowledge, these reconstructions of under-expanded flows presented here and in Ref. [33] have never been obtained before.

Perspectives

Thanks to a joint investment by three scientific departments, 3DPIV and 3DBOS are now mastered at ONERA. To the best of our knowledge, our 3DBOS reconstructions of instantaneous flows have no equivalent in the published literature of experimental fluid mechanics. With regard to 3DPIV, our developments have many more competitors, however, the very recently released results of the 4th PIV Challenge [24] indicate that our results are among the best (for Case C of the Challenge).

This paper focuses on the underlying numerical developments; let us emphasize that several other skills have been acquired along the way, particularly in terms of experimental methodology, relating to subjects such as camera set-up, seeding techniques and background design, illumination tools, camera models, calibration methods, estimation error sources, etc. Some of these advances have already been published [7][8][14][34][26][32] and others will be the subject of future publications.

The perspectives of 3DPIV include algorithmic development for temporal processing, either according to the dense correlation paradigm, as a 3D extension of our recent proposal [46], or within the particle tracking framework in line with [41]. As regards 3DBOS, we are working on a new direct model and associated estimation methods for high gradients or shock wave situations. In this context, we are also currently comparing 3DBOS with other optical measurement methods, such as digital holography [17][35].

As illustrated in Figure 11, joint 3D BOS and PIV measurements have already been conducted: further investigations in this line are planned for the experimental study of compressible flows, a domain where imaging measurement techniques are challenged by aero-optical effects. The information contributed by BOS could be useful for more accurate velocimetry, for instance.

Finally, 3D field measurement is also a gateway towards tighter coupling between experiments and numerical simulations, a field which we believe will be of major importance in future aerodynamic studies ■

References

[1] B. ATCHESON, I. IHRKE, W. HEIDRICH, *et al.* - *Time-Resolved 3D Capture of Non-Stationary Gas Flows*. ACM Transactions on Graphics (TOG), ACM, 2008. p. 132.

[2] C. ATKINSON, J. SORIA - *An Efficient Simultaneous Reconstruction Technique for Tomographic Particle Image Velocimetry.* Experiments in Fluids, 2009, vol. 47, no 4-5, p. 553-568.

[3] E. J. CANDÈS - *Compressive sampling*. Proceedings of the International Congress of Mathematicians, 2006, p. 1433-1452.

[4] F. CHAMPAGNAT, P. CORNIC, A. CHEMINET, *et al.* - *Tomographic PIV: Particles Versus Blobs*. Measurement Science and Technology, 2014, vol. 25, no 8, p. 084002.

[5] F. CHAMPAGNAT, A. PLYER, G. LE BESNERAIS, *et al*. - *Fast and Accurate PIV Computation Using Highly Parallel Iterative Correlation Maximization*. Experiments in Fluids, 2011, vol. 50, no 4, p. 1169-1182.

[6] A. CHEMINET, B. LECLAIRE, F. CHAMPAGNAT, *et al*. - *Accuracy Assessment of a Lucas-Kanade Based Correlation Method for 3D PIV.* 17th International Symposium on Applications of Laser Techniques to Fluid Mechanics, 2014.

[7] A. CHEMINET, B. LECLAIRE, F. CHAMPAGNAT, *et al*. - *On Factors Affecting the Quality of Tomographic Reconstruction*. PIV13, 10th International Symposium on Particle Image Velocimetry, Delft, The Netherlands, July 1-3, 2013.

[8] A. CHEMINET - *Tomographic PIV Development for the Study of Turbulent Flows*. PhD dissertation, Université Paris-Saclay, 2016.

[9] J. H. CITRINITI, W. K. GEORGE - *Reconstruction of the Global Velocity Field in the Axisymmetric Mixing Layer Utilizing the Proper Orthogonal Decomposition*. Journal of Fluid Mechanics, 2000, vol. 418, p. 137-166.

[10] P. CORNIC, F. CHAMPAGNAT, A. CHEMINET, *et al*. - *Computationally Efficient Sparse Algorithms for Tomographic PIV Reconstruction.* PIV13, 10th International Symposium on Particle Image Velocimetry, Delft, The Netherlands, July 1-3, 2013.

[11] P. CORNIC, F. CHAMPAGNAT, A. CHEMINET, *et al.* - *Fast and Efficient Particle Reconstruction on a 3D Grid Using Sparsity*. Experiments in Fluids, 2015, vol. 56, no 3, p. 1-7.

[12] P. CORNIC, F. CHAMPAGNAT, A. PLYER, B. LECLAIRE, A. CHEMINET and G. LE BESNERAIS - *Tomo-PTV with Sparse Tomographic Reconstruction and Optical Flow*. 17th International Symposium on Applications of Laser Techniques to Fluid Mechanics, Lisbon, 2014.

[13] P. CORNIC, B. LECLAIRE, A. CHEMINET, *et al*. - *Two Time Steps Tomo-PTV with Sparse Tomographic Reconstruction Versus Tomo-PIV*. PIV15, 11th International Symposium on Particle Image Velocimetry, Santa Barbara (California, USA), 2015.

[14] P. CORNIC, C. ILLOUL, Y. LE SANT, *et al*. - *Calibration Drift within a Tomo-PIV Setup and Self-Calibration*. PIV15, 11th International Symposium on Particle Image Velocimetry, Santa Barbara (California, USA), 2015.

[15] S. B. DALZIEL, G. O. HUGHES and B. R. SUTHERLAND - *Whole-Field Density Measurements by "Synthetic Schlieren".* Experiments in Fluids, 2000, vol. 28, no 4, p. 322-335.

[16] S. DAVOUST, L. JACQUIN, B. LECLAIRE - *Dynamics of m=0 and m=1 Modes and of Streamwise Vortices in a Turbulent Axisymmetric Mixing Layer*. Journal of Fluid Mechanics, 2012, vol. 709, p. 408-444.

[17] J.-M. DESSE, P. PICART and P. TANKAM - *Digital Three-Color Holographic Interferometry for Flow Analysis.* Optics Express, 2008, vol. 16, no 8, p. 5471-5480.

[18] D. L. DONOHO - *Compressed Sensing*. IEEE Transactions on Information Theory, 2006, vol. 52, no 4, p. 1289-1306.

[19] G. E. ELSINGA, F. SCARANO, B. WIENEKE, *et al*. - *Tomographic Particle Image Velocimetry.* Experiments in Fluids, 2006, vol. 41, no 6, p. 933-947.

[20] G. E. ELSINGA, J. WESTERWEEL, F. SCARANO, *et al*. - *On the Velocity of Ghost Particles and the Bias Errors in Tomographic-PIV.* Experiments in fluids, 2011, vol. 50, no 4, p. 825-838.

[21] N. GAC, A. VABRE, A. MOHAMMAD-DJAFARI, *et al*. - *GPU Implementation of a 3D Bayesian CT Algorithm and its Application on Real Foam Reconstruction*. The First International Conference on Image Formation in X-Ray Computed Tomography, 2010, p. 151-155.

[22] P. C. HANSEN - *Analysis of Discrete Ill-Posed Problems by Means of the L-Curve*. SIAM review, 1992, vol. 34, no 4, p. 561-580.

[23] J. IDIER - *Bayesian approach to Inverse problems*. 2010, vol. 35, Wiley, New York.

[24] C. KAHLER, T. STARITA, P. P. VLACHOS, *et al.* - *Main Results of the 4th International PIV Challenge*. Experiment in Fluids, may 2016.

[25] G. LE BESNERAIS, F. CHAMPAGNAT, A. PLYER, *et al*. - *Advanced Processing Methods for Image-Based Displacement Field Measurement.* AerospaceLab, 2009, no 1, p. 1-12.

[26] Y. LE SANT, V. TODOROFF, A. BERNARD-BRUNEL, *et al*. - *Multi-Camera Calibration for 3DBOS*. 17th International Symposium on Applications of Laser Techniques to Fluid Mechanics, 2014.

[27] B. LECLAIRE, C. BROSSARD, R. COURTIER, *et al*. - *Particle Image Velocimetry : a Key Contributor to Aerospace Research*. Aerospace Lab, 2016.

[28] F. LEOPOLD, M. OTA, D. KLATT, *et al*. - *Reconstruction of the Unsteady Supersonic Flow Around a Spike Using the Colored Background Oriented Schlieren Technique*. J. Flow Control Measur. Vis., 2013, 1(2):69-76.

[29] H. G. MAAS, A. GRUEN and D. PAPANTONIOU - *Particle Tracking Velocimetry in Three-Dimensional Flows*. Experiments in Fluids, 1993, vol. 15, no 2, p. 133-146.

[30] G. MEIER - *Computerized Background-Oriented Schlieren*. Experiments in Fluids, 2002, vol. 33, no 1, p. 181-187.

[31] D. NEEDELL and J. A. TROPP - *CoSaMP: Iterative Signal Recovery from Incomplete and Inaccurate Samples*. Applied and Computational Harmonic Analysis, 2009, vol. 26, no 3, p. 301-321.

[32] F. NICOLAS, V. TODOROFF, A. PLYER, *et al*. - *A Direct Approach for Instantaneous 3D Density Field Reconstruction from Background-Oriented Schlieren (BOS) Measurements*. Experiments in Fluids, 2016, vol. 57, no 1, p. 1-21.

[33] F. NICOLAS, D. DONJAT, O. LEON, *et al*. - *3D Reconstruction of Compressible Flow by Synchronized Multi-Camera BOS*. Submitted to Experiments in Fluids (2nd revision), dec. 2016.

[34] F. NICOLAS, D. DONJAT, A. PLYER, *et al*. - *Experimental Study of a Co-Flowing Jet in ONERA's F2 Research Wind Tunnel by 3D Background Oriented Schlieren*. Submitted in Meas. Sci. Technol., dec. 2016.

[35] F. OLCHEWSKY, J.-M. DESSE, F. NICOLAS, *et al.* - *3D Reconstruction of Helium Jet by Multidirectional Polarized White Light Differential Interferometry*. Symposium on Flow Visualization, Knoxville, USA, 2016.

[36] K. OTA and M. HAMADA - *Quantitative 3D Density Measurement of Supersonic Flow by Colored Grid Background Oriented Schlieren Technique.* ICAS, 2010, p. 23.

[37] S. PETRA and C. SCHNÖRR - *Tomo-PIV Meets Compressed Sensing*. Pure Math Appl., 20(1-2):49-76 2009.

[38] M. RAFFEL, C. WILLERT, S. WERELEY, *et al*. - *Particle Image Velocimetry: a Practical Guide*. Springer, 2007.

[39] F. SCARANO - *Tomographic PIV: Principles and Practice*. Measurement Science and Technology, 2013, vol. 24.

[40] D. SCHANZ, S. GESEMANN, A. SCHRÖDER, *et al*. - *Non-Uniform Optical Transfer Functions in Particle Imaging: Calibration and Application to Tomographic Reconstruction*. Measurement Science and Technology, 2012, vol. 24, no 2.

[41] D. SCHANZ, A. SCHRÖDER, S. GESEMANN, *et al*. - *"Shake The Box": A highly Efficient and Accurate Tomographic Particle Tracking Velocimetry (TOMO-PTV) Method Using Prediction of Particle Positions.* PIV13, 10th International Symposium on Particle Image Velocimetry, Delft, The Netherlands, July 1-3, 2013.

[42] V. TODOROFF, G. LE BESNERAIS, D. DONJAT, *et al*. - *Reconstruction of Instantaneous 3D Flow Density Fields by a New Direct Regularized 3DBOS Method*. 17th International Symposium on Applications of Laser Techniques to Fluid Mechanics, Lisbon, 2014.

[43] D. VIOLATO and F. SCARANO - *Three-Dimensional Evolution of Flow Structures in Transitional Circular and Chevron Jets*. Physics of Fluids (1994-present), 2011, vol. 23, no 12, p. 124.

[44] B. WIENEKE - *Iterative Reconstruction of Volumetric Particle Distribution.* Measurement Science and Technology, 2012, vol. 24, no 2, p. 024008

[45] Z. YE, Q. GAO, H. WANG, *et al.* - *Dual-Basis Reconstruction Techniques for Tomographic PIV*. Science China Technological Sciences, 2015, vol. 58, no 11, p. 1963-1970.

[46] R. YEGAVAN, B. LECLAIRE, F. CHAMPAGNAT, *et al*. - *Lucas-Kanade Fluid Trajectories for Time-Resolved PIV*. Accepted in Measurement Science and Technology, 2016.

AUTHORS

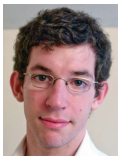
Guy Le Besnerais graduated from ENSTA in 1989 and obtained his PhD degree from *Université Paris Sud* in 1993. Since 1994, he has been working in the Information Processing and Modeling Department at ONERA, the French Aerospace Lab. He obtained the degree of *Habilitation à Diriger les Recherches* (HDR) in 2008 and is affiliated to Paris-Saclay University. Since October 2016, he is a Research Director at ONERA. His research activities include methods for solving inverse problems, performance modeling for imaging measurement systems, and embedded vision for robotics applications.

Frédéric Champagnat graduated from ENSTA in 1989 and obtained his PhD degree from *Université Paris Sud* in 1993. Since 1999 he has been working in the Information Processing and Modeling Department at ONERA, the French Aerospace Lab. Since 2010, he has been a senior scientist and achieved the degree of *Habilitation à Diriger les Recherches* (HDR) in 2011. He is affiliated to Paris-Saclay University and is a member of the board of the first Comité de Pôle of the STIC doctorate school. His research activities include stochastic modeling of signals and images, methods for solving inverse problems, image deblurring and super-resolution and performance modeling for imaging measurement systems. His work is applied in video processing, satellite and airborne imaging, metrology by imaging in fluid mechanics, and the co-design of hybrid sensor+processing systems.

Philippe Cornic graduated from the *École Nationale Supérieure des Ingénieurs Electriciens de Grenoble* in 1989. He has been working on computer vision at ONERA since 1992 and is currently a senior research engineer. His main research interests are geometric computer vision (image to model registration, video and image georeferencing, etc.) and tomographic methods for estimating scalar (BOS) or vector (Particle Image Velocimetry) fields for Fluid Mechanics.

Aurélien Plyer graduated from *Université Pierre et Marie Curie* (Paris 6) in 2008 and received his PhD degree in Image Processing from the *Université de Paris XIII*, in 2013. He is now working in the Information Processing and Modeling Department at ONERA, the French Aerospace Lab. His research deals with high efficiency parallel methods for low-level video processing and 3D environment perception, with application to robotics and experimental physics.

Benjamin Leclaire graduated from the *École Polytechnique*, France, in 2001. He obtained his PhD degree in Fluid Mechanics in 2006. He is now both a researcher at ONERA and an assistant professor at the *École Polytechnique*. His research interests include optical metrology for fluid dynamics, and turbulent flows.



Adam Cheminet graduated from *ENS Cachan*, France, in 2011. He defended his PhD in Fluid Mechanics in May 2016. He is now an ATER at *Université Pierre et Marie Curie*, Paris VI, teaching at undergraduate levels and carrying out research at Institut Jean le Rond d'Alembert-UPMC. His research interests focus on turbulent flows, ranging from optical metrology for fluid dynamics to numerical models for numerical simulation of turbulent flows.

Cédric Illoul joined ONERA in 2012 as a technician for the use and development of optical methods such as PIV, PSP (Pressure Sensitive Paint). He is in charge of assessing new devices, such as using high-power LED for optical methods.

Gilles Losfeld joined ONERA in 1965 as a technician in a supersonic wind tunnel. He then became a member of the ONERA team that developed LDV (Laser Doppel Velocimetry) at ONERA in the early 70's. He has gained a lot of experience and skill in optical methods, such as Schlieren and methods that use laser systems. He is the head of the group within the DAFE department that develops and uses these methods for wind tunnel testing.

Yves Le Sant has been a researcher at ONERA since 1983. His first studies concerned wall interferences and the development of an adaptive test section. He then became involved in applying and developing numerous measurement methods, such as heat flux assessment, Temperature and Pressure Sensitive Paints and the Model Deformation Method. His current activities are in the field of image processing applications, such as Particle Image Velocimetry, and using modern parallel computing architectures such as GPU.

François Nicolas graduated from the *École Nationale Supérieure de Mécanique et d'Aérotechniques* (ENSMA) in Poitiers and received a Master's Degree in mechanical and aerospace engineering in 2013. He achieved his PhD at ONERA Toulouse contributing to the development of 3DBOS and the use of the technique in wind tunnels. He is now in charge of aeroacoustics development for the ONERA large wind tunnel facilities.

Dr. David Donjat graduated from IUSTI in Marseille and received his PhD degree in 2003. From 2003 to 2006 he worked mainly on aerothermal activities, including antiicing systems and internal aerothermal management for Aeroconseil and Airbus. In 2006, he joined ONERA's DMAE department where he currently works on new experimental methods, such as 3DBOS for aerothermal flow characterization and modeling.

After a doctoral degree in Poitiers University, **Francis Micheli** worked with the DAFE's (Fundamental and Experimental Aerodynamics Department) Laser Velocimetry team from 1990 to 2003. He was project leader for LDV measurement campaigns in the Meudon center, with experimental conditions ranging from low velocity to high supersonic. Since 2003, he works in the DMAE (Aerodynamics and Energetics Modeling Department), both with the aeroacoustic and aero-thermal teams. His work mainly concerns the LDV measurements, with contributions in BOS experiments within small or large wind tunnels and more recently about instrumentation on flame test bench.

M. Nafa
(ONERA/Institut UTINAM,
UMR CNRS 6213)

M. Scherman,
A. Bresson, A. Aubin, A. Godard,
B. Attal-Tretout
(ONERA)

P. Joubert
(Institut UTINAM, UMR CNRS 6213)

E-mail: michael.scherman@onera.fr

DOI : 10.12762/2016.AL12-10

Ro-Vibrational Spectroscopy in Hybrid fs/ps-CARS for N_2 Thermometry

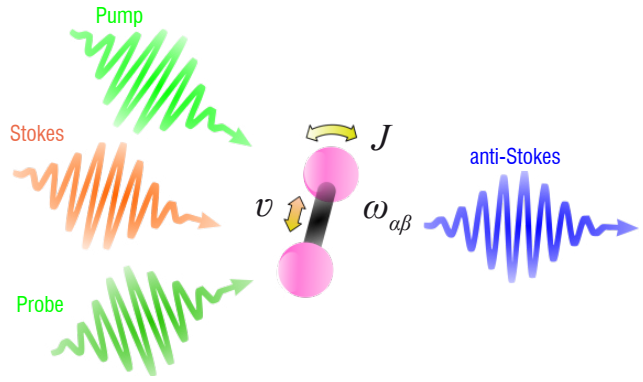
Ro-vibrational CARS spectroscopy is performed on the Q-branch of N_2 in the hybrid fs/ps regime. A new optical setup has been mounted using a single laser source and a spectral filtering stage based on a volume Bragg grating and a Fabry-Perot etalon, in order to achieve a good enough resolution. CARS spectra are recorded in ambient air at a 1 kHz repetition rate. However, a CH_4/air flame is probed at 2 Hz, in order to obtain a strong enough signal to enable thermometry. The theoretical treatment is also detailed and provides the calculated spectrum, properly compared to experimental profiles in order to retrieve the temperature. The accuracy of the hybrid fs/ps-CARS thermometry is satisfactory as regards the state of the art, but further experiments are in progress to increase the repetition rate of the measurement in hot gases.

Introduction, Context

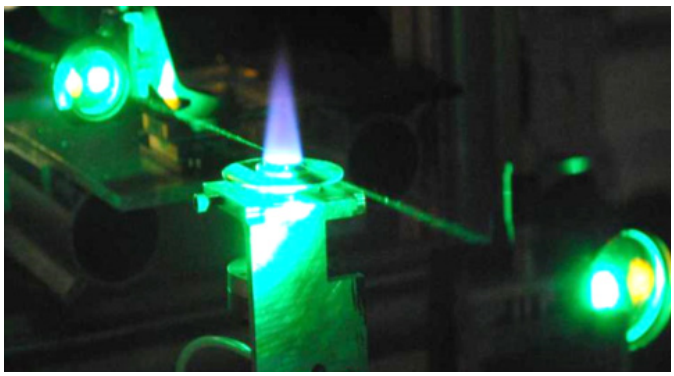
CARS spectroscopy for diagnostics in reactive flows

Laser techniques have high spatial and temporal resolution, enabling locally and temporally resolved measurements of temperature and species concentration. Among the various optical diagnostics, non-linear techniques provide stronger and collimated signals allowing highly luminous media at high temperature and pressure to be studied *in situ* [1]. One of them, coherent anti-Stokes Raman scattering (CARS), was developed at ONERA [2] and extensively used for temperature measurements in hostile reactive environments of interest in the aeronautic field [3] [4].

CARS is a nonlinear spectroscopic technique that was first introduced by Maker *et al.* [5]. As illustrated in Figure 1 (a), a CARS process takes place when three laser beams, namely pump, Stokes and probe, at the respective frequencies ω_P, ω_S and ω_{Pr}, propagate through a medium having Raman active ro-vibrational modes $\omega_{\alpha\beta}$, described by the vibration and rotation quantum numbers υ and J of the ro-vibrational levels $|\alpha\rangle$ and $|\beta\rangle$ associated with the transition. When pump and Stokes fulfill ω_p-ω_S= $\omega_{\alpha\beta}$, a new wave is generated at the anti-Stokes frequency ω_{AS}=ω_{Pr}+$\omega_{\alpha\beta}$. It results from the inelastic scattering of the wave ω_{Pr} by the molecular vibrations, coherently driven by the



(a)



(b)

Figure 1 – (a) Sketch of the CARS process: generation of an anti-Stokes signal by a N_2 molecule, with rotational and vibrational degree of freedom designated by their respective quantum number J and υ , excited by three laser pulses, pump, Stokes and probe. (b) Illustration of CARS diagnostics in a CH_4/air flame.

incident waves at ω_p and ω_S. By scanning the Stokes wavelength, the molecular spectrum can be obtained [6], while a broadband pulse allows it to be recorded in a single laser shot [7].

Over the last three decades, ns and ps-CARS spectroscopy appeared to be very well suited to gas phase diagnostics, such as temperature and concentration measurements in flames, as illustrated in Figure 1 (b). CARS diagnostics were used in this regime at a repetition rate of a few Hz, although 100 kHz CARS thermometry was recently demonstrated with ps-pulses using burst pulse lasers [8]. However, over the last ten years, the development of femtosecond lasers enabled to take advantage of good stability from shot to shot to increase the measurement accuracy, as well as high peak power to enhance the CARS process efficiency and allowing high repetition rate measurements [9]. Also, laser pulses benefit from large bandwidths, which were used to excite and probe wide spectral ranges in a single shot [10] [11] [12], thereby providing the broadband spectroscopy demonstrated previously [7] [13].

However, although gas spectroscopy was straightforward in the nanosecond or picosecond regime, a good enough spectral resolution was not easily achieved in the fs-regime due to the large bandwidth involved in the interaction. Moreover, the stationary approximation is no longer valid when the state lifetime becomes larger than the laser pulse duration and transient solutions of the system evolution thus need to be calculated [14]. Recently, a mixed regime called hybrid fs/ps, has been proposed to take advantage of both the broadband spectral range excited by fs-pulses, and the good resolution provided by ps-pulses [15]. It opens new possibilities for CARS diagnostics, and was successfully used to demonstrate single-shot pure-rotational [16] and under-resolved ro-vibrational [17] thermometry, ultra-broadband kHz punctual and 1D-CARS spectroscopy [18], and finally 10 Hz 2D-thermometry using pure rotational CARS [19]. In that context, our paper is aimed at investigating some specific properties of ro-vibrational CARS spectroscopy in the hybrid fs/ps regime and its potential application to thermometry.

Hybrid fs/ps-CARS spectroscopy for thermometry

The basic principle of temperature measurements by CARS is to probe the relative population of the molecular levels from which the thermodynamic temperature of the molecular system is drawn. The temperature is derived from the observed Q-branch ($\Delta J=0$, $\Delta v=+1$) of the N_2 CARS spectrum, as illustrated in Figure 2 (a), which is fitted to a library of theoretical shapes simulated assuming temperatures, potentially distinct, for vibration and rotation of N_2 molecules [20]. This technique is appropriate for diagnostics in non-equilibrium or very luminous reactive flows, such as plasma and flames, since the CARS process efficiency is high and the Raman shift of N_2 is well separated from those of other chemical species.

Hybrid fs/ps-CARS is described as a pump/probe experiment, and is illustrated in Figure 2. First, the medium is excited by ultra-short pump and Stokes fs-pulses that allow efficient broadband excitation of a set of Raman coherences, depicted in Figure 2 (b). Then, a narrowband ps-probe pulse interrogates the medium during a time that drives the final resolution on the anti-Stokes spectrum, as described in Figure 2 (c). Also, the excitation is separated from the probing of the coherences when using a pump/probe delay τ, which represents a new degree of freedom, compared to ns-CARS, to improve CARS performances, as illustrated in Figure 2 (d). As a result, non-resonant background rejection, which was a main limitation of ns-CARS, is optimized using τ [21] [22], thanks to the long-lived coherences induced from short pulse excitation. The timescale of the process has been well described by time diagrams in [1]. Moreover, as will be described later, in that particular time scale of photon interactions, the final anti-Stokes spectrum is shaped by both the molecular spectral signature and the temporal dephasing between rotational lines. It was used to drive the interferences and improve the diagnostic performances [23], [24], or to perform time-domain spectroscopy [15].

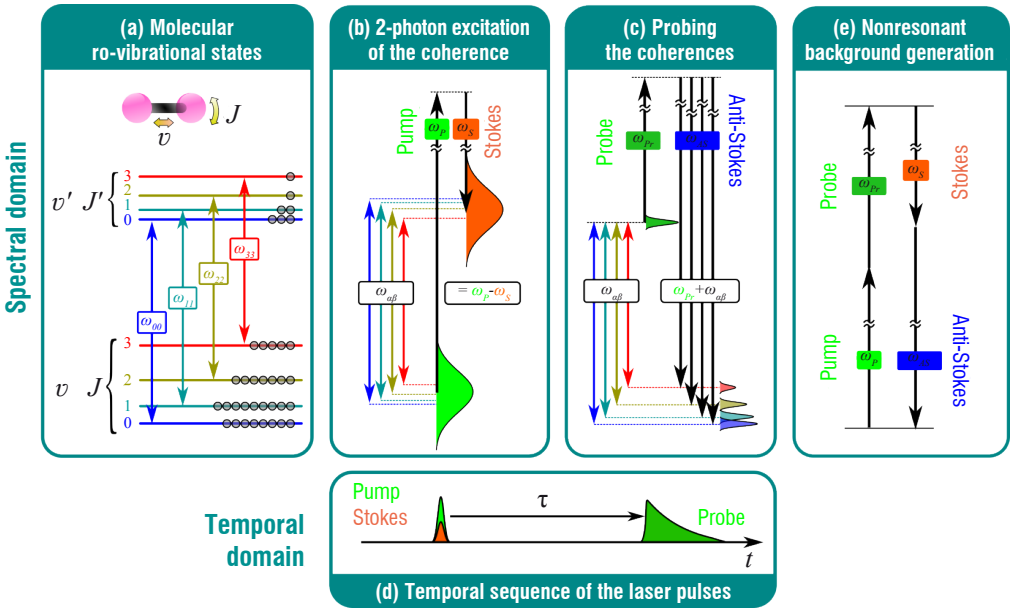


Figure 2 – Sketch of the light-matter interaction in hybrid fs/ps-CARS. (a) Energy level diagram of the N_2 rotational (J) and vibrational (v) structure, and Q-branch associated Raman transitions $\omega_{\alpha\beta}$. (b) Broadband excitation of the molecular coherences using a femtosecond pump and Stokes pulses at the respective frequencies ω_P and ω_S. (c) Narrowband probing using a picosecond probe pulse at frequency ω_{Pr} and generating a well resolved anti-Stokes spectrum at $\omega_{AS}=\omega_{Pr}+\omega_{\alpha\beta}$. (d) Temporal sequence of the laser pulses, in which pump and Stokes overlap and τ is the pump/probe delay. (e) Two-photon absorption (TPA) scheme contributing to CARS signal.

Experimental setup

General description of the experiment

The experimental setup is sketched in Figure 3. The laser beams generation stage, which is illustrated in Figure 3 (a), is based on a single industrial Yb:KGW laser (LightConversion) that delivers 150 fs – 1 mJ pulses at a 1 kHz repetition rate. A commercial optical parametric amplifier (OPA) with a second harmonic converter (LightConversion) is used to generate the tunable Stokes wavelength. It produces a 150 fs broadband (114 cm^{-1}) pulse of 21 μJ energy at the Stokes wavelength of 583.5 nm. A leakage from the OPA pump, at 515 nm, is used to produce both the broadband (130 cm^{-1}) pump pulse of 60 μJ energy, and the picosecond probe pulse. The latter is obtained by spectral filtering using a Fabry-Perot Etalon (FPE) and a volume Bragg grating (VBG, Optigrate). The narrowband pulse (0.7 cm^{-1}) is used as the CARS probe with about 1 μJ energy. Residual spectral parasites are filtered out using high rejection filters (Semrock). The delay between the various pulses is controlled by two motorized translation stages. Recombination of the beams and focalization in a planar folded-BOXCARS geometry [25] are achieved using a dichroic mirror and a 300 mm focal length lens as shown in Figure 3 (b). The anti-Stokes signal is collected using a lens and a collection of interference filters (Semrock). It is then analyzed using a 750 mm spectrometer (Sopra) equipped with a 1800 line/mm grating, and coupled to an electron-multiplying CCD camera (Roper), as shown in Figure 3 (c). The flame was produced out of a McKenna burner, supplied by a CH_4/air mixture at an equivalent ratio $\phi=1$. The flat premixed flame is shown in Figure 1 (b).

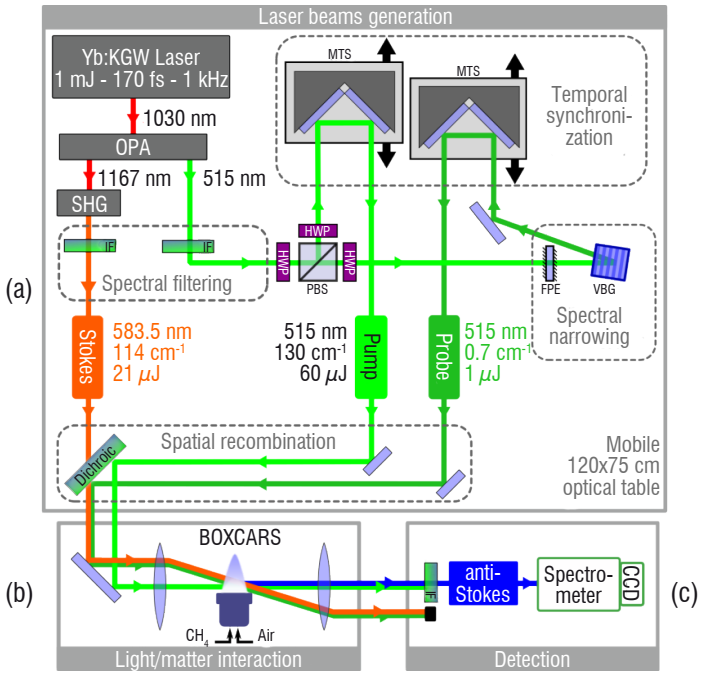


Figure 3 – Experimental setup: (a) laser beam generation, (b) light-matter interaction in the CH_4/air premixed flame with planar BOXCARS geometry, (c) anti-Stokes detection. OPA: Optical Parametric Amplifier, SHG: second harmonic generation, IF: interferential filter, PBS: polarizing beam splitter, HWP: half-wave plate, FPE: Fabry-Perot etalon, VBG: volume Bragg grating, MTS: motorized translation stage.

Ps-probe spectral shaping

The temporal profile of the probe beam acts as a time-window to provide the molecular response. As a result, the probe bandwidth directly drives the resolution of the spectroscopy. Much effort has been devoted to precisely shape this spectral and temporal profile, as detailed in the following.

Various methods have been investigated by the hybrid fs/ps-CARS community, in order to generate the well suited picosecond pulse out of a femtosecond one. They are all aimed at narrowing the pulse bandwidth enough to resolve the fine structure of the spectrum, and to impose the right phase mask to generate an optimal temporal profile for the probe pulse. In particular, spatial light modulators [26] [27] [28], commercial filters [29], Fabry Perot etalons [30] [31], and nonlinear bandwidth compressors [32] were implemented and tested in this double objective. They enabled the generation of pulses with few wavenumbers spectral width, and with asymmetrical temporal shape that proved to be optimal for CARS signal generation [33] [34]. In our case, the bandwidth condition is given by the spacing of the rotational lines, which is less than 1 cm^{-1} in N_2, thus pushing the previously mentioned techniques to a challenging limit.

Monolithic devices are used in our case to shape both the amplitude and phase of the spectral profile. A volume Bragg grating (VBG) was used as a commercially available ultra-narrow amplitude filter [35]. The modulation of the refractive index, encoded in the thick transparent material, reflects a very thin part of the input spectrum that fulfills the Bragg condition. In our setup, a 10 mm thick VBG in photo-thermo-refractive (PTR) glass is used to generate a 0.7 cm^{-1} (FWHM) pulse, as shown in Figure 4 (a).

Phase shaping was achieved using a 100 μm thick FPE that was added upstream from the VBG, to reduce the energy and peak power delivered to the latter. The linewidth of the FPE frequency comb is 1.8 cm^{-1}, thus simplifying the superposition of the VBG reflection peak on the FPE transmission peak. The amplitude profile at the output of the filtering stage is thus fully driven by the VBG filter. However, the phase step that occurs at the FPE resonance peak is encoded into the spectrum of the pulse, and enables the generation of a time-asymmetric probe pulse, as plotted in Figure 4 (b). The temporal profile of the probe pulse was obtained by recording the CARS signal of pure Ar while scanning the pump-probe delay over an adequate range (~100 ps) [16]. In this case, the nonresonant CARS signal is proportional to the instantaneous superposition of the short pump/Stokes excitation (~150 fs FWHM) with the long probe pulse (~30 ps FWHM), thus providing a simple way to sample the temporal intensity profile of the probe (Figure 4(b)).

The Fourier Transform of the temporal profile has been calculated, assuming a flat temporal phase. The intensity (dotted line) and the phase (thick solid line) of the spectrum are plotted in Figure 4 (a). Good agreement between the calculated intensity profile and the experimental spectrum (thin solid line) is observed, validating the flat temporal phase assumption. Moreover, the phase step that characterizes asymmetric pulses generated by FPE could be evaluated. As a result, this set of commercial filters enabled a well suited probe profile to be shaped, *i.e.*, temporally asymmetric, and spectrally narrow enough to probe the rotational lines of N_2, as will be seen later on.

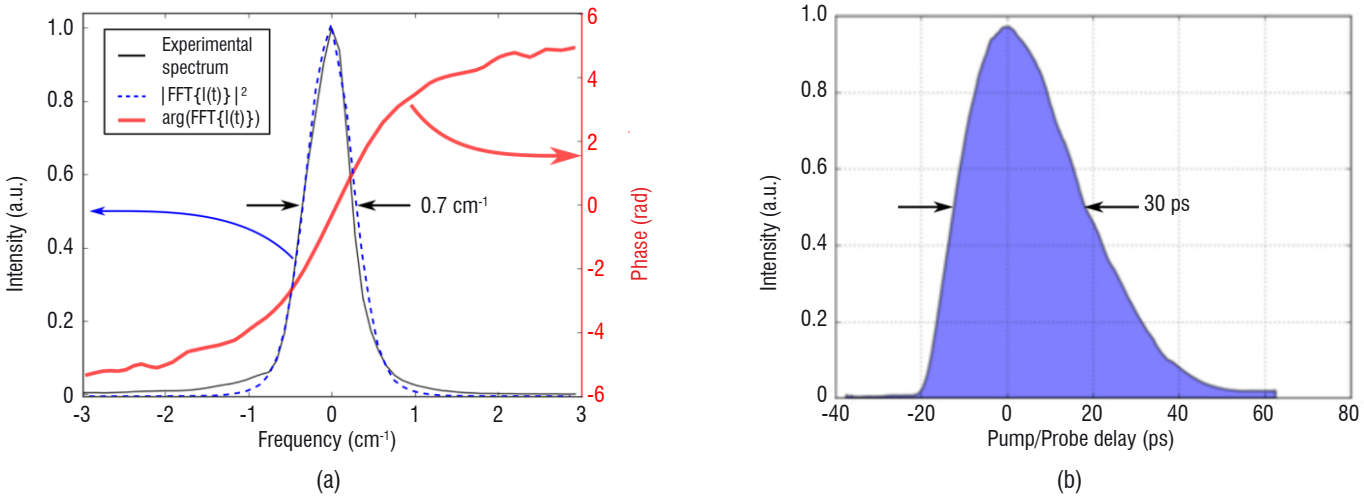


Figure 4 – (a) Spectral and (b) temporal characterization of the generated probe pulse through the filtering stage. (a) The experimental spectrum (thin solid line) is compared to the intensity (dotted line) and phase (thick solid line) of the Fourier transform of the temporal profile.

Characterization of the pump/Stokes excitation

The effective Raman bandwidth excited by the pump/Stokes pulses was measured by recording the nonresonant CARS signal in Ar. The spectral excitation profile (~100 cm^{-1} FWHM) centered on the N_2 first vibrational band is shown in Figure 5. The amplitude distribution of the various ro-vibrational lines of N_2 at flame temperature (~2500 K) is shown in blue for comparison. The excitation bandwidth is thus well suited for N_2 spectroscopy in combustion, since the main intense vibrational bands, $v=0\rightarrow v'=1$, $v=1\rightarrow v'=2$ and $v=2\rightarrow v'=3$, are included in the green excitation window in Figure 5.

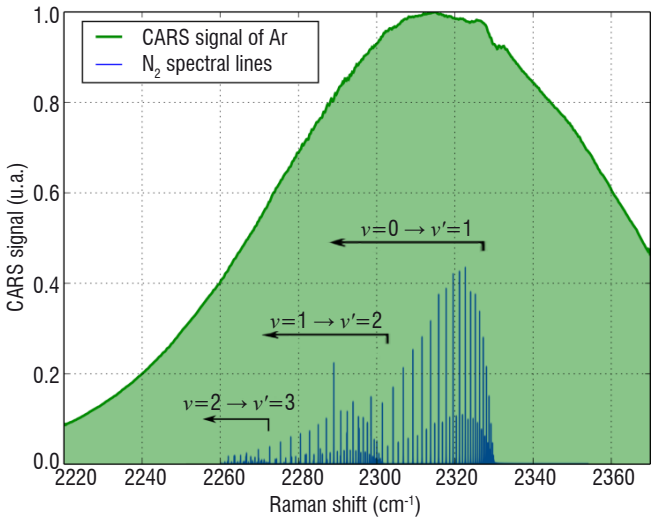


Figure 5 – Nonresonant CARS signal of Ar (green filled curve), illustrating the bandwidth effectively excited by pump and Stokes pulses, compared to the theoretical CARS spectrum of N_2 in a flame (blue lines).

This bandwidth is driven by the pump and Stokes spectral and temporal profiles. Indeed, while Fourier transform limited pump and Stokes pulses allow the excitation of a large spectral range, a chirp on one pulse, which delays the frequencies contained in the optical spectrum, may decrease the bandwidth, which is effectively excited, as illustrated in the spectrograms of Figure 6 (a). In our case, a commercial auto-correlator (Geco, LightConversion) is used to measure the pulse duration. The Stokes pulse proved to be Fourier transform limited, but a chirp was observed on the pump pulse. This aspect is investigated experimentally in Figure 6 (b), by monitoring the

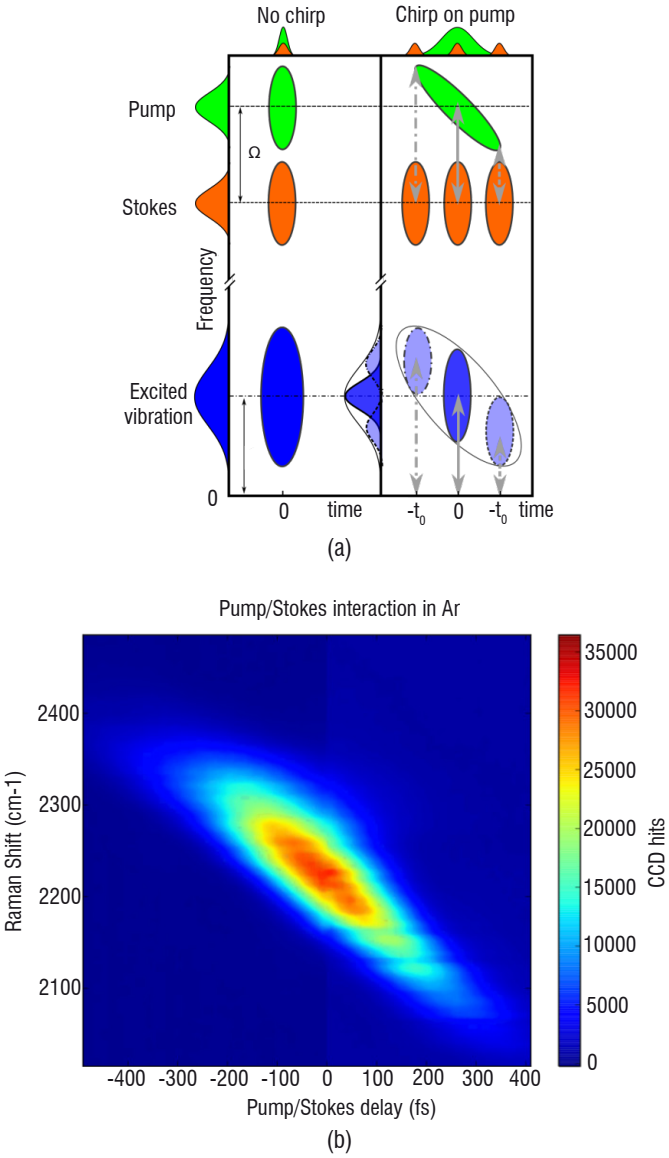


Figure 6 – Effect of a chirp on the CARS spectrum. (a) Illustration of the vibration range excited by a Fourier transform Stokes pulse and a Fourier transform (left), or chirped (right), pump pulse. (b) Experimental CARS spectrum (vertical) collected in Ar versus the pump/Stokes delay (horizontal).

location of the Ar CARS signal when the pump/Stokes delay is scanned. This effect was used in order to finely tune the excitation profile in the selected spectral range, by adjusting the pump/Stokes delay. In addition, it is important to quantify the chirp of the CARS excitation pulses in order to accurately simulate the CARS spectrum. Indeed, it may influence the temporal dephasing of the various excited coherences and thus the shape of the final spectrum, as will be seen later on. Finally, compensating techniques can be implemented to increase the excitation window.

N_2 spectroscopy

Ro-vibrational signature in the hybrid ps/fs regime

Resolving the ro-vibrational band structure

A major species of interest for thermometry is N_2. CARS spectra of N_2 were recorded in ambient air and in a CH_4/air flame, in order to evaluate the performances of the setup for thermometry. Figure 7 shows the experimental spectra recorded (a) in the ns-regime, which represents for us a typical reference spectrum, and (b) in the hybrid fs/ps-regime. An enlargement of the fundamental band of the Q-branch (2300-2335 cm^{-1}) is shown, to outline the rotational features provided by the two techniques. It appears that the two regimes allow both the vibrational and the rotational structures to be seen: three vibrational bands ($v = 0 \rightarrow v' = 1$, $v = 1 \rightarrow v' = 2$ and $v = 2 \rightarrow v' = 3$) can be distinguished in the hot spectra, each one being composed of well-resolved rotational lines, with characteristic 1/4 alternation amplitude of the even/odd lines [36]. It is of major importance in the diagnostics of non-equilibrium media that the two population distributions, corresponding to vibrational and rotational temperatures, are able to be measured independently [20].

As in ns-CARS, the sensitivity of hybrid fs/ps-CARS to temperature can be appreciated by comparing the hot and cold spectra. In ambient air (in blue), the signal is restricted to a 15 cm^{-1} spectral range centered on the first rotational lines of the fundamental band, whereas in the hot spectra (in red) rotational lines are distributed over three vibrational bands. The performance of hybrid fs/ps-CARS ro-vibrational thermometry is thus expected to be comparable to that obtained in ns-CARS, since a similar type of information can be retrieved from the spectra.

Coherent beating of the Raman lines

As mentioned before, one of the specificity of hybrid fs/ps-CARS over traditional ns-CARS is the ability to separate the process into two steps. We consider that the fast excitation of the ro-vibrational coherences through pump and Stokes pulses can be separated from the medium interrogation by the probe pulse in the second step. In the gas phase, introducing a delay τ between these two steps is allowed, since the lifetime of the coherences (~few 100 ps to few ns) is large compared to the duration of the probe pulse (30 ps) [22]. This allows some limitations of ns-CARS to be circumvented, such as nonresonant background interference [21], as well as enabling new measurement strategies to be found [23].

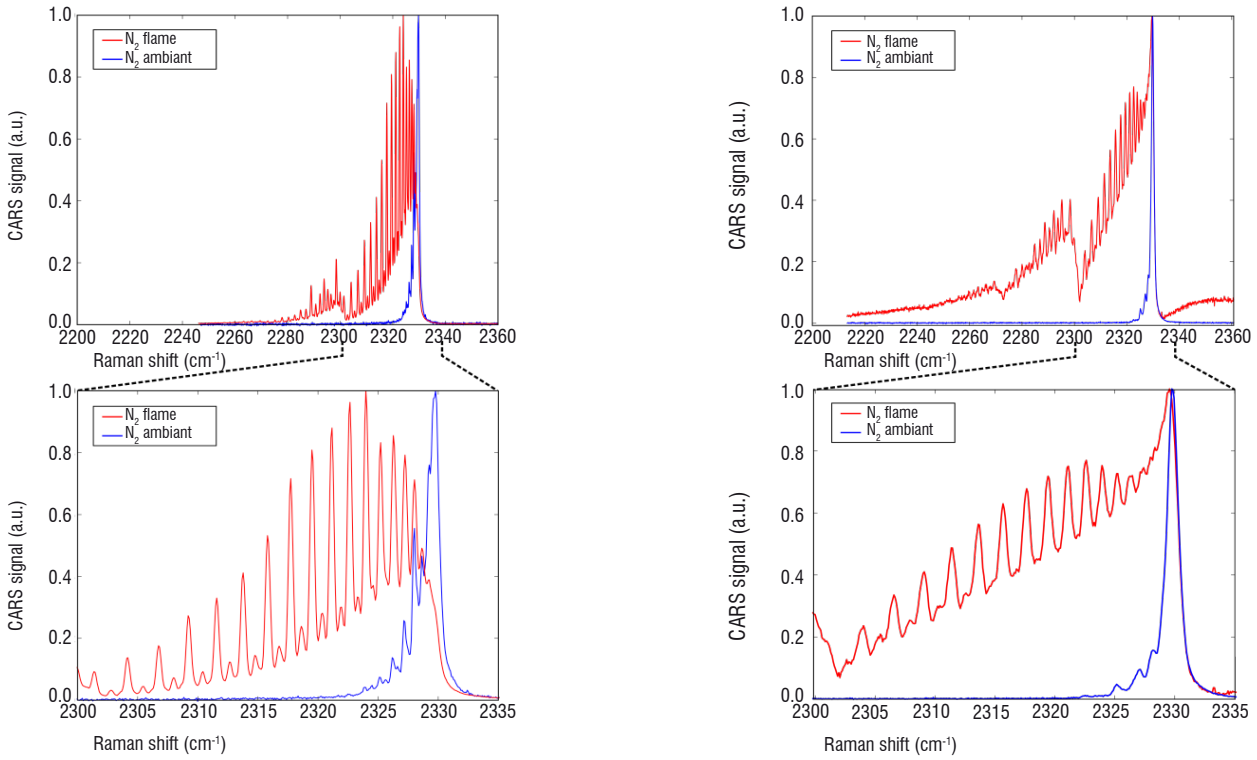


Figure 7 – Comparison of (a) ns-CARS and (b) the hybrid fs/ps-CARS spectra of the N_2 Q-branch obtained in ambient air (blue) and in flame (red). Bottom: enlargement of the fundamental vibrational band.

In Figure 8, the probe pulse delay is continuously scanned from -70 to 270 ps on the vertical axis using motorized delay lines. The CARS spectrum is monitored along the horizontal axis in ambient air, Figure 8 (a), and in the flame, Figure 8 (b). The zero delay position reported in the 2D plot is experimentally checked by maximizing the nonresonant signal. As previously observed in hybrid fs/ps CARS [15] [34] [23] [37] [31], periodic extinctions and revivals of the signal appear at positive delays, more particularly in Figure 8 (b), in which a broader range of ro-vibrational lines is populated by temperature.

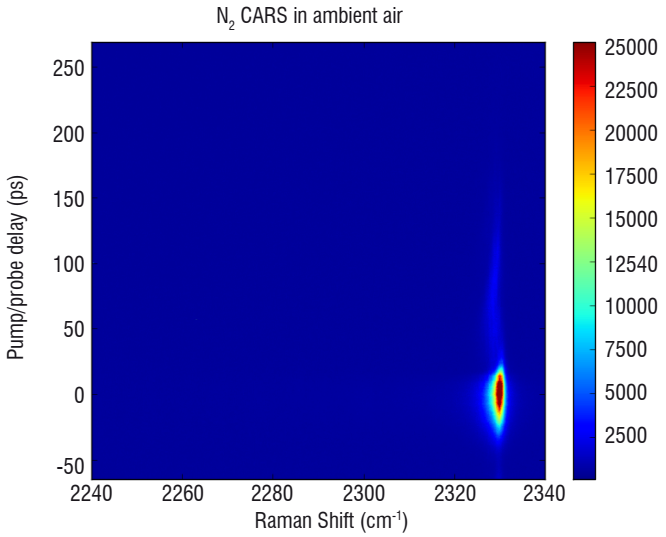


Figure 8 – Experimental investigation of the N_2 CARS spectrum (horizontal axis) evolution versus the pump/probe delay (vertical axis) (a) in ambient air, and (b) in a CH_4/air flame.

In order to follow the modulation of individual Raman lines versus the pump/probe delay, an enlargement of the fundamental band is shown in Figure 9 (a). The evolution of the CARS amplitude at specific Raman shifts, corresponding to several rotational lines ($11<J<37$), is shown in Figure 9 (b). A rising edge is first observed at zero delay, when the probe coincides with the pump/Stokes excitation, which contains resonant and nonresonant contributions to the signal.

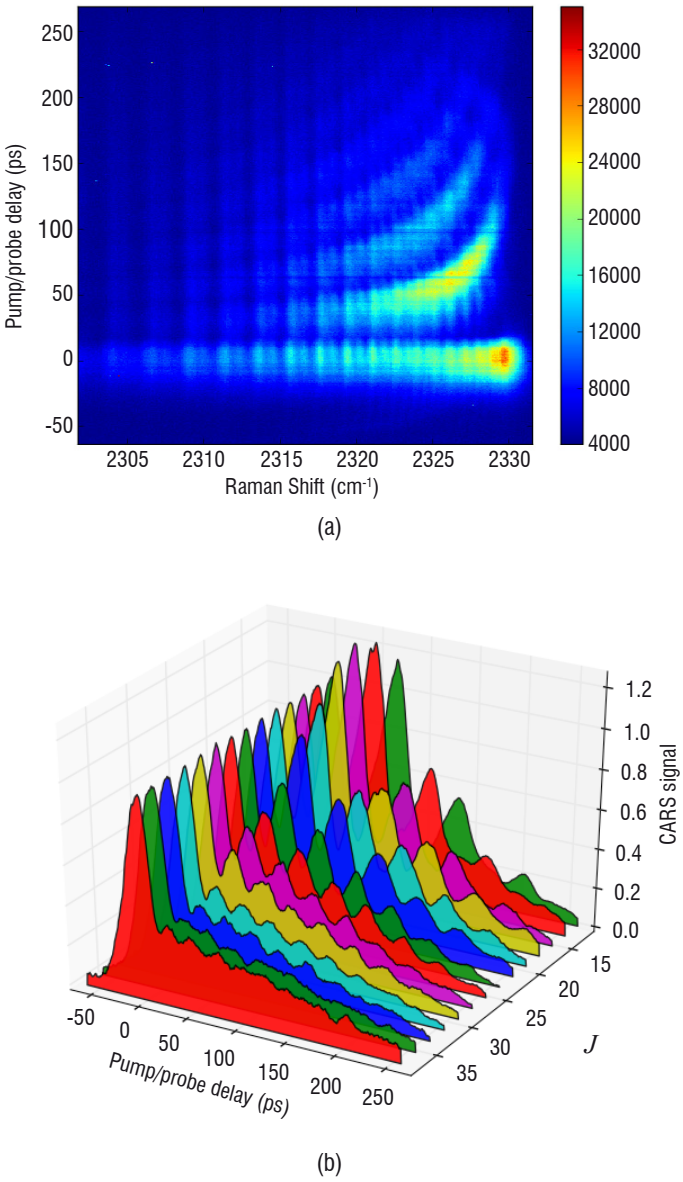


Figure 9 – (a) Evolution of the N_2 CARS spectrum (horizontal axis) centered on the fundamental vibrational band versus the pump/probe delay (vertical axis). (b) Evolution of the CARS intensity of the various rotational lines J of the fundamental vibrational band versus pump/probe delay.

As the probe delay increases, the nonresonant contribution disappears, and the resonant amplitude decays due to the finite lifetime of the coherences. Moreover, the CARS signal exhibits a sinusoidal pattern, whose amplitude and period decrease as a function of the quantum number J.

This behavior is attributed to interferences between rotational lines that occur in the hybrid fs/ps regime, as has already been observed by many authors [12] [37] [34] [23] [31] [38]. Indeed, as illustrated by Figure 10 (a), the multiples ro-vibrational coherences $\omega_{\alpha\beta}$ are excited simultaneously by pump and Stokes pulses. Then, during the free evolution that follows, the phase of each coherence evolves at its own frequency $\omega_{\alpha\beta}$, thus becoming dephased and rephased with respect to each other (Figure 10 (b)). Finally, during the interrogation step operated by the probe pulse, the contributions of the various lines may overlap, depending both on the line separation

and the probe linewidth. Consequently, they interfere constructively or destructively, depending on whether they are in or out of phase, and as a result the anti-Stokes intensity is modulated as plotted in Figure 10 (c).

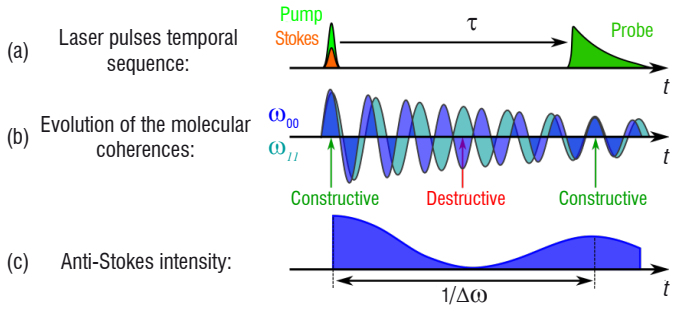


Figure 10 – (a) Typical laser sequence of the hybrid fs/ps-CARS process involving pump, Stokes and probe pulses. (b) Illustration of the dephasing and rephasing that occur between two ro-vibrational coherences, ω_{00} and ω_{11}, when the pump/probe delay τ evolves. (c) Effect of the dephasing on the anti-Stokes intensity when the contributions at ω_{00} and ω_{11} interfere: a modulation is observed with a period of $1/\Delta\omega$, where $\Delta\omega = |\omega_{00} - \omega_{11}|$.

In order to determine the modulation period, the Fourier Transform of the temporal evolution of the even Raman lines is processed in Figure 11 (a). For each J number, a main peak is observed at nonzero wavenumber, which corresponds to the modulation observed in the temporal domain. In Figure 11 (b), the peak spectral position is plotted as a function of the J value. It must be compared to the local spectral interval between rotational lines, which is calculated and also plotted along the vertical axis. It is given at first order in J by the expression:

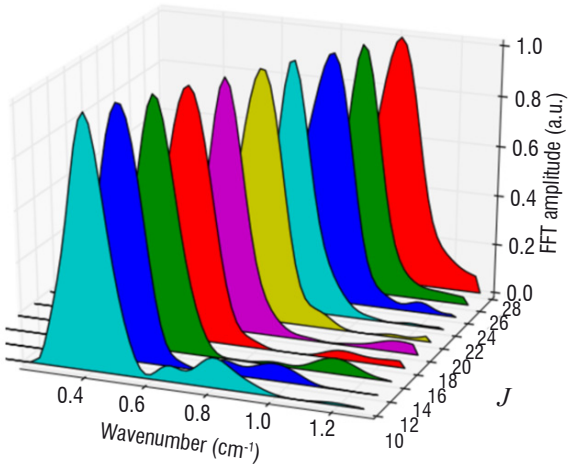
$$\Delta\omega = 2\alpha_e (J+1) \qquad (1)$$

where α_e is the rotation-vibration coupling constant. The good agreement between calculations and experiment confirms the possibility of using the hybrid fs/ps regime to perform spectroscopy with a resolution better than the probe linewidth (0.7 cm^{-1}). Moreover, this interference pattern can be used as a spectral filter, in order to selectively attenuate or enhance specific spectral zones, as proposed by [23] [24].

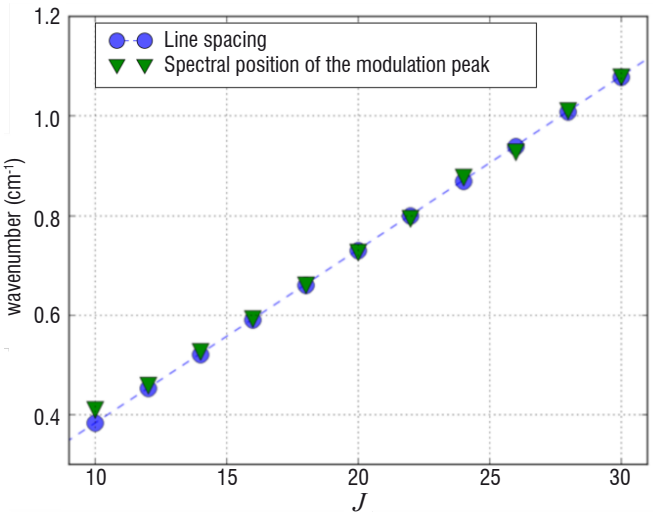
Nonresonant background rejection

In the hybrid fs/ps regime, nonresonant background interference can be reduced by using a pump/probe delay [34], [33] [39] [28], as well as by crossing the polarizations of the pump, Stokes and probe beams, as has previously been done in ns-CARS spectroscopy [40] [41] [1]. CARS spectra in the CH_4/air flame have been recorded for three polarization configurations for the pump, Stokes and probe given in Figure 12 (a), (b) and (c). A Glan-polarizer is added upstream from the spectrometer to analyze the anti-Stokes signal.

The first configuration corresponds to parallel polarizations; the second one corresponds to the conventional technique of a 60° rotated Stokes [1] [42]; in the third one, the probe polarization is orthogonal to those of the pump and Stokes. In the three cases, the orientation of the anti-Stokes analyzer has been rotated every 10°, and CARS spectra have been acquired at each position, with 0° referring to the pump polarization direction. Typical spectra at the maximum and minimum CARS signal amplitude are shown in Figure 12 (a), (b) and (c). Because of a poor signal-to-noise ratio, due to the high degree of polarization of the anti-Stokes beam, the spectrum of Figure 12 (a) is given at 80° while extinction occurs at 90°, and the spectrum of Figure 12 (c) is given at 10° while extinction occurs at 0°. The amplitudes of the resonant (R) and nonresonant (NR) contributions are estimated using the CARS intensity at 2330 cm^{-1} and 2360 cm^{-1} respectively, and are shown in Figure 12 (d), (e) and (f). As expected, R and NR contributions appear respectively in phase and in quadrature for the parallel and the 60° rotated Stokes configuration schemes. This is illustrated in the CARS spectra, which show no distortion for the parallel case (Figure 12 (a)), and whose R and NR relative weight evolves in an opposite way to that of the 60°-Stokes (Figure 12 (b)), when the Glan analyzer is turned. It can be noticed that almost no NR contribution is observed in the 90°-crossed probe scheme and that the spectrum remains unaffected by nonresonant interference when the analyzer is crossed with respect to the probe. Indeed, CARS acquisitions in Ar with the parallel and the 90°-crossed probe configurations revealed that the latter permits the NR amplitude to be reduced by a factor of 8, while the resonant amplitude remained at similar level.



(a)



(b)

Figure 11 – Spectral analysis of the rotational line modulation on the first vibrational band. (a) Spectral modulation peak obtained by Fourier transform of the line CARS intensity temporal evolution as a function of the rotational quantum number J, and (b) comparison of the peak position (triangle) obtained experimentally with the local spectral interval between two successive Raman lines (circle), calculated using spectroscopic constants, as a function of J.

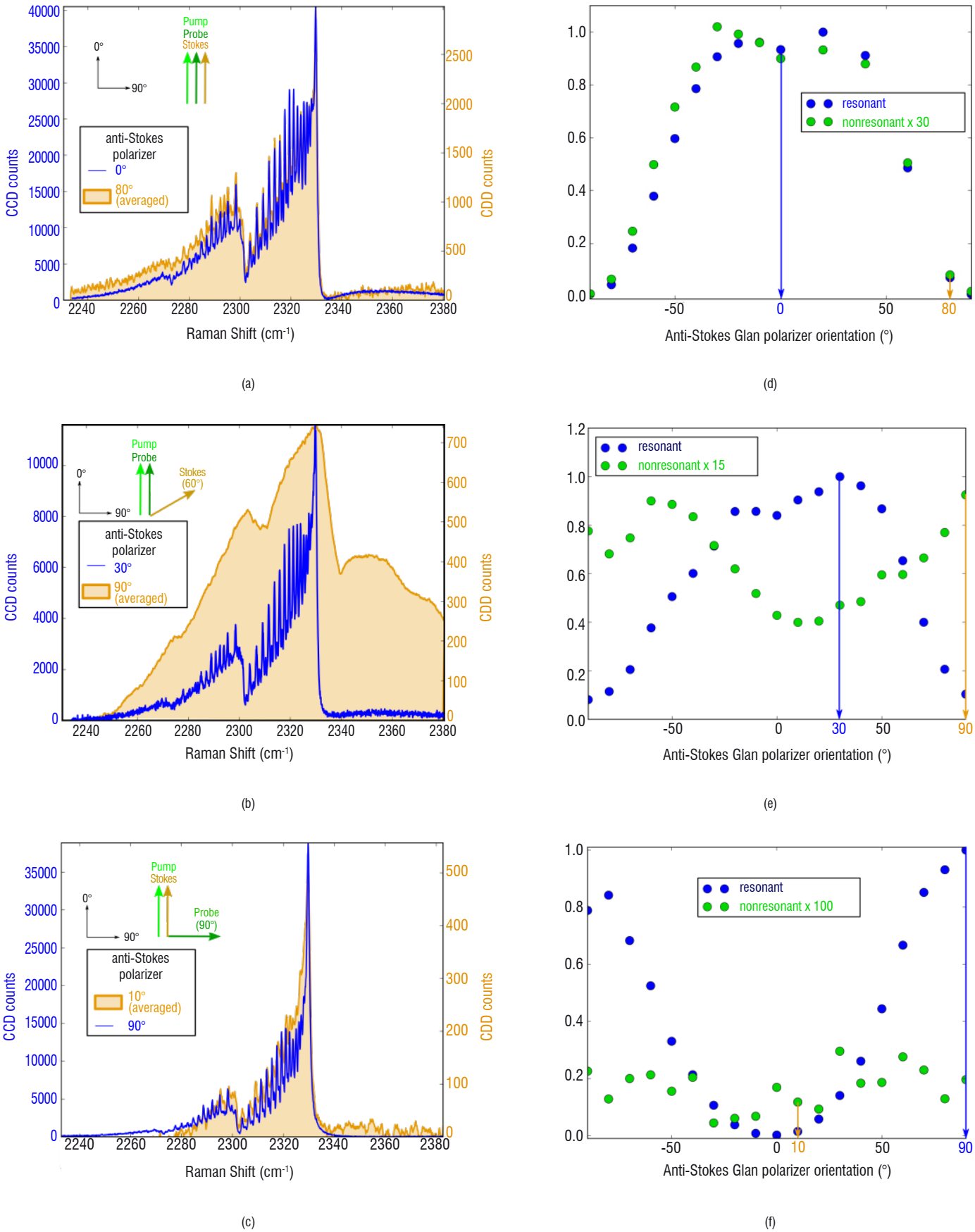


Figure 12 – N_2 CARS spectra at maximum and minimum signal amplitude obtained by rotating Glan-analyzer (a) parallel beams, (b) 60°-rotated Stokes, and (c) 90°-crossed probe. Resonant (blue circles) and nonresonant (green circles) contributions recorded for the 3 configurations are displayed in (d), (e) and (f).

2D plots of the CARS spectra of N_2 in a flame versus probe delay are shown in Figure 13 (a) for parallel polarizations, and in Figure 14 (a) for the crossed-probe configuration. At zero delay, a nonresonant contribution appears at the foot of the resonant line pattern only for parallel polarization. This can be seen in the spectra extracted from the dataset at 0 and 25 ps, shown in Figure 13 (b) and Figure 14 (b), in which the NR amplitude at 2360 cm^{-1} is always zero for the crossed-probe polarization. Moreover, the zero delay spectrum for the crossed polarization shows no distortion by NR contribution, while its shape at 25 ps is very similar to that obtained in the parallel configuration, since the signal is assumed to be purely resonant at this probe delay.

The crossed-probe configuration is observed to drastically attenuate the NR background, while the R signal is maintained. Attenuation is even stronger than expected from the conventional calculations [1], and thus attributed to the short pulse regime properties. Moreover, resonant contributions are maintained in this case, whereas a probe delay may lead to spectral distortions as a result of coherent line interferences. Indeed, polarization configuration, achieved using the 3 beam excitation of the hybrid regime, may provide a strong attenuation of the two photon absorption (TPA) contributions to the CARS signal (Figure 2 (e)), thereby acting in favor of the R/NR ratio optimization. This should increase the measurement accuracy, since the temperature information is mainly enclosed in the resonant part, which is sensitive to the Boltzmann distribution. Moreover, a temperature estimation procedure should gain reliability, since the nonresonant contribution amplitude is usually left as a free parameter of the fit, or estimated from an assumption about the chemical composition of the medium, which introduces a bias in the measurement. Other polarization configurations will be studied in the future, in our laboratory.

Theoretical model and experimental spectra

Description of the model

We use a frequency-domain model equivalent to that reported by Stauffer *et al.* [37]. In the time-domain, the CARS amplitude is obtained by multiplying the probe pulse amplitude by the molecular response, convolved with the pump/Stokes excitation profile. We calculate the CARS amplitude as a function of the Raman shift

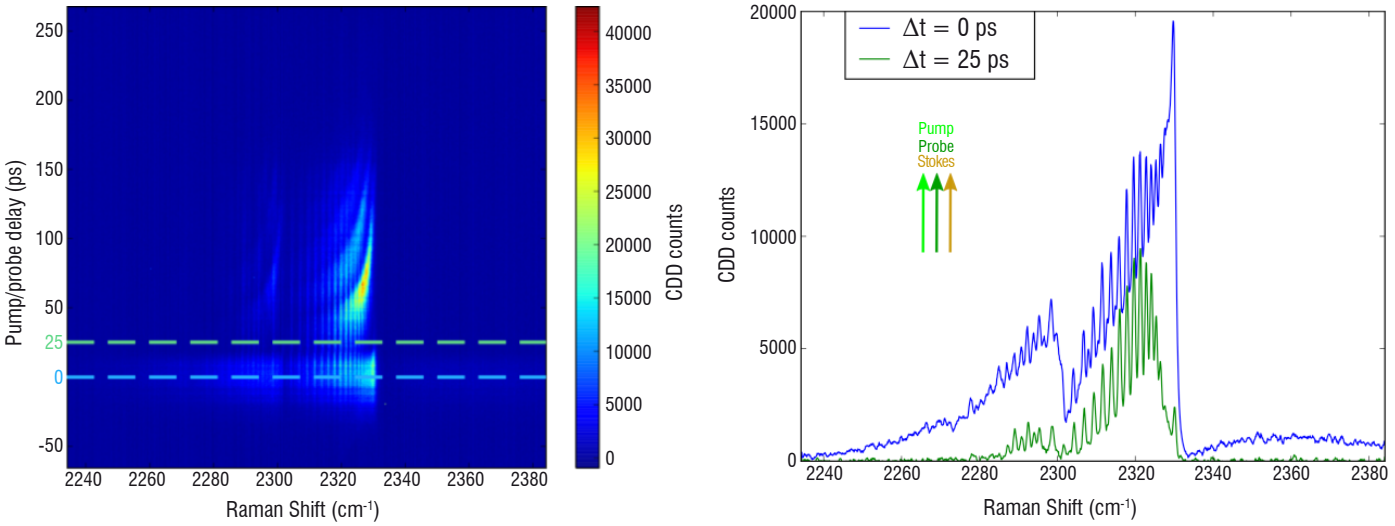


Figure 13 – N_2 CARS spectra recorded with vertical pump, Stokes and probe polarizations. (a) Evolution of the spectrum (horizontal axis) with the probe delay (vertical axis). The two dashed lines indicate the 0 and 25 ps delays. (b) Spectra recorded at the 0 and 25 ps probe delay

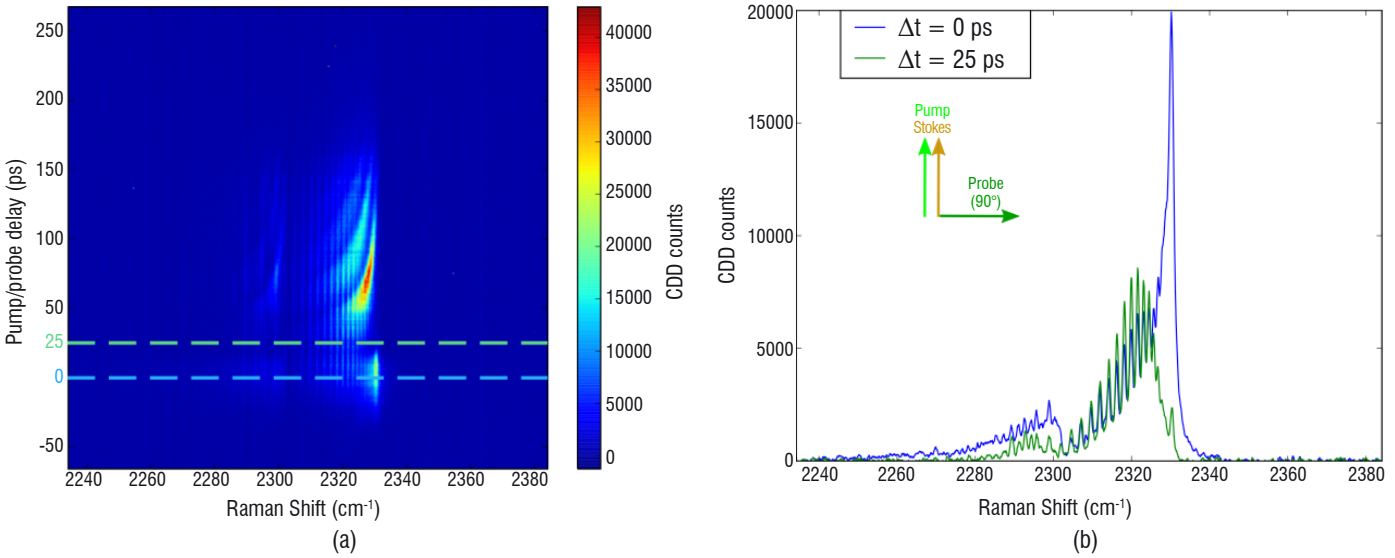


Figure 14 – N_2 CARS spectra recorded with vertical pump and Stokes, and horizontal probe polarizations. (a) Evolution of the spectrum (horizontal axis) with the probe delay (vertical axis). The two dashed lines indicate the 0 and 25 ps delays. (b) Spectra recorded at the 0 and 25 ps probe delay.

$\omega = \omega_{pump} - \omega_{Stokes}$, by convolving two terms: the molecular excitation, $A_{excitation}(\omega,\tau)$, which takes into account the pump/probe delay τ, and the probe spectrum $A_{probe}(\omega)$. The CARS intensity is thus given by:

$$I_{CARS}(\omega,\tau) = \left| A_{excitation}(\omega,\tau) \otimes A_{probe}(\omega) \right|^2 \quad (2)$$

Modeling of the molecular excitation

The molecular excitation $A_{excitation}(\omega,\tau)$ is expressed as the nonlinear susceptibility $\chi^{(3)}(\omega)$, weighted by the effective spectral profile $H(\omega)$ of the pump/Stokes excitation, and dephased by a term $\phi(\omega,\tau)$ that takes into account the free evolution of the coherences during the pump/probe delay τ:

$$A_{excitation}(\omega,\tau) = H(\omega)\chi^{(3)}(\omega)\phi(\omega,\tau) \quad (3)$$

The third order susceptibility $\chi^{(3)}(\omega)$ is composed of resonant and nonresonant contributions. The resonant amplitude depends on the pump, Stokes and probe polarizations, on the temperature, and the number density [1] [43]. Raman cross-sections, together with Doppler and collisional line broadening, are needed to properly simulate the profile. The resonant CARS susceptibility $\chi_R^{(3)}(\omega)$ is a sum of the nearby Raman lines expressed by [43]:

$$\chi_R^{(3)}(\omega) = \sum_{\alpha\beta} \left(\frac{d\sigma_{\alpha\beta}}{d\Omega} \right) . \Delta\rho_{\alpha\beta} . \frac{i\sqrt{\pi}}{\sqrt{2}\sigma_{Doppler}} . W\left(\frac{\omega_{\alpha\beta} - \omega - i\Gamma_{\alpha\beta}}{\sqrt{2}.\sigma_{Doppler}} \right), \quad (4)$$

where:

- $|\alpha\rangle$ and $|\beta\rangle$ are two ro-vibrational levels coupled by a Raman allowed transition,
- $\left(\frac{d\sigma_{\alpha\beta}}{d\Omega} \right)$ is the spontaneous Raman cross-section of the $|\alpha\rangle \rightarrow |\beta\rangle$ transition,
- $\omega_{\alpha\beta}$ is the Raman shift of the $|\alpha\rangle \rightarrow |\beta\rangle$ transition,
- $\Delta\rho_{\alpha\beta}$ is the population difference between the $|\alpha\rangle$ and $|\beta\rangle$ levels, which is a function of temperature (Boltzmann distribution),
- $\Gamma_{\alpha\beta}$ stands for the collisional broadening of the $|\alpha\rangle \rightarrow |\beta\rangle$ transition, the linewidths $\Gamma_{\alpha\beta}$ were taken from a fit of N_2 lines with an energy gap law standing for collisional effects including motional narrowing [44],
- $\sigma_{Doppler}$ is the Doppler broadening,
- $W(a+ib) = \frac{i}{\pi} \int_{-\infty}^{+\infty} \frac{e^{-t^2}}{a-t+ib} dt$ is the complex error function, giving the spectral profile of each transition [43] considered as a complex Voigt profile resulting from Doppler and collisional effects.

Apart from the resonant CARS signal, a nonresonant background is generated, which is modeled as a constant scalar $\chi_{NR}^{(3)}$, potentially complex [37]. The whole susceptibility $\chi^{(3)}(\omega)$ is thus obtained by:

$$\chi^{(3)}(\omega) = \chi_R^{(3)}(\omega) + \chi_{NR}^{(3)} \quad (5)$$

The excitation amplitude profile $H(\omega)$ is introduced to take into account the finite excitation bandwidth of the pump and Stokes pulses. It has been modeled by a real profile, since the dephasing between the different spectral parts, which can be nonzero due to the chirp of the pump pulse, is assumed to be negligible in our case. Indeed, the 2D plot in Figure 8 shows that all Raman lines are located at the same delay $\tau = 0$ along the abscissa, thus proving that coherences are not delayed significantly with respect to each other. Indeed, the excitation time window is limited to ~150 fs by the short duration of the Stokes pulse. In comparison, the smallest temporal modulation period observed in the Raman lines, estimated to be of about 20 ps in Figure 8, is thus two orders of magnitude larger than the excitation window.

In Figure 15, two different experimental conditions are shown. The pump/Stokes excitation profiles are shifted by 40 cm^{-1} to outline the effect on the CARS spectrum. Experimentally, excitation profiles are obtained by recording the CARS signal in Ar and by adjusting the temporal shift of the pump, thus providing $|H(\omega)|^2$.

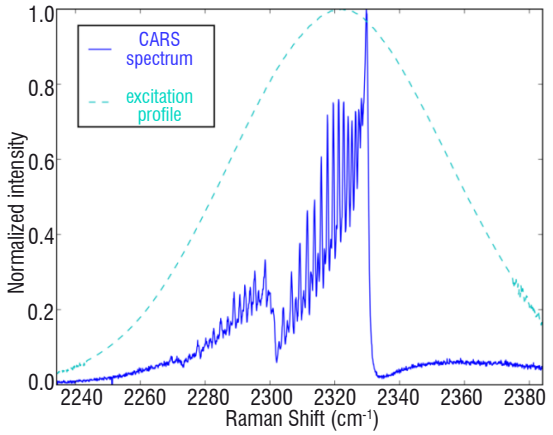


Figure 15 – CARS experimental spectra (blue) acquired over 5000 shots in a premixed CH_4/air flame, with pump/Stokes excitation profiles $|H(\omega)|^2$ (dashed line) centered at (a) 2320 cm^{-1} and (b) 2280 cm^{-1}.

The dephasing term $\phi(\omega,\tau)$ is used to model the free evolution of the coherences during the time delay τ separating the excitation and the probing. It is expressed by:

$$\phi(\omega,\tau) = e^{i\omega\tau} \quad (6)$$

Finally, the complete molecular excitation is expressed by:

$$A_{excitation}(\omega,\tau) = H(\omega)\chi(\omega)e^{i\omega\tau} \quad (7)$$

Modeling of the probing step

During this step, the probe spectral pattern $A_{probe}(\omega)$ is convoluted to the molecular excitation $A_{excitation}(\omega,\tau)$, in order to produce the CARS spectrum. The final resolution is thus driven by the spectral width of the probe. In the temporal domain simulation, the crucial parameter is shown to be the pulse duration. The latter corresponds to the temporal window in which the molecular response is probed [37]. Finally, the CARS intensity can be expressed as follows:

$$I_{CARS}(\omega) = \left| H(\omega)\chi(\omega)e^{i\omega\tau} \otimes A_{probe}(\omega) \right|^2 \quad (8)$$

Since the probe spectral pattern is much narrower than the excitation profile, the factor $H(\omega)$ can be extracted from the convolution product and the calculation is implemented as follows:

$$I_{CARS}(\omega) = \left| H(\omega) \right|^2 \cdot \left| \chi(\omega)e^{i\omega\tau} \otimes A_{probe}(\omega) \right|^2 \quad (9)$$

The probe spectrum $A_{probe}(\omega)$ is measured (Figure 4 (a), thin solid line) and adjusted to fit a Voigt profile. However, a proper approach may be to consider the full complex amplitude of the probe spectrum, which can be obtained by recording the spectral shape with a spectrometer, and/or a combination of interferometric techniques such as FROG [45]. The excitation profile $|H(\omega)|^2$ is obtained by measuring the nonresonant response in Ar, as detailed previously. An overall view of the algorithm, including the convolution by apparatus function of the spectrometer, is summed up in Figure 16.

Temperature measurements

The temperature is measured by comparing the experimental spectrum to a set of simulated spectra. A least-square algorithm, run over a bank of spectra, simulated by varying the temperature in 10 K steps and normalized to the unit maximum, leads to the best fit. The free parameter of the fit is thus the temperature, whereas the pump/probe delay and the nonresonant background amplitude are fixed. The least-square minimization range is chosen between 2250 cm^{-1} (lowest line of v=2 band) and 2350 cm^{-1}, which provides information on both the R and NR contributions, as illustrated in Figure 5. In particular, it enables the choice of the fixed NR amplitude to be validated *a posteriori*.

The accuracy of the thermometric setup has been tested on two sets of experimental data at zero probe delay. The first set is a stack of 1000 single shots spectra, acquired at a repetition rate of 1 kHz in ambient air. The second one is composed of 100 spectra recorded in a premixed-CH_4/air flame with equivalent ratio $\phi = 1.0$ at atmospheric pressure. Each hot spectra is averaged over 500 shots, in order to reach a signal-to-noise ratio (SNR) beyond 120. The criticality of the SNR in hybrid fs/ps-CARS has been recently investigated by Kearney [46]. The probe delay parameter has been respectively fixed to 5.5 ps and 1.5 ps for the first and the second set, due to a small variation in the optical paths during the laser beam alignment.

Figure 17 and Figure 18 show a typical experimental spectrum (blue) superposed to its best fit (green). Temperature histograms, represented in Fig. 17 (b) and Fig. 18 (b), have been processed and were used to quantify the mean temperature $\langle T \rangle$, standard deviation σ_T, and relative precision defined by the $\sigma_T/\langle T \rangle$ ratio, which quantifies

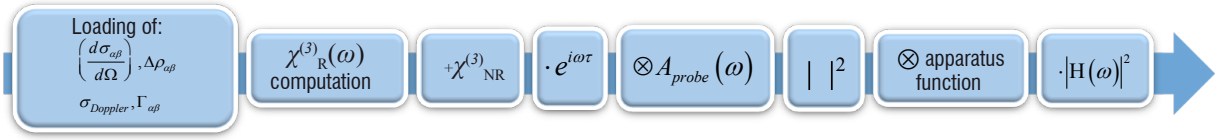


Figure 16 – Schematic view of the hybrid fs/ps-CARS simulation algorithm

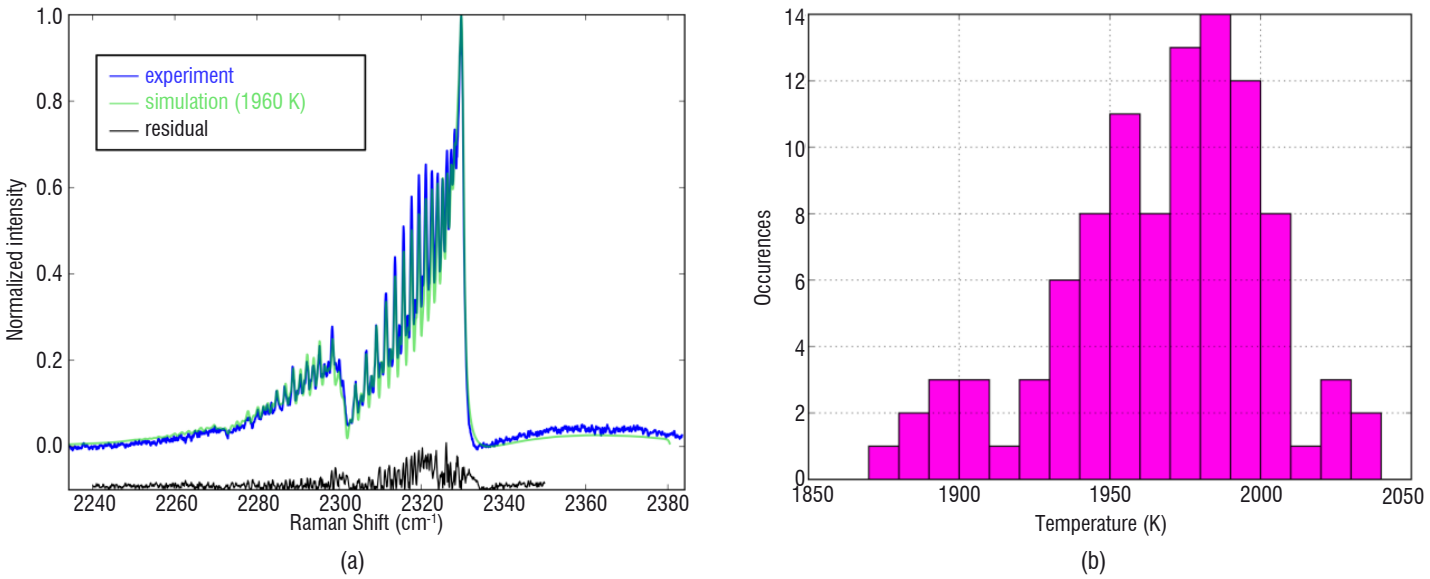


Figure 17 – CARS thermometry in a CH_4/air flame. (a) Typical experimental spectrum (blue), average best fit (green), and residual contribution (black) downshifted by 0.1 along the vertical axis. (b) Temperature histogram: $\langle T \rangle$ = 1960 K, σ_T = 35 K

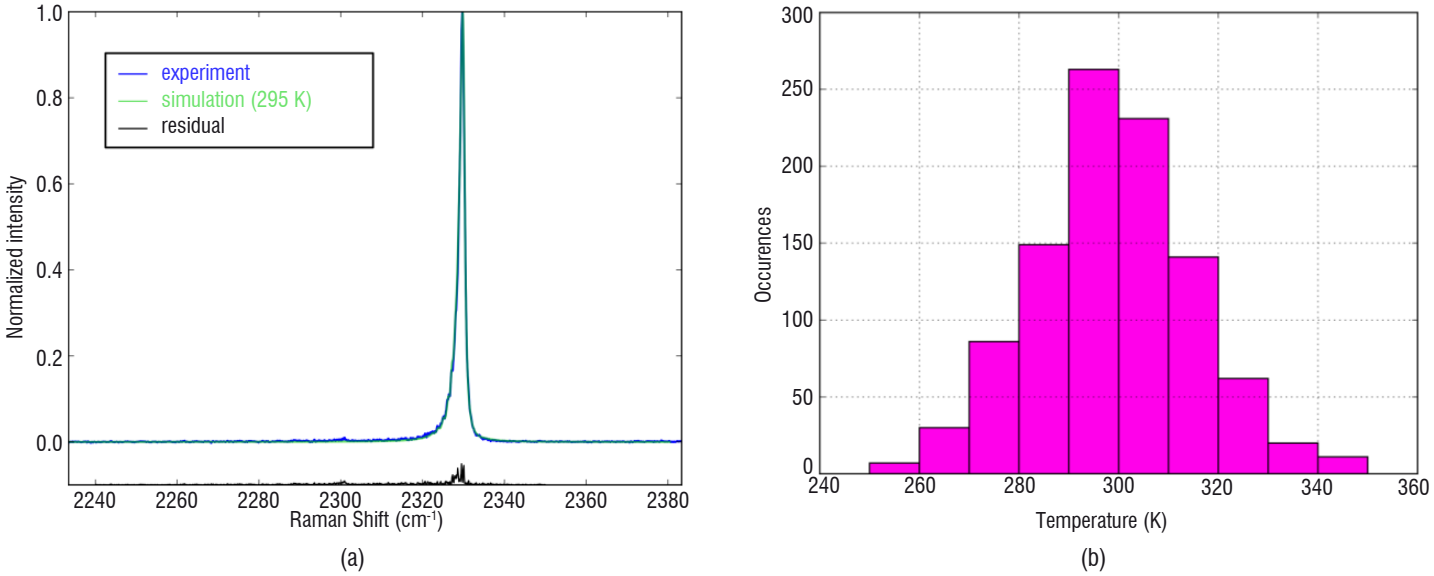


Figure 18 – CARS thermometry in ambient air. (a) Typical experimental spectrum (blue), average best fit (green), and residual contribution (black) downshifted by 0.1 along the vertical axis. (b) Temperature histogram: $\langle T \rangle = 295$ K, $\sigma_T = 15$ K

the repeatability of the measurement. In ambient air and in the flame, the estimated mean temperatures were respectively 295 and 1960 K, with corresponding precisions of 5.2% (15 K) and 1.7% (35 K). The temperature in the flame was measured to be slightly lower than the adiabatic CH_4/air flame temperature (~2200 K). This was attributed to experimental misalignment of the probed volume in the flame, and to nonzero pump/probe delay adjustment. This measurement will be reproduced in a better controlled flame and using a well-calibrated procedure to accurately and rapidly control the pump/probe delay. It is comparable to the few percent precision obtained with hybrid fs/ps rotational CARS [46], fs-CPP-CARS [47], and ns-CARS [48] in a flame. To go further in the characterization of the techniques, single shot thermometry at low and high temperatures would be of great significance, in order to avoid integrating spectra over the flame temperature fluctuations. This new process in our setup development would require a more energetic probe, whose small bandwidth is currently achieved at the expense of an important loss of energy, due to the filtering technique used. Using nonlinear techniques [32], so as to achieve more efficient narrowing, would be one way to achieve this goal.

Perspectives and conclusion

In summary, hybrid fs/ps-CARS was used to properly resolve ro-vibrational structures of the Q-branch of N_2. The resolution was good enough to observe both rotational and vibrational features in a single spectrum, as had already been done in ns-CARS, thus opening the way to instantaneous rotational and vibrational temperature measurements. Moreover, time domain spectroscopy was used, in order to resolve sub-probe bandwidth details in the spectrum.

Temperature measurements were performed in the flame, by comparing simulations to experiments, with good accuracy. The model will be improved by taking into account non-diagonal elements of the relaxation matrix, i.e., motional narrowing effects. The calculation of motional narrowing in hybrid fs/ps-CARS is under study, in order to achieve our final objective, which is to probe both high pressure and high temperature media. The long probe duration that is required for spectral resolution will be further combined to an increase in the energy of the probe, to achieve single shot measurements ■

Acknowledgments:

Frederic Grisch is greatly thanked for his constant support and advice regarding both the experiments and the simulations. We also thank Hans Stauffer and Sean Kearney, who helped us to clarify important theoretical aspects through fruitful discussions.
We are grateful to Jean Pierre Faleni and Thomas Schmid for their technical assistance with the burner design and gas supply.
Funding was provided by the SESAME program of the *Conseil Régional d'Ile de France* (13016393), thanks to which part of the equipment was obtained.

References

[1] S.A.J. DRUET and J.-P.E. TARAN - *CARS Spectroscopy*. Prog.Quant. Electronics, vol. 7, p. 1, 1981.

[2] S.A.J. DRUET, B. ATTAL, T.K. GUSTAFSON and J.P. TARAN - *Electronic Resonance Enhancement of Coherent Anti-Stokes Raman Scattering.* Physical Review A, vol. 18, no. 4, p. 1529, 1978.

[3] BLOEMBERGEN - *Nonlinear Optics*. 31977th ed.: WA Benjamin New York, 1965.

[4] A.C. ECKBRETH - *Laser Diagnostics for Combustion Temperature and Species.* 2nd ed. Amsterdam: Gordon and Breach Publishers, 1996.

[5] P.D. MAKER and R.W. TERHUNE - *Study of Optical Effects Due to an Induced Polarization Third Order in the Electric Field Strength.* Phys. Rev, vol. 137, no. 3A, p. A801, 1965.

[6] P.R. REGNIER and J.P.-E. TARAN - *On the Possibility of Measuring Gas Concentrations by Stimulated Anti-Stokes Scattering.* Applied Physics Letters, vol. 23, p. 240, 1973.

[7] W.B. ROH, P.W. SCHREIBER and J.P.-E. TARAN - *Single-Pulse Coherent Anti-Stokes Raman Scattering.* Single-pulse Coherent anti-Stokes Raman Scattering, vol. 29, no. 3, p. 174, 1976.

[8] S. ROY, P.S. HSU, N. JIANG, M.N. SLIPCHENKO and J.R. GORD - *100-kHz-Rate Gas-Phase Thermometry Using 100-Ps Pulse from a Burst-Mode Laser.* Optics Letters, vol. 40, no. 21, p. 5125, 2015.

[9] S. ROY, W.D. KULATILAKA, D.R. RICHARDSON, R.P. LUCHT and J.R. GORD - *Gas-Phase Single-Shot Thermometry at 1 kHz Using Fs-CARS Spectroscopy.* Optics Letters, vol. 34, no. 24, p. 2857, 2009.

[10] C.C. HAYDEN and D.W. CHANDLER - *Femtosecond Time-Resolved Studies of Coherent Vibrational Raman Scattering in Large Gas-Phase Molecules..* The Journal of Chemical Physics, vol. 103, no. 24, p. 10465, 1995.

[11] P. BEAUD, H.-M. FREY, T. LANG and M. MOTZKUS - *Flame Thermometry by Femtosecond CARS*. Chemical Physics Letters, vol. 344, p. 407, 2001.

[12] B. LAVOREL, H. TRAN, E. HERTZ, O. FAUCHER, P. JOUBERT, M. MOTZKUS, T. BUCKUP, T. LANG, H. SKENDEROVI, G. KNOPP, P. BEAUD and H.M. FREY - *Femtosecond Raman Time-Resolved Molecular Spectroscopy.* Comptes Rendus Physique, vol. 5, p. 215, 2004.

[13] S. ROY, T. MEYER and J.R. GORD - *Broadband Coherent Anti-Stokes Raman Scattering Spectroscopy of Nitrogen Using a Picosecond Modeless Dye Laser*. Optics Letters, vol. 30, no. 23, p. 3222, 2005.

[14] W. ZINTH - *Transcient Coherent Raman Scattering in the Time and Frequency Domain.* Optics Communications, vol. 34, no. 3, p. 479, 1980.

[15] B.D. PRINCE, A. CHAKRABORTY, B.M. PRINCE, and H.U. STAUFFER - *Development of Simultaneous Frequency- and Time-Resolved Coherent Anti-Stokes Raman Scattering for Ultrafast Detection of Molecular Raman Spectra.* The Journal of Chemical Physics, vol. 125, p. 044502, 2006.

[16] J.D. MILLER, S. ROY, M.N. SLIPCHENKO, J.R. GORD, and T.R. MEYER - *Single-Shot Gas-Phase Thermometry Using Pure-Rotational Hybrid Femtosecond-Picosecond Coherent Anti-Stokes Raman Scattering.* Optics Express, vol. 19, no. 16, p. 15627, 2011.

[17] J.D. MILLER, M.N. SLIPCHENKO, T.R. MEYER, H.U. STAUFFER, and J.R. GORD - *Hybrid Femtosecond/Picosecond Coherent Anti-Stokes Raman Scattering for High-Speed Gas-Phase Thermometry.* Optics Express, vol. 35, no. 14, p. 2430, 2010.

[18] A. BOHLIN and C.J. KLIEWER - *Two-Beam Ultrabroadband Coherent Anti-Stokes Raman Spectroscopy for High Resolution Gas-Phase Multiplex Imaging.* Applied Physics Letters, vol. 104, p. 031107, 2014.

[19] A. BOHLIN and C.J. KLIEWER - *Communication: Two-Dimensional Gas-Phase Coherent Anti-Stokes Raman Spectroscopy (2D-CARS): Simultaneous Planar Imaging and Multiplex Spectroscopy in a Single Laser Shot.* The Journal of Chemical Physics, vol. 138, p. 221101, 2013.

[20] D. MESSINA, B. ATTAL-TRETOUT and F. GRISCH - *Study of a Non-Equilibrium Pulsed Nanosecond Discharge at Atmospheric Pressure Using Coherent Anti-Stokes Raman Scattering*. Proceedings of the Combustion Institute, vol. 31, no. 1, p. 825, 2007.

[21] W. ZINTH, A. LAUBEREAU and W. KAISER - *Time Resolved Observation of Resonant and Non-Resonant Contributions to the Nonlinear Susceptibility.* Optics Communications, vol. 26, no. 3, p. 457, 1978.

[22] F.M. KAMGA and M.G. SCEATS - *Pulse-Sequenced Coherent Anti-Stokes Raman Scattering Spectroscopy: A Method for Suppression of the Nonresonant Background.* Optics Letters, vol. 5, no. 3, p. 126, 1980.

[23] J.D. MILLER, C.E. DEDIC and T.E. MEYER - *Vibrational Femtosecond/Picosecond Coherent Anti-Stokes Raman Scattering with Enhanced Temperature Sensitivity for Flame Thermometry from 300 to 2400 K.* Journal of Raman Spectroscopy, vol. 46, p. 702, 2015.

[24] M. SCHERMAN, M. NAFA, T. SCHMID, A. GODARD, A. BRESSON, B. ATTAL-TRETOUT and P. JOUBERT - *Rovibrational Hybrid fs/ps CARS Using a Volume Bragg Grating for N2 Thermometry.* Optics Letters, vol. 41, no. 3, p. 488, 2016.

[25] A.C. ECKBRETH - *BOXCARS: Crossed-Beam Phase-Matched CARS Generation in Gases.* Applied Physics letters, vol. 32, p. 421, 1978.

[26] D. ORON, N. DUDOVICH, D. YELIN AND Y. SILBERBERG - *Narrow-Band Coherent Anti-Stokes Raman Signals from Broad-Band Pulses.* Physical Review Letters, vol. 88, no. 6, p. 063004, 2002.

[27] S. ROY, P. WRZESINSKI, D. PESTOV, T. GUNARATNE, M. DANTUS and J.R. GORD - *Single-Beam Coherent Anti-Stokes Raman Scattering Spectroscopy of N_2 Using a Shaped 7 Fs Laser Pulse.* Applied Physics Letters, vol. 95, p. 074102, 2009.

[28] S.O. KONOROV, M.W. BLADES and R.F.B. TURNER - *Lorentzian Amplitude and Phase Pulse Shaping for Nonresonant Background Suppression and Enhanced Spectral Resolution in Coherent Anti-Stokes Raman Scattering Spectroscopy and Microscopy.* Applied Spectroscopy, vol. 64, no. 7, p. 767, 2010.

[29] O. KATZ, J.M. LEVITT, E. GRINVALD and Y. SILBERBERG - *Single-Beam Coherent Raman Spectroscopy and Microscopy via Spectral Notch Shaping.* Optics Express, vol. 18, no. 22, p. 22693, 2010.

[30] V. KUMAR, R. OSELLAME, R. RAMPONI, G. CERULLO and M. MARANGONI - *Background-Free Broadband CARS Spectroscopy from a 1-MHz Ytterbium Laser.* Optics Express, vol. 19, no. 16, p. 15143, 2011.

[31] S.P. KEARNEY, D.J. SCOGLIETTI and C.J. KLIEWER - *Hybrid Femtosecond/Picosecond Rotational Coherent Anti-Stokes Raman Scattering Temperature and Concentration Measurements Using Two Different Picosecond-Duration Probes.* Optics Express, vol. 21, no. 10, p. 12327, 2013.

[32] S.P. KEARNEY and D.J. SCOGLIETTI - *Hybrid Femtosecond/Picosecond Rotational Coherent Anti-Stokes Raman Scattering at Flame Temperatures Using a Second-Harmonic Bandwidth-Compressed Probe.* Optics Letters, vol. 38, no. 6, p. 833, 2013.

[33] J.D. MILLER, M.N. SLIPCHENKO and T.R. MEYER - *Probe-Pulse Optimization for Nonresonant Suppression in Hybrid fs/ps Coherent Anti-Stokes Raman Scattering at High Temperature*. Optics Express, vol. 19, no. 4, p. 13326, 2011.

[34] H.U. STAUFFER, J.D. MILLER, S. ROY, J.R. GORD and T.R. MEYER - *Hybrid Femtosecond/Picosecond Rotational Coherent Anti-Stokes Raman Scattering Thermometry Using a Narrowband Time-Asymmetric Probe Pulse*. The Journal of Chemical Physics, vol. 136, no. 11, p. 111101, 2012.

[35] J. LUMEAU, V. SMIRNOV, A. GLEBOV and L.B. GLEBOV - *Ultra-Narrow Bandpass Filters Based on Volume Bragg Grating Technologies.* Proc. of SPIE, vol. 7675, p. 76750H, 2010.

[36] HERZBERG - Molecular Spectra and Molecular Structure. I. Spectra of Diatomic Molecules.: Van Nostrand, 1950.

[37] H.U. STAUFFER, J.D. MILLER, M.N. SLIPCHENKO, T.R. MEYER, B.D PRINCE, S. ROY and J.R GORD - *Time- and Frequency-Dependent Model of Time-Resolved Coherent Anti-Stokes Raman Scattering (CARS) with a Picosecond-Duration Probe Pulse.* The Journal of Chemical Physics, vol. 140, p. 024316, 2014.

[38] C.E. DEDIC, J.D. MILLER and T.R. MEYER - *Dual-Pump Vibrational/Rotational Femtosecond/Picosecond Coherent Anti-Stokes Raman Scattering Temperature and Species Measurements.* Optics Letters, vol. 39, no. 23, p. 6608, 2014.

[39] D. ORON, N. DUDOVITCH and Y. SILBERBERG - *Femtosecond Phase-and-Polarization Control for Background-Free Coherent Anti-Stokes Raman Spectroscopy.* Physics Review Letters, vol. 90, no. 21, p. 213902, 2003.

[40] A.F. BUNKIN, S.G. IVANOV and N.I. KOROTEEV - Soviet Technical Physics Letters. vol. 3, p. 182, 1977.

[41] L.A. RAHN, L.J. ZYCH and P.L. MATTERN - *Background Free CARS Studies of Carbon Monoxide in a Flame.* Optics Communications, vol. 30, no. 2, p. 249, 1979.

[42] F. VESTIN, M. AFZELIUS and P.-E. BENGTSSON - *Development of Rotational CARS for Combustion Diagnostics Using a Polarization Approach.* Proceedings of the Combustion Institute, vol. 31, no. 1, p. 833, 2007.

[43] F. EL-DIASTY - *Coherent Anti-Stokes Raman Scattering: Spectroscopy and Microscopy.* Vibrational Spectroscopy, vol. 55, no. 1, p. 1, 2011.

[44] M. DHYNE, M. LEPERE, and P. JOUBERT - *Semiclassical Line Broadening Calculations, Using an Ab Initio Potential Energy Surface, in Q-Branch and S-Branch of N_2 Perturbed by H_2.* Journal of Raman Spectroscopy, vol. 43, p. 2008, 2012.

[45] C. DORRER and M. JOFFRE, *Characterization of the Spectral Phase of Ultrashort Light Pulses.* Compte-rendu de l'Académie des Sciences de Paris, p. 1415, 2001.

[46] S.P. KEARNEY - *Hybrid fs/ps Rotational CARS Temperature and Oxygen Measurements in the Product Gases of Canonical Flat Flames.* Combustion and Flame, vol. 162, no. 5, p. 1748, 2015.

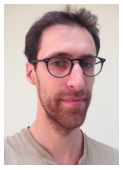
[47] D.R. RICHARDSON, R.P. LUCHT, W. D. KULATILAKA, S. ROY and J.R. GORD - *Theoretical Modeling of Single-Laser-Shot, Chirped-Probe-Pulse Femtosecond Coherent Anti-Stokes Raman Scattering Thermometry.* Applied Physics B, vol. 104, no. 3, p. 699, 2011.

[48] T. SEEGER and A. LEIPERTZ - *Experimental Comparison of Single-Shot Broadband Vibrational and Dual-Broadband Pure Rotational Coherent Anti-Stokes Raman Scattering in Hot Air*. Applied Optics, vol. 35, no. 15, p. 2665, 1996.

Acronyms

CARS	(Coherent Anti-Stokes Raman Scattering)
FPE	(Fabry Perot Etalon)
OPA	(Optical Parametric Amplifier)
NR	(Nonresonant)
R	(Resonant)
VBG	(Volume Bragg Grating)

Malik Nafa, graduated from the *Institut d'Optique* Graduate School in 2014. He is now a PhD student at ONERA working on the hybrid fs/ps CARS setup and on the simulation model described in this paper for thermometry applications in reactive flows.



Michael Scherman received his PhD in 2012 from *Université Pierre et Marie Curie*. He has been a research scientist at ONERA since then, in charge of the development of coherent and spontaneous Raman techniques applied to diagnostics in gaseous reactive flows, material characterization and biological sample studies.

Alexandre Bresson, received his Ph.D. degree in 2000, for his work on "quantitative imaging based on a laser-induced fluorescence technique" at ONERA – the French aerospace lab. He has been working for 4 years at ONERA, specifically on the development of optical diagnostic tools for combustion. Since 2004, he has continued his research on the use of cold atoms for on-board inertial sensing. He currently leads the "laser source and metrology" research unit at ONERA, in Palaiseau.

Alexandre Aubin is a student from the *Institut Universitaire de Technologie d'Orsay*. He built a stabilized CH_4/air flame, which is used in the lab as a reference of temperature for the CARS experiments.

Antoine Godard received his PhD from the *Laboratoire Charles Fabry de l'Institut d'Optique* in 2003. His research interests at ONERA concern the development of innovative laser sources based on nonlinear optics and their applications (gas sensing, onboard applications, etc.).

Brigitte Attal-Trétout performed the first Coherent Anti-Stokes Raman Scattering spectroscopy (CARS) experiments at ONERA. The so-called "resonant" version of CARS was created within this framework and used to record the first spectra of iodine. Since then, the Raman techniques and their degenerated version have been very widely applied to the measurement of traces in reactive environments. Brigitte Attal-Trétout has worked on optical metrology and nonlinear spectroscopy applied to plasmas and flames. She has directed a research unit and oriented field of applications to nano-object detection during their synthesis together with atom concentration measurements during material synthesis and in propellants flames. Finally, she constructed a CARS microscope for biology. Brigitte Attal-Trétout is currently a research director and scientific deputy of the Department of instrumentation and sensing.

Pierre Joubert works on the topic of molecular spectroscopy. He has performed semi-classical calculations of line broadening and line shifting coefficients applied to molecular systems involved in combustion media. He has also developed new spectral line profile models implemented using various thermometry techniques for a wide range of pressures and temperatures. He is the head of a physics and astrophysics research group at the UTINAM Institute. Pierre Joubert is currently a professor at Franche-Comté University.

P. Touboul
(ONERA)

G. Métris
(Geoazur – CNRS/UMR)

H. Sélig
(ZARM Space Science Department
University of Bremen)

**O. Le Traon, A. Bresson,
N. Zahzam, B. Christophe,
M. Rodrigues**
(ONERA)

E-mail: pierre.touboul@onera.fr

DOI: 10.12762/2016.AL12-11

Gravitation and Geodesy with Inertial Sensors, from Ground to Space

Since the years 2000, three space missions, CHAMP, GRACE, and GOCE, have led us to consider the Earth's gravitational field and its measurement in a new light, using dedicated sensors and adequate data processing, revealing the changes in the Earth's field as the true signal rather than the disturbing terms in addition to the geostatic reference field. Besides the possibilities offered by new technologies for the development of inertial sensors, a space environment of course involves special constraints, but also allows the possibility of a specific optimization of the concepts and techniques well suited for microgravity conditions. We will analyze and compare with others the interest in the electrostatic configuration of the instruments used in the main payload of these missions, and we will consider the recent MICROSCOPE mission, which takes advantage of the same mission configuration as a gradiometry mission to test the universality of free fall whatever the mass composition. A few days after launching the satellite in April 2016, we will show how we intend to validate the future result, the existence or not of a violation signal of the equivalence principle, taking into account the laboratory tests, where available, and the in-flight demonstrated performance during the calibration phases and the scientific measurements. With regard to ground measurements, either fixed or mobile, or under marine or aircraft conditions, we will demonstrate the complementary interest of the atomic interferometer. Finally, we will briefly discuss the future envisaged for these technologies, like that already implemented in the Lisa-Pathfinder mission without a gold wire for the electrical control of the charges of the mass, and these types of mission.

Introduction

We have passed the century anniversary of General Relativity [1], which remains the geometrical foundation of gravity, after thirty years of non-conclusive efforts with String theories and others, like quantum loop gravity: no consistent quantum theory of gravity exists as yet. Attempts to go beyond the Standard Model recently confirmed by CERN with the detection of the Higgs boson [2] and to unify the gravitational interaction with the three other interactions, electromagnetic, weak and strong, have mostly led to the discovery of new particles and forces. Accurate cosmological observations have led to the discovery of the existence of dark energy and dark matter for 95 % of our Universe, while gravity wave observatories are now operating after the first modeled signal of two merging black holes [3].

This context motivates the quest for new observations and for laboratory or space experiments to test gravity and in particular to test the Equivalence Principle, which is the basis of the theory of general relativity. In order to perform such an Equivalence Principle test in orbit, we have optimized the instrument configuration of the space accelerometers that were integrated on board three successive geodesy space missions launched during the last decade [4], [5], [6].

All of these instruments take advantage of the space environment. Thus, they have been configured differently from more conventional inertial sensors used for navigation, and are now based on MEMS (Micro-Electro-Mechanical Systems) technologies for size and cost reduction, while preserving high performance for some. Some are based on the old spring-mass concept, but surpass the nominal limits because of their high-resolution capacitive position sensors, their servo-loops and their electrostatic levitation of a specific solid mass. They nevertheless require dedicated ground and flight calibration with specific facilities.

In-orbit performance of several tens of pico-g has been demonstrated, while the detection of the femto-g signal is expected with the MICROSCOPE mission, launched on April 25 this year, in the case of a well-defined sine at a very low frequency of about 10^{-3}Hz. The operation of an inertial sensor, whatever the technology and the concept are, considers the proof-mass in a geodesic motion, which is subject not only to the gravitational field but also to the acceleration field and to any accurately measured force. Thus, in free fall, the inertial sensor output is nullified, as the difference of the gravitational field and the acceleration field (assuming the weak equivalence principle); when the inertial sensor is fixed to a moving body, the sensor output is the difference of the acceleration of the body and the gravitational field. Thus, it is possible to deduce from the provided data, information on either the inertial acceleration or the gravitational field, and on the difference in the MICROSCOPE specific experiment. Atomic wave interferometry involves, in the same way, both fields and will in the future be an interesting new complementary technology. The Lisa-Pathfinder space mission is also the technological way to test new technologies and, in particular, the control of the test-mass charge without any contact [99], [100].

This paper offers an overview of the driving parameters and characteristics of these instruments, the mission concepts and some specific needed procedures, to compare them with other technologies and similar applications, and to analyze their perspective with regard to future space missions or airborne campaigns for the understanding and mapping of gravity.

Inertial sensors

Over this last century, inertial sensors have experienced continuous improvement, from the early guidance and navigation systems and the "famous" German V2 ballistic missile of World War II, followed by the Cold War and the Space Race through the emblematic Apollo program and the first humans on the Moon with Apollo 11 on July 20, 1969, to the recent period with both the arrival of the Global Positioning System and micro/nanotechnologies, enabling today precise navigation with a simple smartphone.

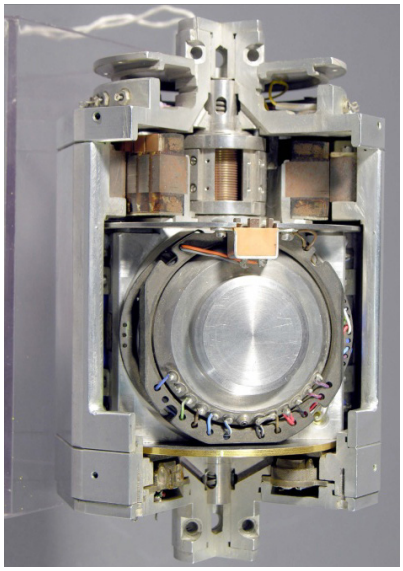
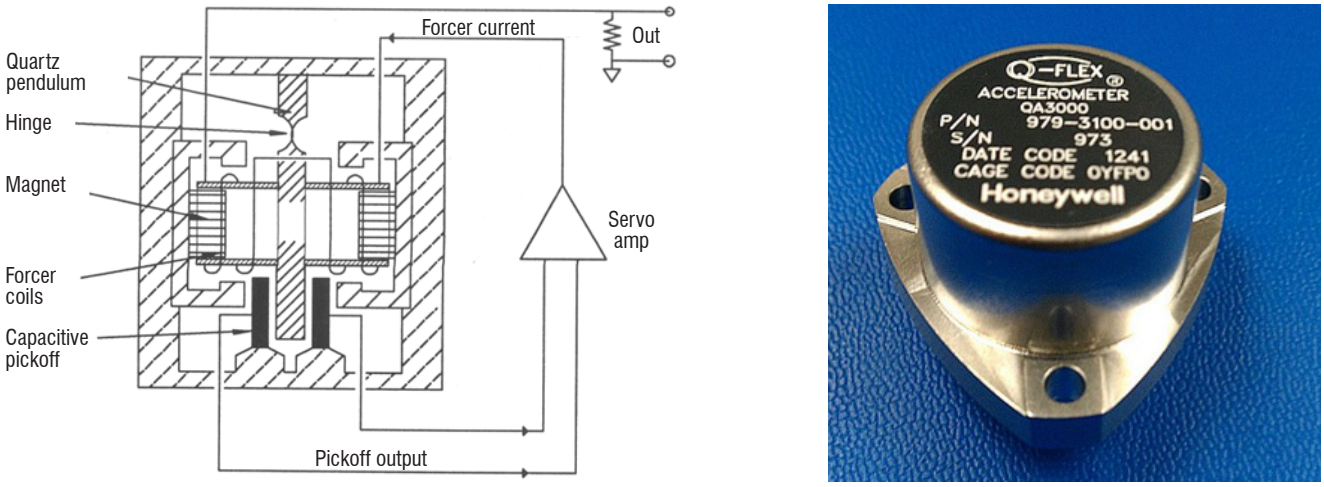


Figure 1 – 25 Piga Accelerometer for the Titan II Intercontinental Ballistic Missile – *Courtesy of AC Spark Plug Division, General Motors Corporation*

The "Pendulous Integrating Gyroscope Accelerometer", PIGA, [7] [8] invented by F.K. Mueller for the V2 missile guidance, remains today the most accurate high dynamic range accelerometer (~50 g range, stability and resolution on the order of 0.1 μg), but is also very complex, bulky (0.5 liter volume and a weight of 3.5 kg for most integrated PIGA) and very expensive to manufacture (Figure 1). The PIGA is based on the conversion of an acceleration input into a gyroscopic torque, induced by a gyroscope in a pendulous configuration. The whole pendulum is mounted into a housing, and the whole pendulum is able to rotate in order to compensate for the gyroscopic torque induced by the acceleration. The rotation rate of the housing is thus proportional to the acceleration input and the rotation angle directly gives the velocity, which is very interesting for precise ballistic systems. However, PIGA accelerometers are sensitive to rotation along the input acceleration axis and need a gimbaled inertial platform, which further complicates the inertial system.



(a) (b)

Figure 2 – (a) closed loop spring mass accelerometer proposed by Sundstrand (now Honeywell), (b) the famous one-inch diameter inertial grade Q-flex QA3000 accelerometer – *Courtesy of Honeywell*

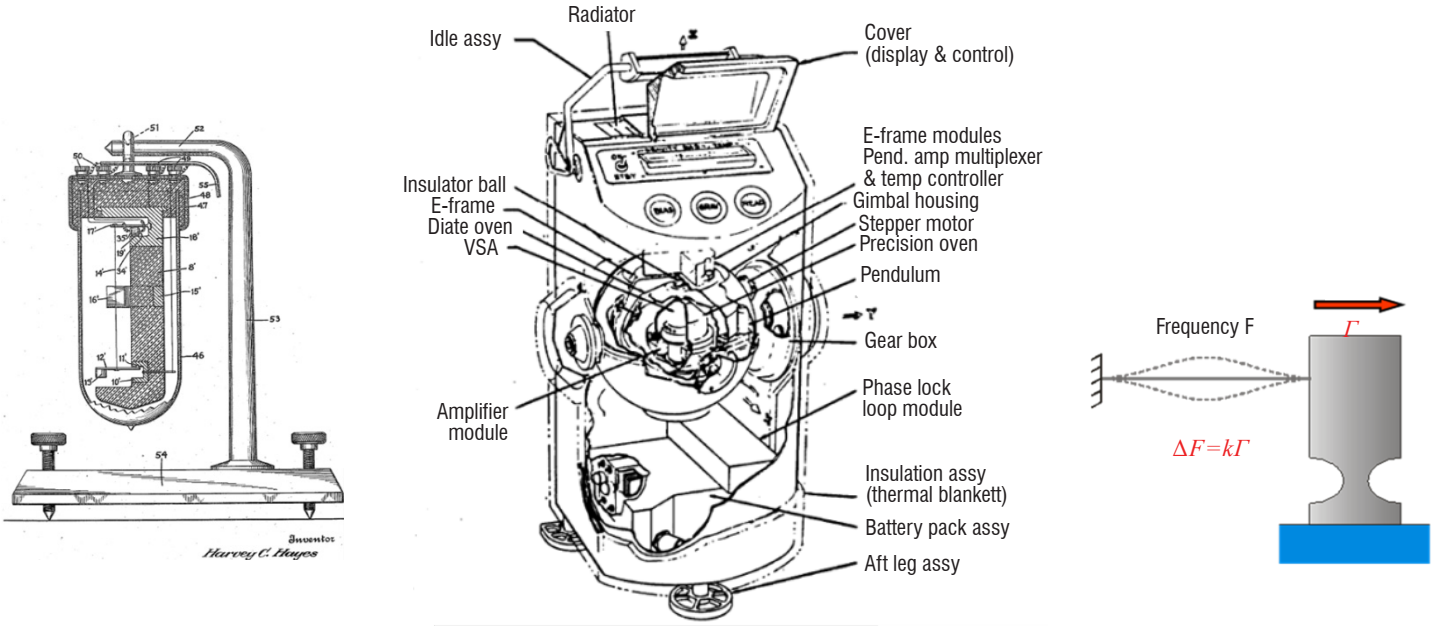


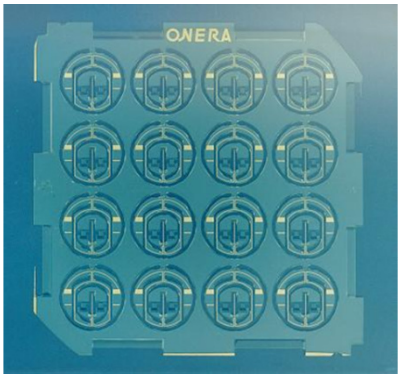
Figure 3 – (a) first vibrating string gravimeter proposed by H. C. Hayes in 1928, (b) lunar VSA gravimeter proposed by the Charles Stark Draper Laboratory for the Apollo 17 mission, (c) vibrating beam accelerometer, the string is changed by a beam in bending mode

The spring mass concept is the conventional configuration of an accelerometer and has known major improvements following the "strap-down" inertial navigation proposed in the early 1970's (i.e., the inertial components are directly mounted on the vehicle, without a gimbaled inertial platform). Figure 2 shows the nominal configuration, either in open or closed loop, and the "Q-flex" accelerometer proposed in the 1990's by Sundstrand [9] (now Honeywell) remains a reference for wide range inertial navigation grade accelerometers. Thanks to its magnetic feedback and to its particularly high-stability quartz hinge, it reaches a high measurement range of ± 60 g, while exhibiting a one year bias stability (all errors combined) better than 40 μg, a scale factor repeatability better than 80 ppm and an intrinsic noise of a few $\mu g/\sqrt{Hz}$.

Another very surprising kind of spring mass accelerometer is the Vibrating Beam Accelerometer (VBA). Vibrating Beam Accelerometers (VBAs) are based on the change in the resonance frequency of a vibrating beam when subjected to acceleration (Figure 3a). The idea of the direct conversion of the acceleration in the change of a resonator frequency is not new and was first proposed in 1928 [10]. At that time, the resonator was a metallic string, excited to its resonance frequency by electromagnetic forces, whence the name of Vibrating String Accelerometer (VSA). A basic limitation of VSA was its inherent poor bias stability (frequency stability under zero acceleration), mainly due to the necessary pre-tension of the string and also to the low quality factor of the string. However, it must be noticed that, within the context of the Apollo program in the 1970's, lunar gravity was successfully measured with an accuracy of 1 m-gal (10^{-5} m/s^2) thanks to a Vibrating String Accelerometer, including a reversal system in order to reject the bias [11].

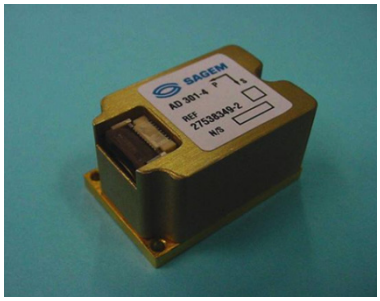
A first major evolution of vibrating accelerometers was achieved by changing the string for a vibrating beam in flexure mode [12]: thanks to its compression stiffness, a vibrating beam doesn't require pre-tension, enabling the improvement of the bias stability. The beam geometry has very often been a research subject in order to reduce the impact of its mounting on the accelerometer structure and to preserve the quality factor of the beam (generally made of quartz crystal). Various beam geometries have been proposed: a simple beam held at its nodal lines [13] [14] [15], a simple beam with decoupling systems on its ends [16] [17] [18] [19], optimized double tuning forks [20] or a triple-beam resonator [21]. Industrial developments of vibrating beam accelerometers (VBA) have been undertaken, leading at the end of the 80s to commercial products, such as the well-known RBA 500 [22], which is still in production today [23].

At the end of the 1980's, riding the waves of the Micro Electro Mechanical System (MEMS) revolution, heavily used by the silicon industry [24], monolithic VBA structures – i.e., the whole accelerometer structure, including not only the proof mass but also the hinges and the beam resonator – made of the same material (generally quartz or silicon crystals) emerged [25] [26] [27] [28]. A monolithic accelerometer structure presents a lot of advantages: reduction of pieces and adjustment, matching of the thermal behavior, elimination of the delicate resonator mounting onto the accelerometer structure, and compatibility with collective micromachining techniques. Thus, all of these points are well suited to optimize the accelerometer accuracy and miniaturization, with a potential of a low manufacturing cost. Quartz and silicon VBA have been developed. The first of these takes advantage of the intrinsic piezoelectricity of quartz crystals, which enables an easy and accurate excitation and detection of the vibrating beam. The second benefits from the amazing developments in the semiconductor and microelectronics industry. Examples of VBAs that have been developed, mainly for military applications, leading or not to industrial products, are given in [29] [30] [31] for quartz devices and in [32] [33] [34] [35], [36] [37] for silicon devices. ONERA was one of the first laboratories to propose a monolithic configuration for the VBA [27] [33], and has developed the VIA accelerometer, which has now been transferred to

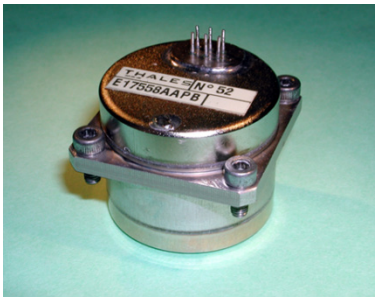


Quartz Wafer (1.5" x 1.5")

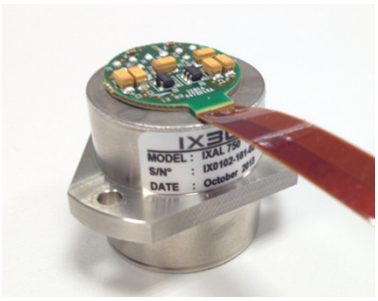
VIA Accelerometer



AD301 SAGEM
Quartz VBA



A100 THALES
Quartz VBA



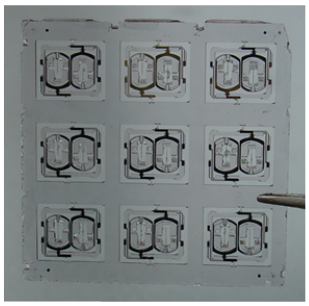
IXAL750 IXBlue
Quartz VBA

Figure 4 – Quartz Vibrating Beam Accelerometer developed at ONERA. (a) quartz wafer with 16 VIA accelerometers, (b) detail of the monolithic VIA accelerometer mounted on a base, (c) Vibrating Beam Accelerometers based on the VIA concept and produced by the companies SAGEM, THALES and IXblue.

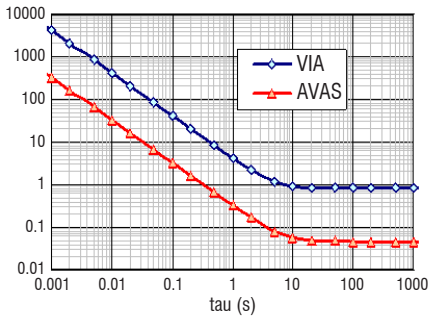
industry [39] [40] [41] (Figure 4). The main originality of the VIA concept lies in the use of a simple beam as a vibrating resonator and of a monolithic insulating system around the beam and the proof mass, allowing a very efficient insulation of the beam vibration. This insulating system preserves the high quality factor of the quartz beam, as well as any thermal stresses due to the mounting of the sensitive monolithic quartz element onto its base, which are two necessary conditions for achieving high bias stability. The VIA Quartz MEMS accelerometer is an excellent tactical grade accelerometer with a measurement range of 100 g, an excellent scale factor stability better than 10 ppm, a bias stability (all errors combined) better than 300 μg, a scale factor stability better than 10 ppm and a resolution of 1 μg @ 10s (until now, no equivalent MEMS device has been produced in the world).

New Quartz MEMS VBAs are under development at ONERA. They are aimed at reaching the navigation and strategic grades, taking advantage on the one hand of the important improvements in quartz microtechnology and on the other hand of innovative concepts [42] [43], as well as progress in analogic/digital electronics able to perfectly control the oscillator circuit phase. Figure 5 shows two of the VBAs being studied: the navigation grade DIVA accelerometer (Differential Inertial Vibrating Accelerometer) with a measurement range of 50 g, a bias stability better than 50 μg (all errors combined) and a noise better than 1 μg @ 10s, and the high-resolution AVAS [44] with a noise of 50 nano- g @ 10s, a bias stability better than 1 μg and a measurement range of 10 g.

Other interesting configurations are also presently being studied at the micrometric scale, including original detection schemes such as optical [45], or electron tunneling effect schemes [46], with an excellent resolution of 20 nano-g, which is near the thermomechanical noise limit due to the small proof mass (in the milligram range) of



DIVA VBA



AVAS VBA

Figure 5 – Quartz MEMS Vibrating Beam Accelerometer under development at ONERA. Left: DIVA (Differential Inertial Vibrating Accelerometer) navigation grade VBA. Right: AVAS high-resolution VBA with a 50 nano-g resolution @10s.

this MEMS device. "On Earth" acceleration measurements are indeed subject to a compromise between measurement range and resolution. Fortunately, Space is another Universe (with other constraints).

Space environment and dedicated sensors

Space can be regarded as a constraint when developing an instrument, considering the importance of demonstrating the instrument reliability, the traceability of all procedures (production, maintenance, tests, etc.), the robustness to the radiation environment, and the resistance to launch vibrations and thermal vacuum. On the other hand, it can also be seen as an opportunity: one space mission can consider a global coverage of the Earth for a limited time and without geographical or political access difficulties. It can take advantage of the fine space environment provided by the satellite: magnetic, with shielding from the Earth's or the spacecraft's residual field; electrical, with shielding also; thermal, with rather well-defined external conditions and internal power sources; vibrations and gravitational fields, thanks to the satellite design and stiffness, without moving or rotating masses; the satellite can take advantage of a drag compensation system when the orbit of the satellite is too low to neglect the effects of the atmospheric drag and its fluctuations [47], or when specific requirements must be met when considering the maximum level of acceleration that the instrument is subjected to.

Micro-gravity operation can also been seen as an opportunity, when weak accelerations have to be detected or measured. In the case of an inertial sensor, it is the possibility of neglecting the normal gravity level, as large as 9.81 ms^{-2}, when considering the full scale range (FSR) of your sensor, thereby increasing its resolution, which is always a limitation for part of the FSR. As a rule of thumb, the atmospheric drag of a satellite equals its radiation pressure at around 700 km, like the altitude of the MICROSCOPE satellite [48], and reaches several 10^{-7} ms^{-2} at the altitude of the GRACE and CHAMP satellites, i.e., 450 km to 500 km [49], [50], and 10^{-5} ms^{-2} at very low altitude, like 270 km in the case of the GOCE satellite [51]. Thus, the FSR of the space inertial sensor can be reduced by a factor 10^6 with respect to those of ground, missile or aircraft sensors, allowing other technologies.

The resolution of an inertial sensor based on the spring-mass concept depends dramatically on the stiffness of the spring and the resolution of the position sensor that will detect the mass motion with respect to its fixing point (see Figure 6). While $N_x(f)$ is the power spectrum of the position detector noise, $N_\Gamma(f)$, the induced power

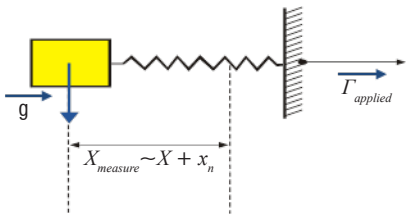


Figure 6 – Inertial sensor mass-spring concept: the acceleration of the accelerometer structure applied in the direction of the spring is measured through the displacement of the inertial mass, as well as the gravitational field applied to the proof-mass when it is in the spring axis

spectrum of the inertial sensor output is simply expressed by taking also into account the Nyquist noise, which is especially important at low frequencies, [52], [53] induced by the spring damping:

$$N_\Gamma(f) = N_x(f).16.\pi^4.\left(k/m + f^2\right)^2 + 1/m^2.\left(4.k_B.T.k/\left(2\pi.f.Q(f,T)\right)\right)$$

where k and Q, define the stiffness and the damping (through its quality factor) of the spring respectively, T is the temperature of the device, k_B is Boltzmann's constant, and f is the frequency of interest. In order to exhibit a very low acceleration noise, N_x and k must be limited. However, the frequency response of the spring-mass device depends mainly on the value of k when the mass is passive and not servo-controlled.

Let s be the Laplace variable expressing the derivation, and $\left(X(f) + xn(f)\right)$ the measurement of the proof-mass position (see Figure 6) from which the acceleration is deduced and that is servo-controlled to null; the frequency response of the sensor can then be expressed by:

$$\hat{\Gamma} \approx -\frac{k}{m}.\left(X(f) + x_n(f)\right)$$

when the mass is passive (in open loop operation) and

$$\hat{\Gamma} \approx -\frac{(2.\pi.f_c)^2}{(2.\pi.f_c)^2 + k/m + s^2}.\left(\Gamma(f) + \left(s^2 + k/m\right).x_n\right)$$

when the mass is servo-controlled (where f_c is the cut-off frequency of the control loop bandwidth simply defining the loop gain; the corrector may be more complex, but here it is only represented by this gain).

Thus, the sensor noise in the last case can be reduced by nullifying the stiffness, while the sensor frequency bandwidth is preserved by the loop control bandwidth. It is also important to point out that the damping provided by the corrector of the electrostatic loop (not considered in the above equation for simplicity purposes) is a "cold" damping that is not considered when computing the Nyquist dissipation [54]. In addition, the perturbing accelerations applied directly to the mass (magnetic, electric, thermal and gravitational), which may depend on the geometry of the mass environment, do not fluctuate because of the motionless mass.

Two types of technology have been proposed for mass levitation and its position sensing. One consists of a magnetic superconducting suspension [55], making use of the sensibility of squid devices for the position detector [56], [57] and operating at a cryogenic temperature, mainly at liquid He temperature, 4.2 K, or even lower. The SQUID detects the motion of the mass in front of a superconductive loop through the variation of the induced magnetic flux. The stability of the suspension and the resolution of the SQUID benefits from the low temperature and the superconducting shielding. However, the mass is not servo-controlled and the operation on board a satellite is very complex due to the helium Dewar and the necessary adjustment of the different currents of the loops corresponding to the right position and attitude of the levitated mass. Superconducting accelerometers [58],[59] have been proposed in the past for various space missions, but have never been selected by NASA or ESA, as gavity-gradiometers [60], [61], that can nevertheless benefit of specific SQUID circuits to perform directly the differential measurements of the test-mass positions and accelerations.

The other consists in the electrostatic levitation of the test mass; this is what has been done for the GPB mission [62] gyrometer ball, with the servo control of its position, only three degrees of freedom (translations), while the attitude of the spinning ball is deduced from the London effect. Given that the ball rotates with respect to the cage electrodes, its precise motion is very sensitive to the patch effects on both conductors [63]. When the six degrees of freedom of the cubic proof-mass are servo-controlled, the levitation is only biased by these potentials, which introduce only DC offsets into the inertial sensor outputs. Thus, the patch fluctuations are the only ones to be considered when one is interested by a signal at a frequency that is not at DC.

This is the electrostatic configuration of the CHAMP, GRACE and GOCE accelerometers [64]. The solid metallic proof-mass is surrounded by at least six pairs of electrodes. Each pair of opposite electrodes performs the capacitive measurement of the mass part between the two, and the differences in voltage between the mass and the electrodes are controlled and generate electrostatic pressures that lead to mass forces and torques. The resultant is finely measured to provide the six acceleration outputs (3 linear and 3 angular) of the inertial sensor; the orthogonal faces of the conductive mass define the frame of the six outputs [65].

The drag-free sensor of the current Lisa-Pathfinder space mission [66] is also based on the electrostatic control of a solid mass. In science mode, along the interferometric direction, the mass must follow a geodesic motion without any non-gravitational force and the detection of the gravitational wave comes from the optical interferometer output. Thus, along the interferometer axis, this is different and the acceleration of the mass does not depend on the servo-electrostatic forces, which must be finely measured by the sensor itself when accelerometers are considered. Thus, the configuration can be optimized in a different way, in particular considering only the sensitivity of the capacitive sensing for measurement [67], [68].

Performance of the GOCE mission, GRACE-FO mission and MICROSCOPE mission electrostatic instruments

One of the main challenges in the realization of the space accelerometer is the verification of the performance. Indeed, due to the presence of gravity on the ground, it is not possible to verify this performance through a dedicated test. The strategy for the verification relies on a combination of the mathematical formulation of the impact of each contributor, specific tests for assessing the level of this contributor, and finally flight verification through the post-processing of the flight data.

For each type of mission, the first step to determine the performance of the instrument is to write the measure equation. For the CHAMP, GRACE or GRACE-FO missions, the accelerometer mainly measures the non-gravitational forces (residual drag, solar radiation pressure, etc.) exerted on the spacecraft. For the MICROSCOPE and GOCE missions, the use of several accelerometers inside the spacecraft enables the measurement of the differential acceleration between accelerometers: this measurement enables either the Eötvös parameter (related to the Einstein equivalence principle) or the gravity gradient to be determined. To simplify, we will focus on the first type of missions, the measurement of the non-gravitational acceleration. For the GOCE or MICROSCOPE missions, the accelerometers also measure this acceleration (used for example for the drag compensation), and the difference in acceleration allows just the impact of the common acceleration to be reduced.

In principle, the output of an electrostatic accelerometer, like those developed by ONERA, is the relative acceleration between the proof-mass and the electrode cage, which is applied through the electrostatic forces to maintain the proof-mass at the center of this cage. The cage is servo-controlled and fixed to the spacecraft, so its acceleration is due to the acceleration applied to the center of gravity of the spacecraft and to the inertial acceleration if the accelerometer is not at the center of gravity. The proof-mass is nominally subjected only to the gravitation acceleration at its location. Taking into account parasitic acceleration on the proof-mass and the spacecraft, or deformation of the spacecraft, the acceleration as seen by the accelerometer is expressed by the following equation:

$$\underline{a}_{ACC} = \underline{a}_{NG} - [U]\underline{r} + \left([\dot{\Omega}] + [\Omega^2]\right)\underline{r} + \underline{a}_{para\_SC} + \underline{a}_{para\_acc} + 2[\Omega]\dot{\underline{r}} + \ddot{\underline{r}}$$

Where $\underline{a}_{NG}$ represents the non-gravitational acceleration,

$[U]$ is the gravity gradient tensor,

$\underline{r}$ is the vector between the proof-mass location and the center of gravity, $\dot{\underline{r}}$ and $\ddot{\underline{r}}$ being the velocity and acceleration of this vector in the spacecraft reference frame (therefore representing the deformation of the spacecraft),

$[\dot{\Omega}]$ and $[\Omega^2]$ are the angular rate and acceleration tensors,

$\underline{a}_{para\_SC}$ is the parasitic acceleration due to the spacecraft (e.g., the magnetic perturbation),

$\underline{a}_{para\_acc}$ apara_acc is the parasitic acceleration due to the instrument (e.g., the radiometric effect).

The 2 first terms generally represent the signal that we want to measure (for GRACE or GRACE-FO, we want to measure the residual drag, for GOCE we want to measure the gravity gradient tensor). The other terms represent errors to be minimized or corrected.

Beyond these terms, it is also necessary to take into account the fact that the acceleration seen by the accelerometer is not perfectly measured, due to the imperfect scale factor ($[dK]$), the instrument noise (n) or bias (b), the non-linearity ($[K2]$) or the imperfect alignment or the coupling between the axes ($[R+S]$), leading to additional sources of errors:

$$\underline{a}_{meas} = \underline{a}_{ACC} + [dK]\underline{a}_{ACC} + \underline{n} + \underline{b} + [K2]\underline{a}^2_{ACC} + [R+S]\underline{a}_{ACC}$$

Often, the performance of an accelerometer is defined by its noise in a frequency-amplitude space. Most of the engineering design activity consists in first reducing, and second optimizing and tuning the noise of the various contributors with respect to the target bandwidth and the required level specification. The noise from the electronics is precisely measured on the ground and converted into acceleration. The parasitic noise contributions are deduced from a specific test ([69], [70]).

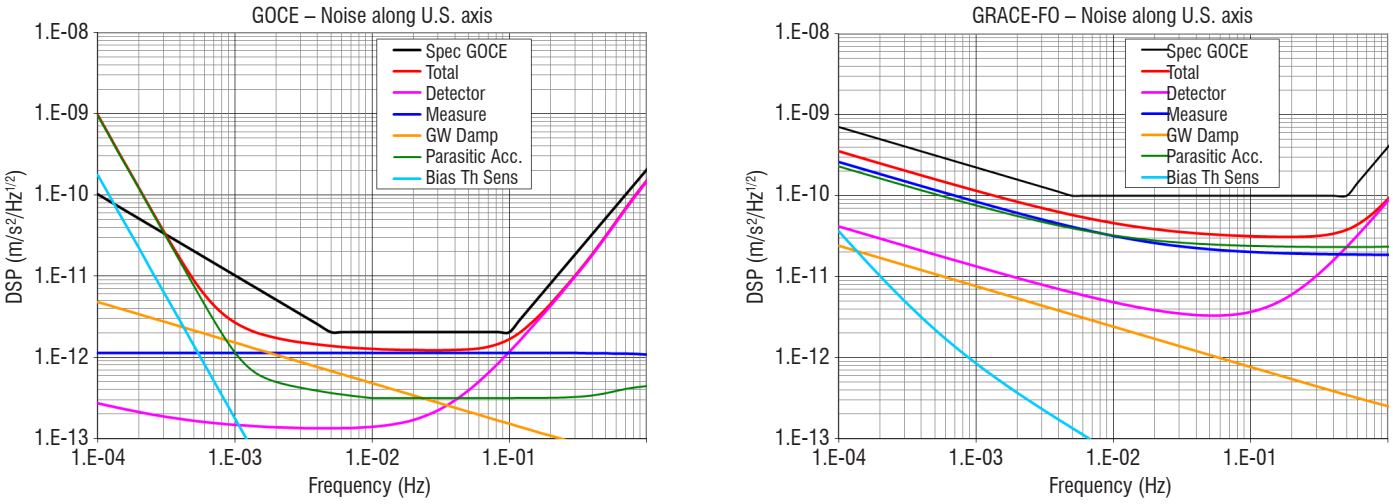


Figure 7 – Noise performance of the GOCE accelerometer (a) and GRACE-FO accelerometer (b) along their ultra-sensitive axis

The result is shown in Figure 7, for the GOCE [71] and GRACE-FO [72] accelerometers, at the same scale for comparison, where only the main contributors have been selected:

- The detector noise (pink). This is basically due to the electronic noise in the detector circuit; the control loop induces a detector noise increase with frequency.
- The wire damping (orange): a thin gold or platinum wire is used to impose the mass electrical potential, whatever the space radiation. The wire stiffness and damping is minimized by the use of a 5 μm diameter gold wire on GOCE and a 7 μm diameter platinum wire on GRACE-FO; the effect of the wire damping in the control loop produces a noise level limit that fortunately remains far below that required.
- The measurement readout and digitalization noise (dark blue). Usually, in digital circuits the quantification step of the measurement device is designed well below the noise level of the instrument and is not a limiting point. The present situation is the best trade-off between the range needs and the measurement noise.
- The thermal sensitivity of the bias (light blue). Despite the accelerometer thermal control, the temperature variation leads to the variation of the electronic bias at low frequencies. In the case of GOCE, the gradiometer manufacturer announced a great increase in the temperature stability at very low frequency, which was pessimistic.
- The parasitic acceleration inside the accelerometer core (green). The main contributors are the thermal sensitivity of the radiometer effect, due to the difference of temperature between the faces of the proof-mass, and the patch effect due to the electrode and proof-mass surface states.

Since the performance is not achievable on the ground, flight post-processing enables the performance prediction to be confirmed. When only one accelerometer is present (like in the case of the GRACE mission), the in-flight verification will be done during a quiet period and will suffer from the low frequency drag. Nevertheless, the accelerometer performance has been verified for GRACE ([73]). In the case of GOCE, the combination of several accelerometer outputs eliminates the common mode (the residual drag) and yields information on the intrinsic performance. It has also been possible with GOCE to verify the electronic noise or internal stiffness in orbit, thanks to the versatility of the digital loop ([74]).

Validation of the inertial sensor performance

During the development of the sensor, three types of tests are performed to validate the expected performance. On the one hand, the geometries of the mechanical parts of the sensor, and in particular the test-mass, are finely controlled after their machining and throughout their integration in a clean room (size accuracy: down to the micron, parallelism, orthogonality, as well as the material properties: homogeneity of density, magnetic susceptibility, conductivity, cleanliness, etc.). On the other hand, the performance of the electronic units [81] is verified (sensitivity, bandwidth, linearity, noise spectrum, and thermal sensitivity) leading, for instance for the MICROSCOPE sensor configuration, to the following flight model results:

- Capacitive position sensors

Axis	Internal Mass (1.4 kg)	External Mass (0.4kg)
X	2 μVHz$^{-1/2}$ = 4 10^{-11} mHz$^{-1/2}$	6 μVHz$^{-1/2}$ = 2.5 10^{-11} mHz$^{-1/2}$
Y, Z	6 μVHz$^{-1/2}$ = 2.5 10^{-11} mHz$^{-1/2}$	3 μVHz$^{-1/2}$ = 1 10^{-11} mHz$^{-1/2}$

- Electrostatic actuators (electric potentials on electrodes)

Axis	Internal Mass	External Mass
X	1.1 μVHz$^{-1/2}$ = 20 10^{-15} NHz$^{-1/2}$	1.6 μVHz$^{-1/2}$ = 52 10^{-15} NHz$^{-1/2}$
Y, Z	2.3 μVHz$^{-1/2}$ = 160 10^{-15} NHz$^{-1/2}$	2.3 μVHz$^{-1/2}$ = 710 10^{-15} NHz$^{-1/2}$

In addition, the perturbation forces considered in the established error budget [80] must be validated for each instrument: gas damping, radiometer effects, radiation pressure, magnetic susceptibility,

electrostatic patch effects, and damping and stiffness of the thin gold wire used to control the mass electrical potential [82], [83], etc. However, this is not the integrated flight model test.

Inertial sensors for space missions like GRACE [5], GOCE [6] and MICROSCOPE are designed and constructed for operation in a zero-g environment. It is a matter of fact that these sensors and their performance cannot be tested under ordinary laboratory conditions at 1g. In principle, there are various solutions to overcome this general problem: (i) Electrostatic levitation of the test mass in order to enable sensor tests in the horizontal plane, (ii) mechanical levitation of the test mass with a wire and, finally (iii), free-fall tests of the sensors to allow sensor tests in all degrees of freedom at the same time.

The first possibility has been considered in the development of the first mentioned instruments. The proof-mass is levitated in a specific configuration along the vertical axis, while the two other axes are tested near the horizontal plane. The laboratory levitation of the mass along the vertical direction requires the smallest gap between the mass and the electrodes, 30 μm instead of 300 μm in the case of the GOCE accelerometer, reducing the resolution of the sensor in space along this axis. On the ground, the voltage applied on the electrodes is also increased due to the presence of normal gravity and can reach levels greater than 1000 V. Furthermore, the sensor must be tested, mounted on an anti-seismic platform, in order to eliminate horizontal disturbances and to maintain the instrument references with respect to the vertical direction. This approach is very interesting when the ground requirements do not interfere with the space performance, because the duration of the tests is not a constraint.

The second possibility has been considered in the development of the LISA-Pathfinder drag-free sensors [75]. However, it takes a huge effort to mimic the configuration of the sensor considered and to measure the torque on the torsion pendulum wire, which represents the phenomenon to be tested. Improvements to the facility are continuously being considered [76], in order to better model the sensor behavior, but the flight models are not actually tested.

This is not the case with the last possibility, which in particular uses the ZARM facility. The drop tower at ZARM allows free fall tests with a zero-g period of between 4.7 and 9.3 seconds, depending on the operation mode of the facility (normal drop mode or catapult mode) [77]. In both modes, a drop capsule with its payload undergoes nearly perfect zero-g conditions inside an evacuated drop tube. For each drop test, the tube must be evacuated by means of high power vacuum pumps. At the end of the free fall phase the capsule is captured by a deceleration chamber filled with small polystyrene balls. For the recovery of the capsule, the drop tube has to be flooded with dried air. The number of drops is limited to three per day. The free fall duration is limited by the height of the drop tube and the corresponding free fall height (110 m for the drop tower in Bremen). Although the best zero-g quality is rather good ($\sim 10^{-6}$g), residual accelerations act on the capsule and the payload. The aerodynamic drag due to the residual air pressure of 10 Pa inside the evacuated drop tube generates acceleration in the opposite direction to that of the capsule velocity vector. The corresponding maximum acceleration is around 10^{-5}g. The second residual acceleration is the centrifugal acceleration due to the residual capsule spin rates. Although these rates are rather small, the payload is mounted near the capsule center of mass, in order to minimize the centrifugal acceleration level. Finally, a certain level of vibrations is induced by the release of the capsule in normal mode or by the initial capsule acceleration in catapult mode. In order to reduce the vibration acting on the payload, a passive damping system is used, which has been tested and optimized in a number of drop tests with a SuperStar sensor provided by ONERA.

The best possible zero-g quality is achieved by using a free flyer platform inside the drop capsule. Due to technical reasons, this option is only available in the normal drop mode. The free flyer is released shortly after the capsule release and captured before impact by a pneumatic system. This technique is interesting, especially for experiments that demand a very high zero-g quality. The zero-g quality is improved by 1-2 orders of magnitude compared to the normal drop capsule. Many drop tests have been carried out for the MICROSCOPE sensor validation using the free flyer technique. The drawback of the free flyer is the relatively short zero-g duration of around 4 seconds, which is insufficient for testing all sensor parameters. Therefore, the final payload tests for MICROSCOPE have been performed in catapult mode. The main advantage of the catapult mode is the long zero-g duration of 9.3 seconds. Another advantage is the possibility of measuring the sensor bias of an accelerometer directly at the highest point of the capsule trajectory [78]. At this point, the capsule velocity is zero and therefore the residual acceleration induced by the drag also vanishes. The time when the capsule reaches the highest point of the trajectory can be calculated very precisely.

The ZARM drop tower team has been cooperating with ONERA for many years, especially in the area of free fall tests of inertial sensors for space missions. The correct zero-g-operation has been demonstrated for the GRACE and GOCE sensors, and finally for the MICROSCOPE

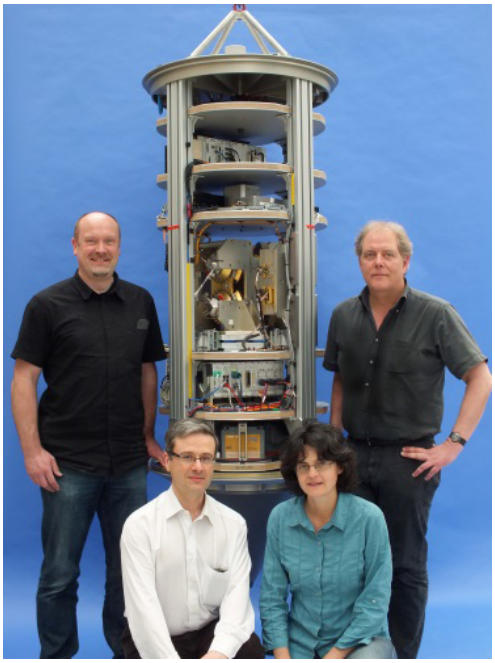


Figure 8 – Catapult capsule for the MICROSCOPE free fall tests at the ZARM drop tower in Bremen. The MICROSCOPE sensor unit is installed close to the center of mass of the capsule. The various platforms carry a computer system, batteries and radio link systems for communication with the control room. For the free fall tests inside an evacuated drop tube, the setup is covered with a pressure-tight aluminum cylinder, which is not shown in this figure.

sensors. According to the corresponding rules, the sensors have been tested before and after vibration, thermal vacuum and shock tests, in order to demonstrate their proper function in space. The tests of the MICROSCOPE sensors required some effort especially, due to having the highest sensor sensitivity and due to the corresponding demand for a very quiet environment and low disturbance level.

However, free fall tests with a duration of a few seconds cannot replace the complete and precise sensor calibration, which always has to be carried out in special calibration measurement sessions during the mission in space. However at least the proper functioning of space accelerometers in all degrees of freedom can be demonstrated very well with the help of the ZARM drop tower facility. Also, some sensor parameters can be roughly evaluated, such as sensor bias and scale factors.

Mission requirements, MICROSCOPE

In the case of MICROSCOPE, which is aimed at testing the Equivalence Principle (EP) in space with an accuracy of 10^{-15}, the mission requirements are directly derived from the test-mass equation of measure, as described in §4. The main difference with other missions is that we distinguish here the inertial mass m_i from the gravitational mass m_g in the equations.

The inertial mass defines the force of inertia to accelerate a body: $\vec{F} = m_i \vec{\Gamma}$.

The gravitational mass defines the weight in a gravitational field: $\vec{F} = m_g \vec{g}$.

For a body in free-fall in a uniform gravitational field g, Newton's laws lead to the acceleration being expressed as $m_i \vec{\Gamma} = m_g \vec{g}$.

Einstein postulated as a principle in his General Relativity theory [79] that in a uniform gravitational field the acceleration is equal to the gravity. In other words, in an accelerated vehicle far from any gravitational field, the passenger should be able to perform any physics experiment as if he was in a gravitational field. The consequence is, first, that inertial mass and gravitational mass are equal and second that all bodies in free fall follow the same trajectory regardless of their mass or composition.

In the case of two perfectly concentric bodies in free-fall, the difference in the acceleration felt by the two bodies is:

$\overrightarrow{\Gamma_1} - \overrightarrow{\Gamma_2} = \left(\frac{m_{1g}}{m_{1i}} - \frac{m_{2g}}{m_{2i}} \right) \vec{g}$ as, since the two bodies are subjected to the same gravitational field $\delta = \frac{2\left(\frac{m_{1g}}{m_{1i}} - \frac{m_{2g}}{m_{2i}} \right)}{\left(\frac{m_{1g}}{m_{1i}} + \frac{m_{2g}}{m_{2i}} \right)}$ is commonly defined as the Eötvös parameter and expresses the deviation from the Equivalence Principle. In a uniform gravitational field, the Eötvös parameter is calculated with the ratio between the differential acceleration and the mean acceleration of the 2 bodies:

$$\frac{\overrightarrow{\Gamma_1} - \overrightarrow{\Gamma_2}}{\left(\overrightarrow{\Gamma_1} + \overrightarrow{\Gamma_2} \right)/2}$$

Performing the EP test at 10^{-15} accuracy involves measuring and interpreting a difference in acceleration of at least 10^{-15}g, and the previous Eötvös parameter is computed for the two materials with 10^{-15} accuracy.

Actually, to perform the test, two different materials are used: a Platinum alloy (PtRH10: 10% of Rhodium) versus a Titanium alloy (TA6V: 6%Al + 4%V).

Testing the universality of free-fall in orbit, around the Earth, should be an ideal test if one is able to perfectly manage the initial conditions of the free-fall. In order to be more accurate, the free-fall of the test-masses is controlled by electrostatic forces and the test-masses are kept relatively motionless with respect to the surrounding control electrodes. The trajectory is maintained in the same way, and is identical for the two test-masses and identical to the drag-free satellite orbit, thanks to the drag compensation system; and if they are accelerated in the same way, Universality is verified, a dissymmetry in the applied acceleration reveals a violation of the Equivalence Principle. The principle of the electrostatic control is exactly the same as that of the electrostatic accelerometers used for the CHAMP, GRACE or GOCE payloads.

The satellite payload is thus composed of two differential accelerometers, each including two cylindrical and concentric test masses. The masses are made of the same material (PtRh10) for the first one, which is dedicated to assessing the accuracy of the EP experiments and the level of systematic disturbance errors. The mass materials are different for the second one. Then, the experimental procedure is based on a double comparison, in order to eliminate systematic errors. The selection of the mass material is a compromise between the instrument accuracy requirements, the theoretical interest and the technological feasibility.

The first step in establishing the requirements was focused on the test-mass shape. The electrostatic forces applied to the test-masses result from voltages applied on the surrounding electrodes. The geometry of the test-masses defines the capacitive environment used as input to the digital servo control of their motion. The volume forces, like gravity, act at the center of gravity G, which is defined for the volume of the mass by: $\int_V \overrightarrow{GP} \wedge \vec{g}(P)\, \rho dV = \vec{0}$.

The gravitational force exerted by a punctual source Ms (like a spacecraft device) at the test mass center of gravity G_0 is given by

$$\int_V \frac{G_0 Ms}{r^3} \rho(\vec{r}) \vec{r} dV$$

Given that each element of the test-mass is not located at the same distance from the punctual disturbing source, the equation becomes of first order: $mg_z(G) \approx \frac{G_0 Ms}{R^2} \left[m + \frac{3}{2R^2} \left(I_{xx} + I_{yy} - 2I_{zz} \right) \right]$, where z is the direction of the source at a distance R (considered much bigger than the dimensions of the test-mass), and I_{kk} is the main moment of inertia of the test-mass.

In order to neglect the first order term, the spacecraft gravitational field and fluctuation effects on the test-mass acceleration measurements, the cylinders have been manufactured so that their moment of inertia is identical around all axes. The cylindrical test mass can then be considered at first order as a sphere from the point of view

of gravity. The equality of all moments has been achieved at 0.1% by considering variations of 10^{-16}m/s^2 of the spacecraft gravity at orbital frequency (which is the modulation frequency of the Earth's gravity g when the spacecraft is inertial pointing and thus the EP frequency).

Thanks to the hypothesis of spherical inertia of the test-mass, the equation of measure is established as follows for two concentric test-masses [80]:

$$\begin{aligned}\vec{\Gamma}_{mes,d} = \underbrace{\vec{b}_{0d}}_{bias} &+ \left(\left[1+\underbrace{dK_{1c}}_{scale}\right]+\underbrace{[\eta_c]}_{coupl.}\right)\cdot\vec{b}_{1d} \\ &+\left(\left[1+\underbrace{dK_{1c}}_{scale}\right]+\underbrace{[\eta_c]}_{coupl.}+\underbrace{[d\theta_c]}_{align.}\right)\cdot\vec{\Gamma}_{app,d/sat} \\ &+\left(\left[\underbrace{dK_{1d}}_{scale}\right]+\underbrace{[\eta_d]}_{coupl.}\right)\cdot\vec{b}_{1c} \\ &+\left(\left[\underbrace{dK_{1d}}_{scale}\right]+\underbrace{[\eta_d]}_{coupl.}+\underbrace{[\theta_d]}_{align.}\right)\cdot\left(\vec{\Gamma}_{app,c/sat}+\frac{\vec{F}ext_{/sat}}{M_{Isat}}+\frac{\vec{F}th_{/sat}}{M_{Isat}}\right) \\ &+\frac{1}{2}\underbrace{K_{21}}_{quad}\Gamma_{app,1/sat}{}^2-\frac{1}{2}\underbrace{K_{22}}_{quad}\Gamma_{app,2/sat}{}^2+O\left(dK,d\eta,d\theta,K_2\right)^2\end{aligned}$$

The index d represents the differential mode (the difference between the measured or true accelerations or the mismatching of the sensitivities, scale factors, alignments, etc.); on the other hand, the index c corresponds to the common mode (mean value of the acceleration of the two concentric test masses, or of the scale factors, etc.), $\vec{b}$ is the acceleration bias (due to the mechanics or the electronics defects: these can be affected or not by the lack of knowledge of the scale factor K_1 or the couplings η). Θ represents the misalignment of the test-masses. The nonlinear terms or quadratic terms are noted K_2.

The Eötvös parameter is included in the applied acceleration difference:

$$\Gamma_{app,d/sat} = \frac{1}{2}\delta\vec{g}_{/sat}+\frac{1}{2}\left(T_{/sat}-In_{/sat}\right)\cdot\vec{\Delta}_{/sat}-\frac{1}{2}\left(2[\Omega]_{/sat}\dot{\vec{\Delta}}_{/sat}+\ddot{\vec{\Delta}}_{/sat}\right)$$

This equation shows the effect of the miss-centering Δ of the two test-masses, which are supposed to be concentric: it depends on the local gravity gradient T, or the attitude motion of the satellite In (angular acceleration and centrifugal acceleration).

Finally, we have to consider the possible *a posteriori* corrections of the measurement. Indeed, some of the instrument parameters can be evaluated in orbit by specific motions or stimuli of the satellite to enhance the effect of the defect. Once the parameters are calibrated, it is possible to subtract the contribution of the disturbing effect.

The second step in establishing the requirements consists in distributing the errors on the payload (measurement noise, scale factors, bias fluctuations, etc.) and on the satellite (thruster noise for the drag-free satellite and the attitude control, star sensor noise, position error, etc.). These errors have an impact either directly on the measured acceleration or on the *a posteriori* correction of the measurement.

The frequency distribution of the error has also been considered. Given that the measurement is performed at a well-known frequency fep, the stochastic noise can be partially rejected with the increase in the time integration Ti. The expression of the error budget is summarized as follows:

$$\left(\sqrt{\sum_{n_{ep}}S_{fep}^2}+\sqrt{\sum_{n'_{f'}}\left(\frac{S(f')}{R(f')}\right)^2}+\sqrt{\sum_{n''_i}\left(\frac{S(i\cdot fep)}{R(i\cdot fep)}\right)^2}\right)\oplus\sqrt{\sum_{n_r}\frac{S_r^2}{Ti}}<8.10^{-15}\,ms^{-2}$$

S_{fep} depicts each error source in phase at the EP frequency, which is the frequency of the Earth's gravitational field, modulated by the projection on the instrument axes and evaluated as 8 m/s^2 at 710 km altitude. $S(i.fep)$ are the error sources at the harmonic frequencies of fep, they are eliminated by a ratio $R(i.fep)$ depending on the harmonic. These ratios have been evaluated considering a single source at each harmonic frequency. $S(f')$ are the tone errors at frequencies different from multiples of the fep and eliminated by the ratio $R(f')$. All tone errors have been estimated and added together. Finally, the stochastic error S_r is eliminated by the integration time and summed quadratically with the tone errors.

The error distribution contains more than 80 sources for sine wave errors and 80 for stochastic sources. These are evaluated and validated during the manufacturing of the instrument, because they are deduced from one of the major functions of the instrument [81] or require specific experimentation to confirm the driven physical parameters [82], [83]. This error budget is derived to express more than 150 satellite and instrument major mission requirements, including the DC value of diving parameters that could be combined with the fep variations to generate a disturbance.

While the satellites were launched successively on April 25, the instrument was quickly switched on, on May 2nd, and the four masses were automatically electrostatically levitated, with the inertial sensors thereby providing their data. As from this date, the satellite quickly entered into a partial eclipse period that does not allow scientific measurements to be made, but only the operation assessment of the satellite and the payload. This is what is fruitfully being done. In order to achieve the MICROSCOPE mission objective, not only must the eventual violation measurement be performed, but also the Eötvös parameter accuracy obtained must be demonstrated.

Atomic interferometer

A new generation of complementary instruments, relying on the manipulation of matter waves through atom interferometry, appears nowadays as very promising for highly precise and accurate inertial measurements. Cold Atom interferometers have indeed proven on the ground to be very high performance sensors, with the development in recent decades of cold atom gravimeters [84], gravity gradiometers [85] and gyroscopes [86]. This promising technology has demonstrated performances that compete with other state-of-the-art gravimeters (superconducting or mass-spring devices) and is only expected to reach its full potential in space based applications. In such a micro-gravity environment, the interrogation time, and therefore the measurement scale factor, can be increased by orders of magnitude compared to ground-based sensors.

In a cold atom inertial sensor, the test mass is a gas of cold atoms obtained by laser cooling and trapping techniques. This cloud of cold

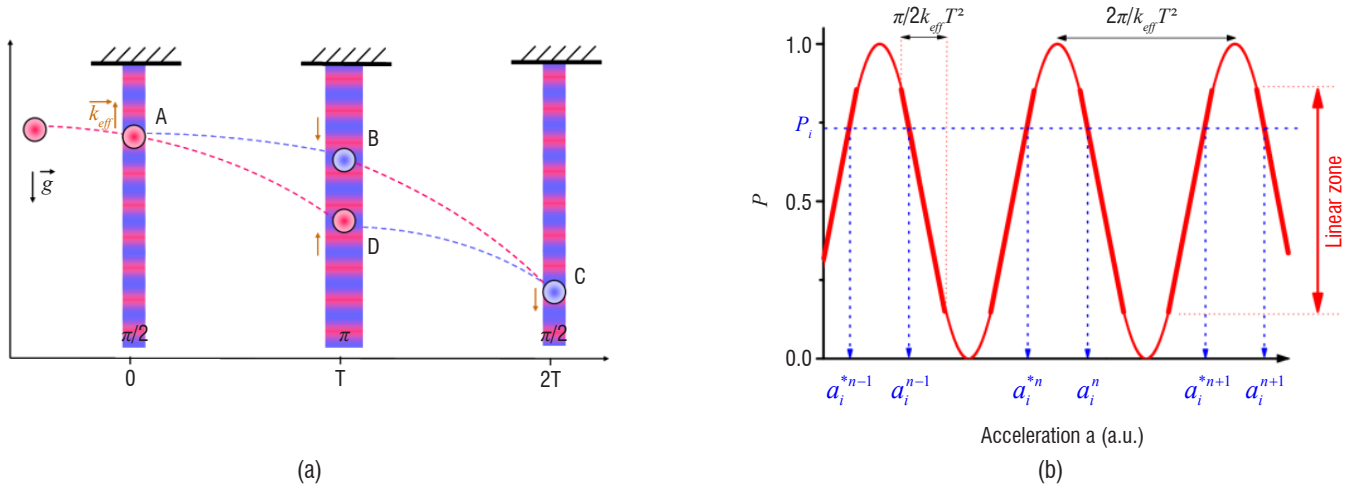


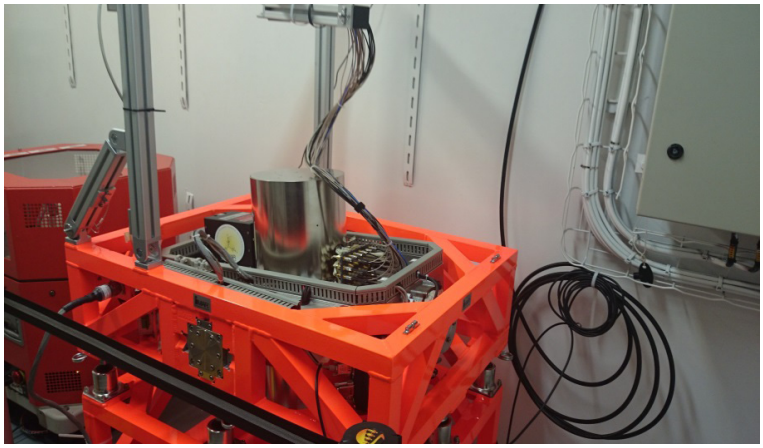
Figure 9 – (a) Mach-Zehnder type atom interferometer for gravity acceleration measurement. For such a gravimeter configuration, the three light pulses are separated only in time and not in space. (b) Output signal P_i of a cold atom accelerometer. One measurement P_i corresponds to several possible acceleration values a_i^n , a_i^{*n}.

atoms is released from a trap and, during free fall in a vacuum chamber, its displacement due to inertial forces is measured by an atom interferometry technique. Typically, an atom Mach-Zehnder type interferometer is made, consisting in a sequence of three equally spaced laser pulses allowing the atomic wave to be either equally split, deflected or re-combined (see Figure 9-a). The output signal from the instrument P depends sinusoidally on the phase of the interferometer (see Figure 9-b), which is proportional to the acceleration $\vec{a}$ of the atoms along the laser direction of propagation $\vec{k}_{eff}$: $P = \frac{1}{2} - \frac{1}{2}\cos(\vec{k}_{eff} \cdot \vec{a}\, T^2)$, where $\vec{k}_{eff}$ is the effective wave vector associated with the laser and T is the time between laser pulses. In the case where the acceleration to be measured is not subjected to large variations ($\Delta a << \frac{\pi}{2 k_{eff} T^2}$), it is possible thanks to a specific technique to identify the fringe index corresponding to the measurement and to retrieve the true acceleration value unambiguously. Otherwise, for large shot-to-shot acceleration variations ($\Delta a > \frac{\pi}{2k_{eff}T^2}$), the true acceleration value can no longer be recovered. Note that the measurement rate is typically of a few Hz for this kind of instrument. This ambiguity concerning the acceleration determination consequently reduces the shot-to-shot measurement range of cold atom sensors. With regard to this limitation, it is typical to associate a mechanical accelerometer to the cold atom instrument. This mechanical accelerometer allows the identification of the fringe index corresponding to the atomic acceleration measurement [87].

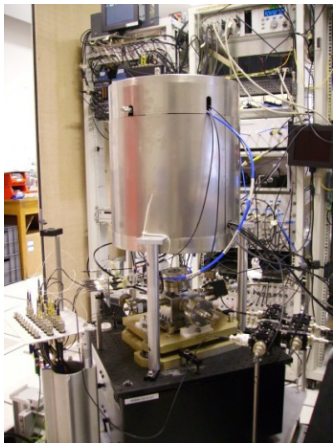
Following these measurement principles, cold atom interferometers have demonstrated on the ground performances comparable to those of other existing state-of-the-art technologies, especially concerning the development of gravimeters. At the present time, laboratory cold atom gravimeter performances have already exceed those of the conventional corner cube gravimeter, especially in terms of sensitivity, reaching 4.2 μGal/Hz$^{1/2}$ [88] (1 μGal = 10^{-8} m/s^2) and an accuracy of 5 μGal [89]. For most gravimeters, the sensitivity remains limited by the vibration noise background. Compared to conventional corner cube gravimeters, cold atom gravimeters can achieve a higher repetition rate, up to 300 Hz [90], and do not have movable mechanical parts. These features make cold atom gravimeters more suitable for on-board applications. In this context, in 2009 ONERA emphasized the on-board potential of this new generation of instruments, through the development of the GIRAFE cold atom gravimeter, designed especially for on-board applications. Despite the compactness of this instrument, the gravimeter has demonstrated a sensitivity of 42 μGal/Hz$^{1/2}$ and an accuracy of 25 μGal, close to that of state-of-the-art gravimeters [91]. Gravity measurements in an elevator were also conducted, leading to the determination of the Earth's gravity gradient along the vertical direction with a precision of 4 E (1 E = 0.1 μGal/m), and to first measurements of an atom gravimeter on a mobile platform.

Recently, ONERA has continued its research efforts in the field of on-board applications by developing a second generation cold atom gravimeter, GIRAFE 2, dedicated to boat gravity measurements (see Figure 10-a). In October 2015 and January 2016, ONERA in collaboration with SHOM tested and characterized this gravimeter when integrated on a gyro-stabilized platform during two marine campaigns, demonstrating successful results in terms of gravitational field mapping, even exceeding those obtained with the previous existing mechanical technology [92].

In the context of future Earth's gravity measurements in space, ONERA is also involved in a preliminary ESA study aimed at assessing the potential of combining the electrostatic technology and atom interferometry. These two technologies are clearly identified as very good candidates for future spatial missions dedicated to Earth observation. Each of these two types of instruments have their own assets which are, for electrostatic sensors, their demonstrated short term sensitivity and their maturity regarding the space environment and, for Atom Interferometers among others, the absolute nature of the measurement and therefore the lack of need for calibration processes. These two technologies seem in some aspects to be very complementary, and a hybrid sensor bringing together all of their assets could be the opportunity to take a big step in this context. ONERA, which has developed an expertise in each of these two technologies, is initiating a first experimental demonstration of a hybrid electrostatic-atomic instrument (see Figure 10-b) to begin the exploration of such an original sensor and to study its full potential for future gravity missions or aircraft applications.



(a)



(b)

Figure 10 – (a) The cold atom gravimeter GIRAFE 2 with its gyrostabilized platform installed in the gravimetry room of the French ship Beautemps-Beaupré owned by SHOM. (b) The cold atom gravimeter GIRAFE coupled to an electrostatic accelerometer.

Mission and sensor perspectives

Concerning fundamental physics in space, it is clear that the MICROSCOPE mission must be followed by another similar space experiment with the same type of mission. If no EP violation is confirmed at the level of 10^{-15}, going beyond this accuracy must be sought and a cryogenic mission would help to gain another two orders of magnitude [93]. If a violation of the Equivalence Principle is detected, the violation must be confirmed and to analyzed, in particular according to different pairs of mass compositions. The same technology and concept can also be used for the detection of gravity waves in space as an inertial sensor. This is what is currently being demonstrated with promising results by the Lisa-Pathfinder payload of the eponym satellite for the e-Lisa mission [66], [94].

A reduced power and size sensor has been developed over the last years in our laboratory devoted to planetology. Interplanetary spacecraft require the mass to be reduced and the necessary power. This stand-alone instrument presents a cubic proof-mass, a volume of about 1 liter for 1 Watt consumption and its offsets along two of its sensitive axes are estimated and corrected in orbit. It can be used to map the gravitational field of celestial bodies, to measure residual atmosphere density or winds through satellite drag, or to perform long range Newton law tests during the transfer. This GAP sensor has been proposed to ESA in previous calls [95], [96].

Obviously, the common history of space geodesy and space accelerometers is not finished. Electrostatic accelerometers in space have been the corner stones of drastic improvements in the knowledge of the Earth's gravitational field: the GOCE mission (2009-2013) brought a strong improvement of the static part of the gravitational field, both in terms of space resolution (about 100 km) and in terms of accuracy, whereas since 2002 the GRACE mission has allowed the temporal evolution of gravity to be monitored, typically with a set of Stokes coefficients every month at a space resolution of about 300 km. These two kinds of concept will drive the needs of inertial sensors in space geodesy in the short and medium term. The results of the GRACE mission are so rich that there is a definite consensus from the community to continue this kind of monitoring with the least possible interruptions; following the example of space altimetry, the objective is to reconcile the robustness and the improvements by means of progressive evolutions. For example, GRACE Follow On (GFO), the successor of GRACE, will embark a new system based on laser interferometry to measure the relative velocity between the 2 craft, but this will be redundant with the microwave link already used on GRACE. GFO will use an accelerometer that is practically to the same as the GRACE instrument and with the same performances, at the 10^{-10} $ms^{-2}Hz^{-1/2}$ level. For the next step, when the reliability of the laser link is proved with an expected accuracy of 50 nanometers, the aim would be to have an accelerometer compatible with this level of performances, i.e., at the 10^{-11} $ms^{-2}Hz^{-1/2}$ level. This is the heart of the e.motion^2 mission proposed to ESA within the framework of the Explorer-9 program. There are also high hopes for cold atom accelerometers, but we do not currently know the exact limits of these sensors. If they prove to be competitive, a first step could be a mission embarking cold atom accelerometers together with electrostatic accelerometers.

The GOCE mission could also have a successor, but the context will not be the same as for GRACE. Continuity is not a main goal and the interest in a new gradiometric mission would be based on a substantial improvement of the performances. The need would be typically to go from the 10^{-12} $ms^{-2}Hz^{-1/2}$ level achieved for GOCE to 10^{-13} $ms^{-2}Hz^{-1/2}$. Since geodesists do not lack creativity, some of them have even thought about a hybrid concept associating gradiometric and SST techniques; this would involve 2 satellites with an inter-satellite link, with each satellite (or at least one of them) including one or several gradiometric arms.

However space accelerometers are not only useful to study the gravitational field. They are also very helpful to measure non-gravitational forces acting on the spacecraft. This has both a general and a specific interest. The general interest is to improve the quality of the dynamical model used to compute the trajectory, since gravitational forces are very tricky to compute; for example, errors of some tens of per cents are not rare in the prediction of the drag force due to the atmosphere. Thus, future geodetic missions such as the GRASP concept, which is aimed at the sub-centimeter level in terms of positioning, would be hardly achievable without an accelerometer [97]. The specific interest is to be able to perform in situ observations of the non-gravitational forces, in order to obtain information about their sources; this is the case of the atmosphere densities (the source of the drag force), which have been better modeled using the measurements of the STAR accelerometer on board the CHAMP satellite (2000-2010); this is also the case of radiative flux from the Earth (one of the sources of radiation pressure forces), which motivates mission concepts such as BIRAMIS, which is aimed at monitoring the Infra-Red and albedo radiations by means of accelerometric measurements [98]. This concept could be an option studied for the GRASP mission mentioned above.

Conclusion

The electrostatic levitation of a solid test mass is an old technology, but nevertheless a fruitful concept to design an ultra-sensitive space accelerometer. Squid detectors and superconductive magnetic levitation of the test mass have been proposed recently, but seem to have been forgotten because of their complexity for a space technology. In electrostatic sensors, the test mass motion is not only controlled with regard to the position but also with regard to the attitude, giving information on the angular acceleration of the satellite on which it is integrated. This the case of the GOCE satellite and would also be the case of the MICROSCOPE satellite, for which a fine attitude stability is required. The optimization of the accelerometer design requires mechanics and electronics experts, as well as physicists to deal with the disturbing forces acting on the test mass: this is an integrated team that Onera can offer for the described missions. The actual configuration of the sensor depends on the mission, the required full-scale range and the required resolution in a specific bandwidth. In order to improve the latter, it is now necessary to manage in orbit the test mass charging by the radiation flux through a solution other than the thin gold wire used and its disturbance damping: a photoelectric device could be the solution; it has already been implemented in the Lisa-Pathfinder inertial device configuration [99],[100],[101]. A gain by two orders of magnitude can be expected. The association with an atomic interferometer is also an avenue for the future, by taking advantage of the outstanding stability of the latter at low frequencies and the resolution of the electrostatic sensor at higher frequencies. What seemed to be old technology at the end of the last century now appears to be the key for future space missions at the beginning of the new century ■

Acknowledgment

The authors would like to thank all space team members who have contributed to the success of the mentioned space missions and, in particular, the MICROSCOPE mission. The development of the sensors has been party supported by ONERA and CNES.

References

[1] A. EINSTEIN - *The Meaning of General Relativity*. Princeton University Press, New Jersey, 5th edition, 1956.

[2] M. E. PESKIN - *Theoretical Summary Lecture for the Higgs Huning 2012*. arXiv:1208.5152v2 [hep-ph.] 8 Sep 2012.

[3] B. P. ABBOTT *et al.* - *Observation of Gravitational Waves from Binary Black Hole Merger*. PRL 116, 061102, 2016.

[4] CHAMP - http://op.gfz-potsdam.de/champ/

[5] GRACE - http://science.nasa.gov/missions/grace/

[6] GOCE - http://www.esa.int/Our_Activities/Observing_the_Earth/GOCE/

[7] A. LAWRENCE - "*Modern Inertial Technology: Navigation, Guidance and Control*". Springer, second edition, 1998.

[8] R. E. HOPKINS, F. K. MUELLER, W. HAEUSSERMANN - *The Pendulous Integrating Gyroscope Accelerometer (PIGA) from the V-2 to Trident D5, the Strategic Instrument of Choice*. Navigation, and Control Conference & Exhibit, 6-9 August 2001 Montreal, Canada, 2001.

[9] S.A. FOOTE, D. B. GRINDELAND - *Model QA3000 Q-Flex Accelerometer High Performance Test Results*. IEEE Aerospace and Electronic Systems Magazine, 1992, Volume: 7, Issue: 6, p. 59-67, DOI: 10.1109/62.145120.

[10] H. C. HAYES - *Method and Apparatus for Determining the Force of Gravity*. US Patent n° 1,995,305 filed Oct. 1928, 1935.

[11] G. MAMON - *A Traverse Gravimeter for the Lunar Surface*. IEEE Transactions on Geoscience Electronics, Vol. GE-10, n° 1, p. 64-72, 1972.

[12] J. KRITZ - *Accelerometer*. US Patent n° 2,984,111, filed June 1959, 1961.

[13] J. KRITZ *et al.* - *Accelerometer and Parts Therefor*. US Patent n° 3,190,129, filed July 1961, 1965.

[14] F.A NORRIS *et al.* - *Piezoelectric Force Transducer*. US Patent n° 3,479,536, filed March 1967, 1969.

[15] L. WEISBORD - *Single Beam Force Transducer with Integral Mounting Isolation*. US Patent n° 3,470,400, filed Dec. 1967, 1969.

[16] J. M. PAROS *et al.* - *Longitudinal Isolation System for Flexurally Vibrating Force Transducers*. US Patent n° 4,321,500, filed Dec. 1979, 1982.

[17] W. C. ALBERT *et al.* - *Non-Prismal Beam Resonator*. US Patent n° 4,445,065, filed Sep. 1981, 1984.

[18] W. C. ALBERT *et al.* - *Vibrating Beam Force Transducer with Angled Isolator Springs*, US Patent n° 4,658,174, filed Mar. 1986, 1987.

[19] H. FIMA & D. JANIAUD - *Vibrating String Resonator*. US Patent n° 4,710,668, filed Dec. 1985, 1987.

[20] E. NISSE - *Miniature Quartz Resonator Force Transducer*. US Patent n° 4,215,570, filed Apr. 1979, 1980.

[21] R. G. KIRMAN *et al.* - *Force sensors*. US Patent n° 4,594,898, filed Jun. 1984, 1986.

[22] B. L. NORLING *et al.* - *Accelerometer with Isolator for Common Mode Inputs*. US Patent n° 4,766,768, filed Oct. 1987, published Aug. 1988.

[23] https://aero1.honeywell.com/inertsensor/rba500.shtml

[24] K.E. PETERSEN - *Silicon as Mechanical Material*. Proceeding of the IEEE, Volume:70, Issue:5, Pages:420-457, 1982.

[25] R. T. HOWE *et al.* - *Resonant Accelerometer*. US Patent n° 4,805,456, filed May 1987, 1989.

[26] T. V. ROSZHART - "*Silicon Micromachined Accelerometer*", US Patent n° 4,945,765, filed Aug. 1988, published Aug. 1990.

[27] D. JANIAUD *et al.* - *Accelerometric Sensor with Flectional Vibratory Beam*. US Patent n° 5,170,665, filed Aug. 1990, 1992.

[28] W. C. ALBERT - *Monolythic Resonator Vibrating Beam Accelerometer*. US Patent n° 4,804,875, filed Sep. 1982, 1989

[29] O. LE TRAON *et al.* - *The VIA Vibrating Beam Accelerometer: Concept and Performances*. Proceedings of the Position, Location and Navigation Symposium, Palm springs, 1998.

[30] G.R. NEWELL *et al.* - *Accelerometer and Method of Manufacture*. US Patent n° 5,5755,978, filed Dec. 1996, 1998.

[31] R. JAFFE, T. ASHTON, A. MADNI - *Advances in Ruggedized Quartz MEMS Inertial Measurement Units*. IEEE/ION PLANS, Coronado, CA, April 2006.

[32] T. V. ROSZHART, H. JERMAN, J. DRAKE, C. DE COTIIS - *An Inertial-Grade Micormachined Vibrating Beam Accelerometer*. Transducers'95 & EuroSensors IX, 1995.

[33] R. LEONARDSON, S. FOOTE - *SiMMA Accelerometer for Inertial Guidance and Navigation*. Position Location and Navigation symposium IEEE, p. 152-160, Palm Spring, USA, 1998.

[34] M. HELSEL, G. GASSNER, M. ROBINSON, J. WOODRUFF - *A Navigation Grade Micro-Machined Silicon Accelerometer*. Position Location and Navigation Symposium, 1994, IEEE 11-15, p. 51-58, 1994.

[35] R. HOPKINS *et al.* - *The Silicon Oscillating Accelerometer: A High-Performance MEMS Accelerometer for Precision Navigation and Strategis Guidance Applications*. ION 61st Annual Meeting, Cambridge, MA, June 2005.

[36] S. F. BECKA, M. NOVACK, S. SLIVINSKY, C.PAUL - *A High Reliability Solid State Accelerometer for Balistic Missile Inertial Guidance*. AIAA Guidance, Navigation and Control Conference, Aug. 2008.

[37] O. LEFORT, S. JAUD, R. QUER, A. MILESI - *Inertial Grade Silicon Vibrating Beam Accelerometer*. Symposium Gyro Technology, Sept. 2012.

[38] O. LE TRAON, D. JANIAUD, B. LECORRE, S. MULLER, M. PERNICE - *Mechanical Decoupling Device for Monolithic Differential Vibrating Sensor*. US patent n° 7448268 B, 2008

[39] J. LECLERC - *MEMS for Aerospace Navigation*. Aerospace and Electronic Systems Magazine , IEEE, Volume: 22, Issue: 10, p. 31-36, 2007.

[40] A. JEANROY, P. FEATONBY, J-M. CARON - *Low-Cost Miniature and Accurate IMU with Vibrating Sensors for Tactical Applications*. Symposium Gyro Technology 2003, Stuttgart.

[41] T. LORET, G. HARDY, C. VALLÉE, V. DEMUTRECY, T. KERRIEN, S. COCHAIN, D. BOUTOILLE, R. TAÏBI, R. BLONDEAU - 2014 DGON Inertial Sensors and Systems (ISS).

[42] O. LE TRAON, D. JANIAUD, S. MULLER - *Monolithic acceleration transduce*. US patent n° 5962786 A, 1999.

[43] R. LEVY, O. LE TRAON, S. MASSON, O. DUCLOUX, D. JANIAUD, J. GUÉRARD, V. GAUDINEAU, C. CHARTIER - *An Integrated Resonator-Based Thermal Compensation for Vibrating Beam Accelerometers*. Sensors, IEEE 978-1-4877-1787, 2012.

[44] R. LEVY, D. JANIAUD, J. GUERARD, R. TAÏBI, O. LE TRAON - *A 50 Nano-g Resolution Quartz Vibrating Beam Accelerometer*. 2014 International Symposium on Inertial Sensors and Systems (ISISS).

[45] A. PEREZ, A. M. SHKEL - *Design and Demonstration of a Bulk Micromachined Fabry–Pérot μg – Resolution Accelerometer*.

[46] C.-H. LIU, T. W. KENNY - *A High-Precision, Wide-Bandwidth Micromachined Tunneling Accelerometer*. 1057-7157, IEEE, 2001.

[47] J. LI *et al.* - *On-Orbit Performance of Gravity Probe B Drag-Free Translation Control and Orbit Determination*. Advances in Space Research, 40, 1-10, 2007.

[48] P. TOUBOUL, G. MÉTRIS, V. LEBAT, A. ROBERT - *The MICROSCOPE Experiment, Ready for the In-Orbit Test of the Equivalence Principle*. Class. Quantum Grav. 29, 18, 2012.

[49] E. DOORNBOS *et al.* - *Neutral density and crosswind determination from arbitrarily multiaxis accelerometers on satellite*. Journal of spacecraft and rockets, 47, 4, 2001.

[50] B. D. TAPLEY *et al.* - *GRACE Measurements of Mass Variability in the Earth System*. 305, 5883, p. 503-505, 2004.

[51] E. CANUTO, L. MASSOTTI - *All-Propulsion Design of the Drag-Free and Attitude Control of the European Satellite GOCE*. Acta astronautica, 64,2-3, p. 325-344, 2009.

[52] H. NYQUIST - *Thermal Agitation of Electric Charge in Conductors*. Phys. Rev. 32,110,1928.

[53] C.W. MC COMBIE - *Flucutation Theory in Physical Measurements*. Reports on Progress In Phys., 16, 266, 1953.

[54] F. GRASSIA, J.-M. COURTY, S. REYNAUD, P. TOUBOUL - *Quantum Theory of Fluctuations in a Cold Damped Accelerometer*. The European Phys. Journal D 8, 101-110, 2000.

[55] S. T. RUGGIERO, D. A. RUDMAN - *Superconducting Devices*. Academic Press Inc., 1990.

[56] R KLEINER, D. KOELLE, F. LUDWIG, J. CLARKE - *Superconducting Quantum Interference Devices: State of the Art and Application*. IEEE 92, 10, 2004.

[57] J. CLARKE, A. I. BRAGINSKI - *The SQUID Handbook*. Wiley-VCH Verlag mbH & Co.LGaA, 2006.

[58] H. J. PAIK, J-P. BLASER, S. VITALE - *Principle of the STEP Accelerometer Design*. Adv. Space Res., 32, 7, p. 1325-1333, 2003.

[59] J. OVERDUIN, F. EVERITT, P. WORDEN, J. MESTER - *STEP and Fundamental Physics*. Class Quantum Grav. 29, 18, 2012.

[60] H. A. CHAN, H. Y. PAIK - *Superconducting Gravity Gradiometer for Sensitive Measurements*. Physical Review. I Theory, 35, 12, 1987.

[61] H. A. CHAN, M. V. MOODY, H. Y. PAIK - *Superconducting Gravity Gradiometer for Sensitive Measurements.* II Experiment, Physical Review, 35, 12, 1987.

[62] http://www.nasa.gov/mission_pages/gpb/

[63] S. BUCHMAN, J. P. TURNEAURE - *The Effects of Patch-Potentials on The Gravity Probe B Gyroscopes*. Rev. Sci; Instrument, 82, 074502, 2011.

[64] P. TOUBOUL, B. FOULON, E. WILLIEMENOT - *Electrostatic Space Accelerometers for Present and Future Missions*. Acta Astrautica, 45, 10, p 605-617, 1999.

[65] P. TOUBOUL, B. FOULON, M. RODRIGUES, J.-P. MARQUE - *In Orbit Nano-g Measurements, Lessons for Future Space Missions.* Aerospace Science and Technology 8, p 431-441, 2004.

[66] http://sci.esa.int/lisa-pathfinder/

[67] P. MCNAMARA, S. VITAL, K. DANZMANN - *LISA Pathfinder*. Class. Quantum Grav., 25, 114034, 2008.

[68] V. JOSSELIN, M. RODRIGUES, P. TOUBOUL - *Inertial Sensor Concept for the Gravity Wave Missions*. Acta Astronautica 49, 2, p.95-103, 2001.

[69] E. WILLEMENOT, P. TOUBOUL - *On-Ground Investigation of Space Accelerometers Noise with an Electrostatic Torsion Pendulum*. Rev. Sci. Instrum. 71, 302 (2000); doi: 10.1063/1.1150197.

[70] A. CAVALLERI *et al.* - *Direct Force Measurements for Testing the Lisa Pathfinder Gravitational Reference Sensor*. Class. Quantum Grav. 26 (2009) 094012.

[71] J.-P. MARQUE, B. CHRISTOPHE, B. FOULON - *Accelerometers of the GOCE Mission: Return of Experience from One Year of In-Orbit*. Proc. 'ESA Living Planet Symposium', Bergen, Norway 28 June - 2 July 2010 (ESA SP-686, December 2010).

[72] B. CHRISTOPHE, D. BOULANGER, B. FOULON, P.-A. HUYNH, V. LEBAT, F. LIORZOU, E. PERROT - *A New Generation of Ultra-Sensitive Electrostatic Accelerometers for GRACE Follow-On and Towards the Next Generation Gravity Missions*. Acta Astronautica 117 (2015) 1-7.

[73] J. FLURY, S. BETTADPUR, B. D. TAPLEY - *Precise Accelerometry Onboard the GRACE Gravity Field Satellite Mission*. Advances in Space Research 42 (2008) 1414-1423.

[74] CHRISTOPHE, B., J. MARQUE, B. FOULON - *In-Orbit Data Verification of the Accelerometers of the ESA GOCE Mission*. SF2A-2010: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics. Vol. 1, 2010.

[75] G. RUSSANO - *A Torsion Pendulum Ground Test of the LISA Pathfinder Free-Fall Mode.* Doctoral Thesis, University of Trento, 2015.

[76] M. BASSAN *et al.* - *Approaching Free Fall on Two Degrees of Freedom: Simultaneous Measurement of Residual Force and Torque on a Double Torsion Pendulum*. PRL 116, 051104, 2016.

[77] Drop Tower Operation and Service Company - *ZARM Drop Tower Bremen User Manual.* ZARM FABmbH, Bremen, 2012.

[78] H. SELIG *et al.* - *Drop Tower Microgravity Improvement Towards the Nano-g Level for the MICROSCOPE Payload Tests*. Microgravity Sci. Technol. 22, 2010.

[79] A. EINSTEIN - *Die Grundlage der allgemeinen Relativitätstheorie*. Annalen der Physik 354 (7), 769-822, 1916.

[80] P. TOUBOUL - *The MICROSCOPE Mission and its Uncertainty Analysis*. Space Sci Rev 148: 455-474, 2009.

[81] V. JOSSELIN, P. TOUBOUL, R. KIELBASA - *Capacitive Position Sensing for Space Accelerometers Applications*. Sensors and Actuator, 78, 2-3, p 92-98, 1999.

[82] E. WILLEMENOT, P. TOUBOUL - *Electrostatically Suspended Torsion Pendulum*. Rev. Scien. Instr. 71, 310, 2000.

[83] E. WILLEMENOT, P. TOUBOUL - *On-Ground Investigation of Space Accelerometers Noise with an Electrostatic Torsion Pendulum*. Rev. Scien. Instr. 71, 302, 2000.

[84] A. PETERS, K. Y. CHUNG, S. CHU - *High-Precision Gravity Measurements Using Atom Interferometry.* Metrologia 38, 25, 2001.

[85] J. M. MCGUIRK, G. T. FOSTER, J. B. FIXLER, M. J. SNADDEN, M. A. KASEVICH - *Sensitive Absolute-Gravity Gradiometry Using Atom Interferometry*. Phys. Rev. A 65, 033608, 2002.

[86] T. L. GUSTAVSON, P. BOUYER, M. A. KASEVICH - *Precision Rotation Measurements with an Atom Interferometer Gyroscope*. Phys. Rev. Lett. 78, 2046, 1997.

[87] R. GEIGER, V. MÉNORET, G. STERN, N. ZAHZAM, P. CHEINET, B. BATTELIER, A. VILLING, F. MORON, M. LOURS, Y. BIDEL, A. BRESSON, A. LANDRAGIN, P. BOUYER - *Detecting Inertial Effects with Airborne Matter-Wave Interferometry*. Nature Communications 2, 474, 2011.

[88] Z.-K. HU, B.-L. SUN, X.-C. DUAN, M.-K. ZHOU, L.-L. CHEN, S. ZHAN, Q.-Z. ZHANG, J. LUO - *Demonstration of an Ultrahigh-Sensitivity Atom-Interferometry Absolute Gravimeter.* Phys. Rev. A 88, 043610, 2013.

[89] A. LOUCHET-CHAUVET, T. FARAH, Q. BODART, A. CLAIRON, A. LANDRAGIN, S. MERLET, F. PEREIRA DOS SANTOS - *The Influence of Transverse Motion within an Atomic Gravimeter*. New J. Phys. 13, 065025, 2011.

[90] H. J. MCGUINNESS, A. V. RAKHOLIA, G. W. BIEDERMANN - *High Data-Rate Atom Interferometer for Measuring Acceleration*. Appl. Phys. Lett. 100, 011106, 2012.

[91] Y. BIDEL, O. CARRAZ, R. CHARRIÈRE, M. CADORET, N. ZAHZAM, A. BRESSON - *Compact Cold Atom Gravimeter for Field Applications*. Appl. Phys. Lett. 102 , 144107, 2013.

[92] http://www.onera.fr/fr/actualites/cartographie-marine-haute-precision-pesanteur-atomes-froids

[93] P. TOUBOUL, M. RODRIGUES, E. WILLEMENOT, A. BERNARD - *Electrostatic Accelerometers for the Equivalence Principle Test in Space*. Class. Quantum Grav. 13, A67-A78, 1996.

[94] https://www.elisascience.org/

[95] B. CHRISTOPHE *et al.* - *Odyssey: A Solar System Mission*. Exp Astron 23:529-547, 2009.

[96] B. CHRISTOPHE *et al.* - *OSS (Outer Solar System): A Fundamental and Planetary Physics Mission to Neptune, Triton and the Kuiper Belt*. Exp Astron 34:203-242, 2012.

[97] Y. BAR-SEVER, B. HAINES, W. BERTIGER, S. DESAI, S. WU - *Geodetic Reference Antenna in Space (GRASP) – A Mission to Enhance Space-Based Geodesy*. ilrs.gsfc.nasa.gov/docs/GRASP_COSPAR_paper.pdf

[98] G. DUCHOSSOIS - *Potential European Climatological Satellite Missions: SEOCS and BIRAMIS*. Acta Astronautica Vol 7, 385-399, 1980.

[99] T. SUMNER *et al.* - *LISA and LISA Pathfinder Charging* . Class. and Quant. Gravity, 26,094006, 2009.

[100] D. HOLLINGTON – thesis: *The Charge Management System for LISA and LISA Pathfinder*. High Energy Physics Group, Department of Physics, Imperial College London, 2011.

[101] D. HOLLINGTON, J. T. BAIRD, T. J. SUMNER, P. J. WASS - *Characterising and Testing Deep UV LEDs for Use in Space Applications*, Class. and Quant. Gravity vol.32, N°23, 2015.

Acronyms

AVAS	(Accéléromètre Vibrant pour Applications Spatiales)
CERN	(Conseil Européen pour la Recherche Nucléaire)
CHAMP	(CHAllenging Minisatellite Payload)
eLISA	(European Laser Interferometer Space Antenna)
ESA	(European Space Agency)
FSR	(Full Scale Range)
GAP	(Gravity Accelerometer Package)
GPB	(Gravity Probe B)
GOCE	(Gravity field and steady-state Ocean Circulation Explorer)
GRACE	(Gravity Recovery And Climate Experiment)
GRASP	(Geodetic Reference Antenna in SPace)
MEMS	(Micro-Electro-Mechanical System)
MICROSCOPE	(MICRO- Satellite pour l'Observation du Principe d'Equivalence)
NASA	(National Aeronautics and Space Administration)
PIGA	(Pendulous Integrating Gyroscope Accelerometer)

SQUID	(Superconducting Quantum Interference Device)
SST	(Standard Star Tracker)
STAR	(Space Three-Axis Accelerometer for Research)
VBA	(Vibrating Beam Accelerometer)
VSA	(Vibrating String Accelerometer)
ZARM	(Zentrum für Angewandte Raumfahrttechnologie und Mikrogravitation)

AUTHORS

Pierre Touboul is since 2010 Scientific Director of the Physics Branch at the French Aerospace Lab, ONERA. He is recipient in 2004 of the Award in Physics, "Prix Montgolfier" and of the "Grand Prix Marcel Dassault" from French Science Academy. He awards in 2010 the annual "Grand Prix" from French Air and Space Academy for the success of the space accelerometers on board the international CHAMP, GRACE and GOCE missions. His major scientific contribution, as instrumentalist, concerns space accelerometry, geodesy and fundamental physics space missions, with major contributions to the scientific payloads of already four successful missions. He is presently Prime Investigator of the Cnes-ESA MICOSCOPE mission, co-author of patents, more than seventy peer-reviewed publications, over one hundred Onera reports and international conference papers.

Gilles Metris is Astronomer at laboratory Géoazur of Nice University, Côte d'Azur Observatory and CNRS. Its themes of research concern celestial mechanics including modeling orbital dynamics of spacecraft and applications to the knowledge of environment in which they move: gravity field of the Earth, small non-gravitationnal forces, tests of gravitation in weak field… Gilles Metris is also Co-PI of the MICROSCOPE mission.

Hanns Sélig. After his degree in physics and astrophysics at the University of Hamburg and some work at the Observatory Hamburg-Bergedorf and DESY, he moved to ZARM (University of Bremen) where he started to work on MICROSCOPE in 2001.

Olivier Le Traon obtained an Engineering degree from INSA Lyon (*Institut National des Sciences Appliquées*) in 1986 and a master degree in Solid Mechanics and Modeling. He joined the Physics and Instrumentation Department of ONERA (the French Aerospace Lab) in 1991. His research interests concern micro and nanotechnologies and resonators, in particular Inertial Vibrating MEMS (Vibrating Beam Accelerometers and Coriolis Vibrating Gyros). Since 2012, he is Deputy Director of Physics and Instrumentation Department at ONERA.

Alexandre Bresson. Engineer graduated from SupOptique (France), PhD in Physics (2000), he is currently Research Scientist at ONERA, in charge of atom inertial sensor development. He Heads the research unit SLM (Laser Sources and Metrology) at DMPH (Physics and Instrumentation department).

N. Zahzam was born in 1979. In 2005, he obtained his Ph. D degree devoted to the study of Bose-Einstein condensation, cold atoms and molecules trapping. Since 2006, he has worked at ONERA in the field of cold atom interferometry

Bruno Christophe graduated from ISAE/SupAéro Engineer school in 1990. He joined ONERA in the Syntheses Study Direction in 1990. In 2000, he joined the Physics and Instrumentation Department at ONERA for GOCE accelerometer development, as Technical Manager. Since 2011, he is head of Instrumentation and Aerospace Equipment, involved in the space accelerometer development and project manager of GRACE-FO accelerometer. He is member of CNES advisory group for fundamental physics since 2011.

Manuel Rodrigues is research engineer in physics from the *Ecole Supérieure de Physique et de Chimie Industrielles de la ville de Paris*. In ONERA since 1991, he was first involved in test bench development for thin film sensor for aeronautics and space equipment, and then became MICROSCOPE Project manager.

S. Duzellier, J.P. David,
C. Inguimbert, T. Nuns,
G. Hubert, L. Artola
(ONERA)

E-mail : sophie.duzellier@onera.fr

DOI : 10.12762/2016.AL12-12

Radiation Testing of Electronics Systems: How Can Simulation Tools Help in the Definition and Optimization of Test Plans in Labs?

On-board space systems are exposed to harsh environmental conditions including radiations, extreme temperatures and high-vacuum… Each component of this space environment, either separately or in synergy, induces short or long term degradation of electronics or materials, thereby impacting reliability. The wide diversity of observed disruptions, potential anomalies (and therefore of their underlying physical-chemical mechanisms) and the large set of technologies involved require the development of modeling tools to support experimentation, in order to achieve a representative simulation of device response and a reliable prediction of the life, as well as to limit ground testing.

Introduction

On-board space systems are exposed to a harsh environment composed of highly diversified components. At the electronics and materials level, radiations (Figure 1), temperature, vacuum and other conditions, locally induced specific conditions such as contamination, and charging conditions affect the overall performance of sub-systems. They separately cause short term functional disruptions in electronic circuits or a more gradual degradation of materials and, moreover, their synergy adds to the criticality of this problem with increased loss of performance.

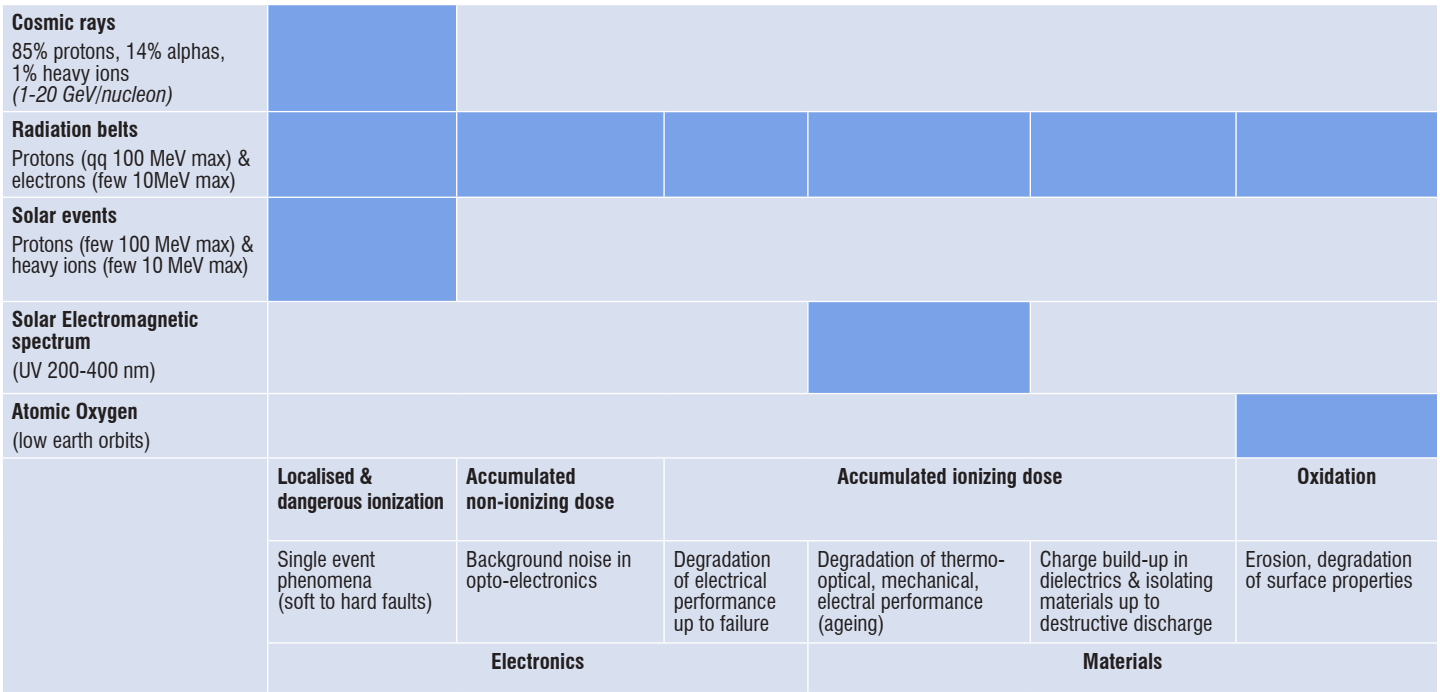


Figure 1 – Space environment components and associated effects on space systems

Consequently, the complexity and wide diversity of the mechanisms at play in observed anomalies and degradations render the study of these phenomena difficult and, moreover, the available ground capabilities for the simulation of these mechanisms are limited and incomplete.

In this game, coupling modeling and experimentation is often mandatory to obtain a realistic and representative description of the device response. The reliability challenges are manifold: event mechanisms or large shift of electrical parameters due to the ionizing dose may lead to functional failure of circuits embedded in critical sub-systems (attitude and orbit control, scientific data acquisition, etc.), and performance loss in thermal control or solar array sub-systems leads to a shortened mission duration.

The physics behind a failure in a circuit or the degradation of passive material is different and therefore the methods for their evaluation and qualification are also. Norms and standards exist with a very generic approach for electronics and more specific to the application for materials. In any cases, testing is complex and costly because many experimental parameters affect the representativeness and thus the validity of the measured data (Table 1).

The wide range of technologies (materials and electronics) systems, and processes involved in any single one of these categories adds to the difficulties in carrying out experimental activities: this wide diversity is also encountered in the observed responses under radiation exposure.

In electronics, lot-to-lot down to part-to-part variation is observed in the total ionizing dose response. The measurement of the single event sensitivity of digital devices entails technological variability that affects the event occurrence threshold (the minimum charge needed to induce event is actually measured as a distribution). Then, depending on the circuit function, the mode of failure may be critical or not (a transient signal can be interpreted as a valid signal, or filtered in the circuit).

Therefore, the ground evaluation of a device for its specific application requires very complex experimental investigation and cannot be deduced from a generic approach. In such a context, numerical tools are very helpful for identifying anomalies or degradation mechanisms and thus 1) support the definition and optimization of a test plan, 2) develop adequate prediction codes or extrapolate test results to complete an available data set, thereby limiting the number of qualification tests.

In the following, several study cases are described to illustrate how modeling and experimentation can combine successfully to investigate radiation effects in electronics, providing advances and discussing the benefits that can be reaped by ensuring the reliability of space systems.

Single Event Modeling with MUSCA SEP3

SEE (Single Event Effect) modeling is complex because many physical mechanisms intervene between the incoming particle and SEE phenomena, each playing an important role in the event occurrence. The ionizing particles of the environment (composed of protons and heavy ions for the space environment and of neutrons, protons and muons for the atmospheric environment) pass through the vehicle structure, shielding and circuit package all the way through to the sensitive structure. Then, primary or secondary ions induced by nuclear reactions generate electron-hole pairs along their paths in the semiconductor, which progress into the media according to transport mechanisms and are collected at the electrodes of the device. Finally, transient currents (SET, Single Event Transient) disturb the circuit operations and lead to the occurrence of SEE.

Experimental parameters critical for the ground simulation of radiation effects			
Single event (functional anomaly)	Degradation of electrical perform		Degradation of thermo-optical performance of surface coatings (ageing of polymer)
	Shift of electronical parameters (ionizing dose)	Background noise in opto-electronics (non-ionising dose)	
> Particle species & energy (deposited local charge density) > Particle flux & fluence (anomaly statistics, testbench detection capability) > Bias Conditions (collected charge quantity) > temperature (internal gain and parasitic signal amplification) > *Topology (circuit reponse to parasitic signal, functional anomaly criteria for occurence)*	> Particle species & energy (deposited local charge density) > Dose rate (charge generation charge) > Bias Conditions (electric field, charge transport) > Temperature (charge transport and annealing) > *Oxide quality (charges mobility and trapping rate)*	> Particle species & energy (species, size and distribution of defects) > Bias Conditions (charge carriers assist in the rearrangement of defects) > Temperature (re-arrangement/annealing of defects) > *Topology (edge effect, dark current distribution, etc.)*	> Particle species & energy (dose profile, degradation mode) > Dose rate (free radicals generation charge) > Test methods (combined/sequential exposure, sequence and timing) > Temperature (scission/xlinking ratio of molecular chains, annealing) > Vacuum (annealing of free radicals, absence of oxidation)

Table 1 – Critical experimental parameters in 4 typical cases. In italics, some technological parameters (device-dependent) directly affecting circuit response

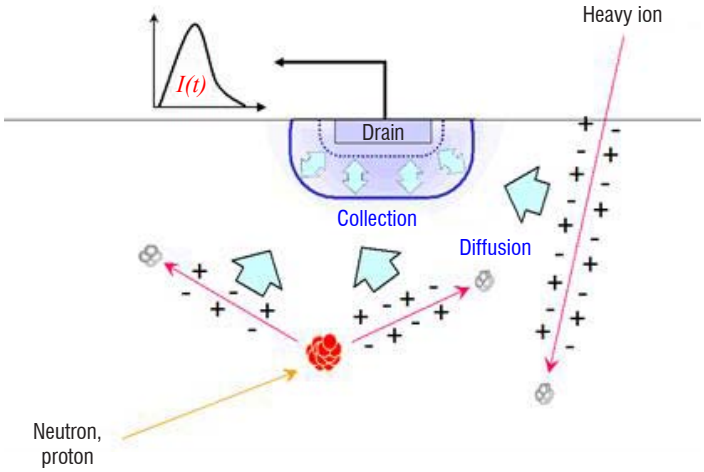


Figure 2 – Principles for charge generation and collection leading to SET occurrence at the transistor drain. This SET is at the origin of the electrical disruption that may trigger an SEE event at the circuit (functional) level

The SEE phenomenon is probabilistic, *i.e.*, it is possible to associate a threshold and probability parameters to describe the phenomena (the sensitivity or cross-section is determined by the ratio of the events count to the received particle flow).

The MUlti-SCAles Single Event Phenomena Predictive Platform (MUSCA SEP3), in development at ONERA since 2007, aims at evaluating the SEE risks within the framework of research and industrial applications. Therefore, MUSCA SEP3 has been designed to be part of the overall device qualification methodology, to allow emerging risks or problems to be anticipated, as expected from the technology roadmap (ITRS -International Technology Roadmap for Semiconductors) or new applications, and to support the analysis of in-flight anomalies. In addition, the chosen approach accounts for structure geometry and design rules and thus helps the investigation of the SET characteristics and sensitivities as a function of a cell design, representing an opportunity to define and validate solutions for SET mitigations or hardening by design process.

SEU (Single Event Upset) estimate is performed by means of MUSCA SEP3, whose detailed framework is presented in previous works [Ref. 1, Ref. 2, Ref. 3]. It is based on a Monte Carlo approach, and consists in sequentially modeling all of the physical and electrical mechanisms, from the global system down to the semiconductor target: (a) the radiation field, (b) the transport mechanisms of radiation particles (protons, neutrons or heavy ions) through the materials comprising the shielding and the Back-End Of Line (BEOL), (c) the generation of electron-hole pairs in the semiconductor via direct or indirect ionization mechanisms, (d) the charge transport and collection mechanisms in the Front-End Of Line (FEOL), (e) the circuit electrical response. The modeling of radiation effects in nanoscale devices implies taking into account a high-level physical description. Thus, realistic primary or secondary ion track structures from databases generated by GEANT4 can be coupled with MUSCA SEP3 to model the 3D structure charge deposition by the incident particle [Ref. 4]. 3D carrier morphology evolves according to mechanisms like drift (electric field), diffusion (carrier concentration gradient), collection and recombination processes. Bipolar amplifications can also be considered. Thus, models describing the transport and collection mechanisms mainly resulted from TCAD simulations, and were calibrated for investigated technological nodes. The particle-induced parasitic currents disturb the circuit response; this will depend on the transient characteristics (duration, amplitude, shape and multiplicity). Transient currents can be injected on each collection node, *i.e.*, the drains of each transistor.

With this tool, it is possible to address numeric circuits (up to a few hundreds of transistors), and destructive effects in power devices. Moreover, the continuous evolution towards a more technological integration brings with it a new set of error mechanisms or problems requiring basic research and scientific investigation to propose alternatives to the obsolete models.

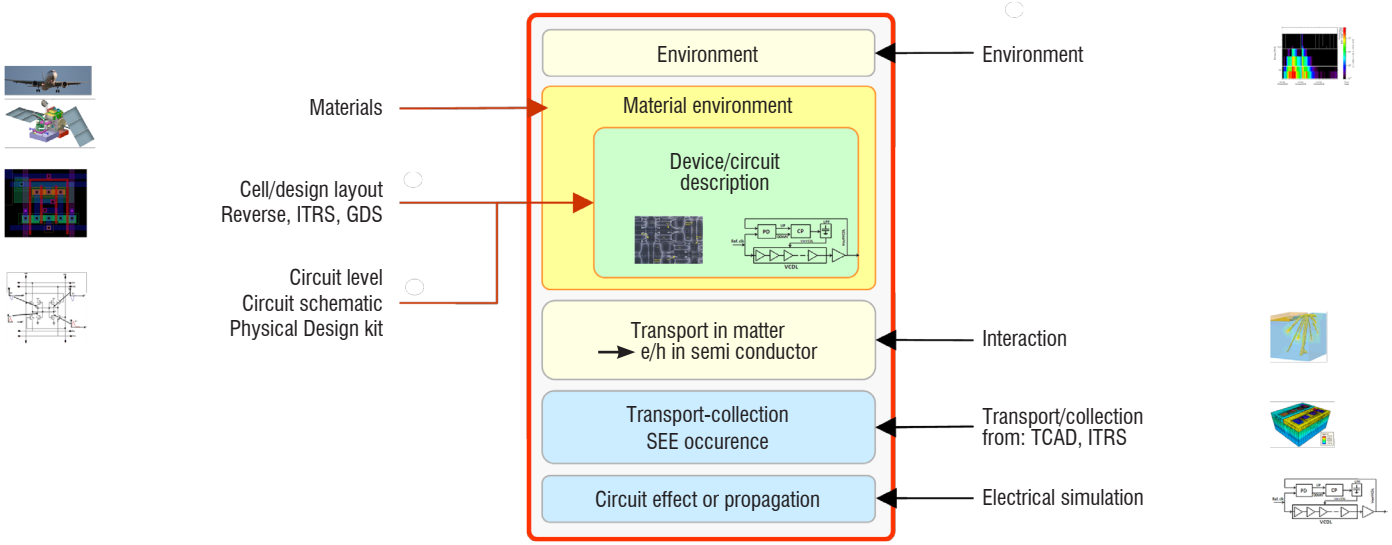


Figure 3 – The MUSCA SEP3 platform integrates the whole physical modeling chain (sequentially) from charge injection down to the event mechanism accounting for the sensitive structure within its global and local environments (hard – material and soft – application)

Data assimilation

Within the framework of a standard qualification, the experimental evaluation of SEE sensitivity for a device requires a complex test plan under many measurement conditions: different particle species and energies (definition of the mode of charge injection), many flux and fluence levels for plotting sensitivity curves, covering the response to most of the expected environmental conditions. This approach is expensive and time-demanding.

The MUSCA SEP3 core approach is based on a technological description of the target structure associated with representative physical parameters, in order to describe the device response to any type of environment (space, atmospheric and terrestrial applications). Thus, the MUSCA SEP3 can be considered as a virtual irradiator. This set of physical parameters can be reduced to a set of the most critical parameters within the framework of an industrial application and project context, without much affecting the validity of the final results, in order to meet the industrial constraints (timescale, cost, available information, etc.). However, as for any modeling approach, experimental data are required to validate the models and obviously the level of reliability and precision of the tool will be defined by the quantity of measurements that can be considered for this "calibration" step.

Figure 4 shows some results of memory sensitivity to protons determined from heavy ion experiments and vice versa (by data assimilation, the first results from this approach can be found in Ref. 1). Indeed, the proton cross-section is used to determine the critical charge by fit process, and then this critical charge value is considered to calculate the heavy ion response. This assimilation process can be applied for any radiation fields.

With such use, the MUSCA SEP3 contribution helps to optimize a test plan (selection of a minimum set of beam conditions to obtain a satisfactory radiation response description) and to complete a data set without performing a large number of tests (gain cost and time for a project).

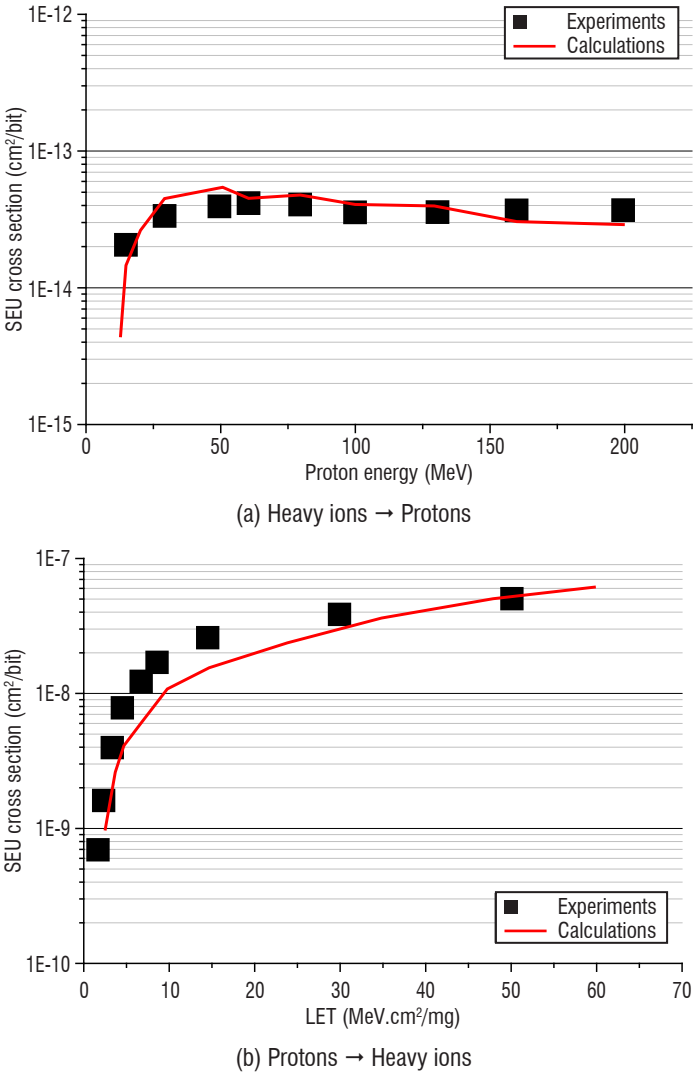


Figure 4 – Comparison of measured and calculated cross-sections in the case of a RAM memory. (a) Heavy ion-to-proton calculation compared to proton experimental data, (b) Proton-to-heavy ion calculation compared to heavy ion experimental data.

Test plan definition

Temperature effect on SET feature induced by ionizing particles, such as heavy ions, have mostly been studied for a high temperature range over the last decades. For infrared technologies, investigations have been dedicated to very low temperatures down to 77K (supported by CNES). This study was focused on the SEU sensitivity of D Flip-Flops used in the readout circuit of a CMOS image sensor developed by Sofradir.

The SEU sensitivity estimate obtained by calculations with MUSCA SEP3 has been compared with success to preliminary experimental data obtained from an irradiation test campaign performed by the CNES in 2014. However, a strong variability of the experimental data was observed although the error bars of the measured events were limited (statistical uncertainties). The interest of the MUSCA SEP3 prediction platform is, in addition to the estimation of SEE cross-sections, to allow for a failure analysis at the design and transistor levels, as shown in Figure 5 for a DFF cell.

It is interesting to note here that the critical features of the SEE sensitive zones associated with the logical states "1" (in red) and "0" (in blue) of the DFF are diversified (area and shape) and scattered (location).

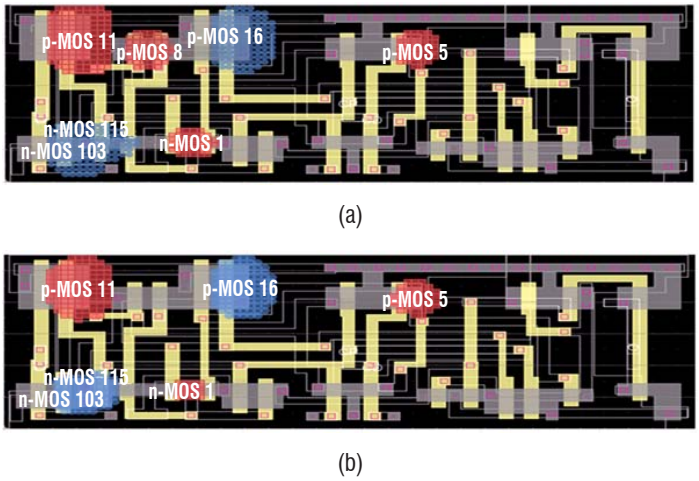


Figure 5 – SEU sensitivity mapping of the DFF cell reference design, as a function of the stored data, a) "0" (blue areas) and b) "1" (red areas) at 300K. The locations, shapes and areas of the critical zones determined by multi-collection and circuit effects affect the global measured SEU cross-section.

SEE phenomena are probabilistic by nature; any disparity in the topological origins of the events can induce strong variability of the measured SEU cross-sections, if the number of measured events is not statistically significant relative to the DFF cell area and the number of critical zones. Indeed, only 20 errors were measured during the first irradiation campaign in June 2014, while over 160 events have been estimated by MUSCA SEP3 for a relevant mapping of SEU events (Figure 5). Thus, in 2015 the experimental setup of the second irradiation campaign was defined using the MUSCA SEP3 calculations as inputs, in order to achieve the best trade-off for the relevant particle fluence to be used (which was actually increased by a factor of 5). Figure 6 shows the impact of fluence on the measured SEU variability (a) and compares the measurements and calculations for both irradiation campaigns (b) [Ref. 5].

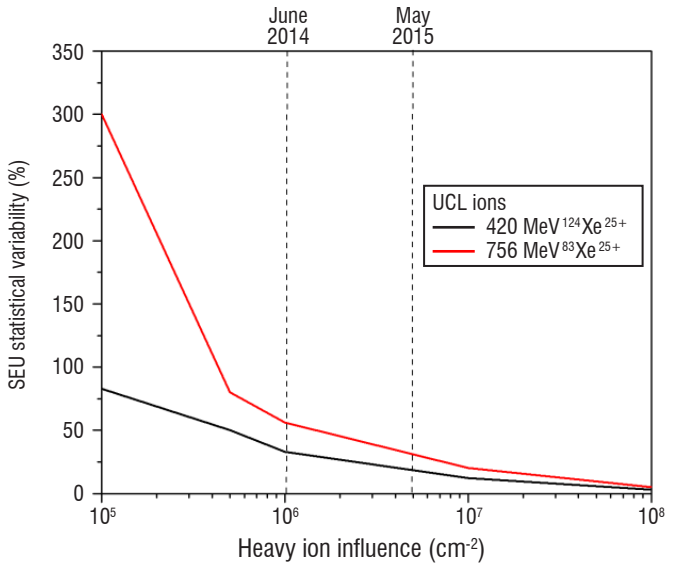


Figure 6 – (a) SEU Statistic variability as a function of fluence for 420 MeV Xe and 756 MeV Kr ions, calculated by MUSCA SEP3 for the reference design of the DFF. (b) Comparison of the experimental data obtained from both irradiation campaigns (2014 and 2015) with the SEU cross-sections calculated by MUSCA SEP3 as a function of the logic state of the DFF.

Thus, the interest of such a SEE prediction platform has been shown: first for the estimation of the SEE sensitivity, second for failure analysis (impact of the design), and third for the definition of the experimental setup of irradiation campaigns, constituting a major asset to reduce the cost defined by the space industry roadmap.

Modeling of ionizing dose phenomena with the AC-DC code

The AC-DC code (Analytical Computing of Dose -induced Charges) was developed at ONERA-DESP with the aim of developing a physical tool to finely describe the TiD (Total Ionizing Dose) mechanisms in the circuits. This degradation effect is related to the trapping of charges in the insulating zones (often SiO_2) of transistors and/or isolation of the active areas of the circuits. Although the mechanisms of generation, recombination, transport, and charge trapping in insulators and interfaces are well known [Ref. 6, Ref. 7], due to the number of parameters and sensitivity it is difficult to implement models capable of predicting the response of a circuit at the functional level.

This "physical modeling" approach is undeveloped in the community, but was chosen as the back track for experimental investigations in the department. The AC-DC code has been developed with the aim of answering questions arising from the circuit evaluation methods, as well as regarding the influence of critical experimental parameters (temperature, dose rate, bias, etc. [Ref. 8]). It describes in a SiO_2 layer the charge generation and initial recombination, the free electrons and the hole diffusion and transport, the deep and shallow trapping/de-trapping processes and the [H+] ions generation and transport. At the SiO_2/Si interface, the interface state generation and the holes trapping on these interface states are modeled [Ref. 9]. Finally, the interface trapped holes profile is calculated, taking into account thermal and tunnel-assisted annealing processes [Ref. 10]. The main physical processes are electrical field-dependent; a coupled solving of Poisson and drift–diffusion equations has been implemented in a 1D MOS structure. All of the well-known dependencies on temperature are explicit for all of the physical processes, in order to evaluate the effect of temperature on electrical degradation.

The AC-DC code was first used to study the ELDRS (Enhanced Low Dose Rate Sensitivity) mechanism observed in bipolar technology essentially. When ELDRS occurs, an electrostatic shielding mechanism (Figure 7) limits the degradation at a high dose rate and therefore becomes critical at low rates since the degradation in a spatial application is stronger than under ground evaluation conditions.

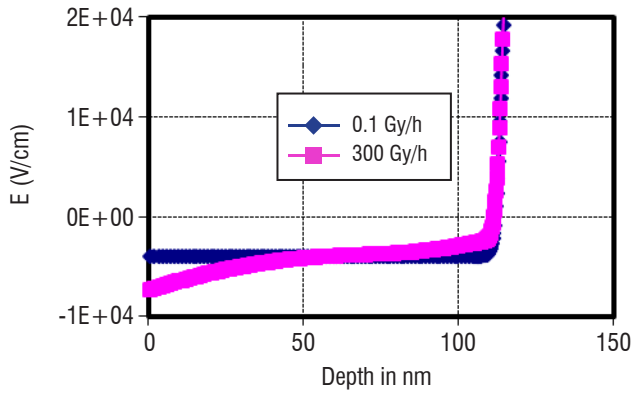


Figure 7 – Electric field profiles in the silica volume of an irradiated MOS for 2 different dose rates. These differences are the cause of the phenomenon ELDRS.

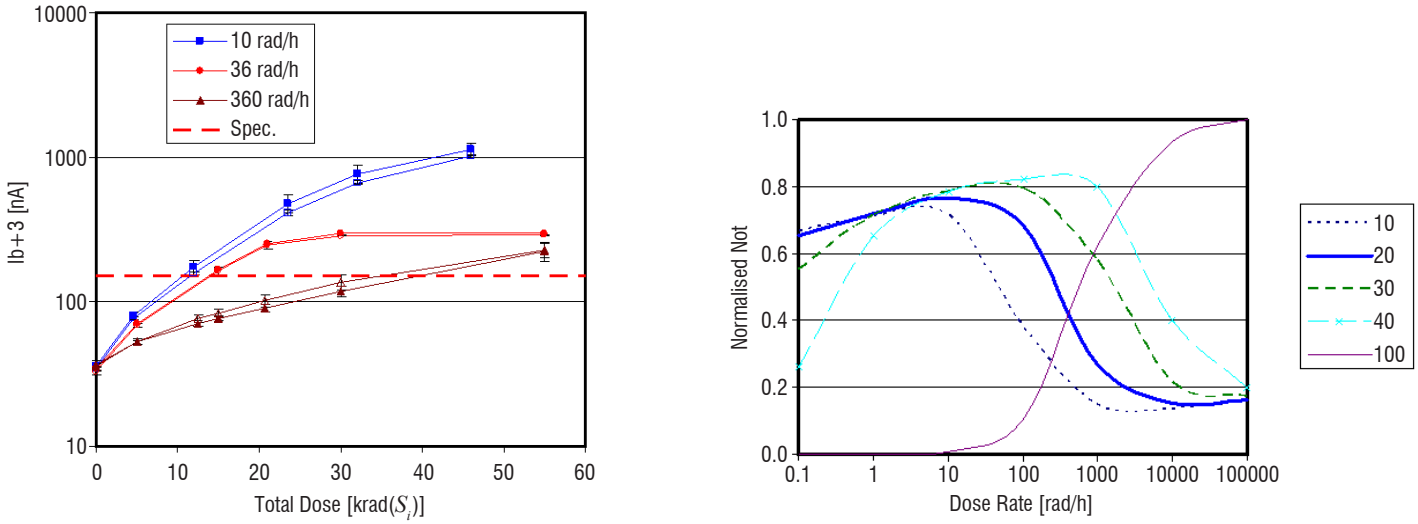


Figure 8 – Test results on a LM111 operational amplifier. (Left) influence of the dose rate on the positive bias current (experiment under biased and un-biased conditions), (Right) trapped charge density (normalized) in the oxide as a function of the dose rate for different irradiation temperatures (AC-DC calculations).

In 2012, tests on a representative set of components were conducted for ESA as the basis for the definition of recommendations, as part of the establishment of the new Test Standard ECSS22900 (dose rates: 36-360rad/h, Figure 8 left).

AC-DC modeling has shown that accelerated configuration could be proposed. When a temperature of 40°C is applied during irradiation, the experimental standard dose rate provides a conservative estimate of the TiD resistance of devices relative to the space dose rate; indeed, the quantity of trapped charges induced at ground testing remains a worst case (Figure 8 right). Obviously, this approach requires a "calibration phase", since many physical parameters are technology-dependent (quality of the oxide). However, such a tool is helpful for proposing an optimized test configuration.

Another example is that of IR sensors operating at temperatures near 80 K (study classified as confidential). The TiD response at such temperatures is not well known and requires modeling as support for analysis (impact of charge trapping and transport).

The experiment developed to answer these questions is based on the evolution of the amount of charge during a rise in temperature after irradiation at 80 K on appropriate test vehicles (MOS transistors) and various topologies (design influence). For this type of topic, where experimentation is quite heavy (irradiation of components in cryostats, use of liquid nitrogen) and few results are available in the literature, modeling the physics with AC-DC is a valuable tool to select the critical experimental parameters and to help in the interpretation of results. In particular, it enables the evaluation of the theoretical response of a MOS structures irradiated at 80 K during a temperature ramp (Figure 9).

In the figures below, the calculated degradation is represented by the two characteristic quantities, volume (Qot) and interface (Qit) charge concentrations, and for three dose rates. The 80K degradation that is observed immediately after irradiation is not the worst-case situation (as opposed to 150 K), both from the point of view of the volume or interface charges. The use of experimental results obtained at 80 K will have to consider this behavior.

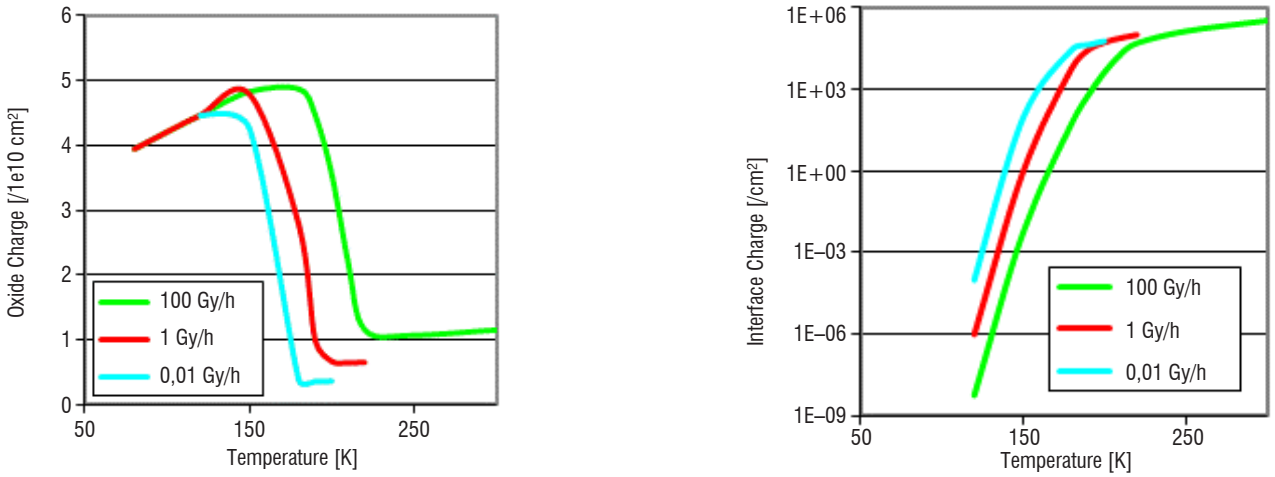


Figure 9 – Calculation with AC-DC changes in the amount of volume (left) and interface (right) charges depending on irradiation temperature and for three dose rates.

Displacement defects and the effective NIEL concept

The displacement damage and ionizing dose often combine to affect the performance of image sensors and, more generally, optoelectronics (Figure 10). The atoms displaced by proton or electron irradiation are at the origin of degradation mechanisms involving a non-ionizing dose. These defects can be electrically-active, and affect the charge carrier physics (generation, recombination, trapping, diffusion, etc.), which in turn degrades the electrical performance of circuits (for instance, an increase in the background dark current noise).

Generally, mean degradations vary linearly with NIEL (Non-Ionizing Energy Loss) when applying the NIEL scaling law, which considers that damage is proportional to the product of NIEL by the particle fluence. But deviations are observed, especially with electrons and also in cases of III-V semiconductor devices.

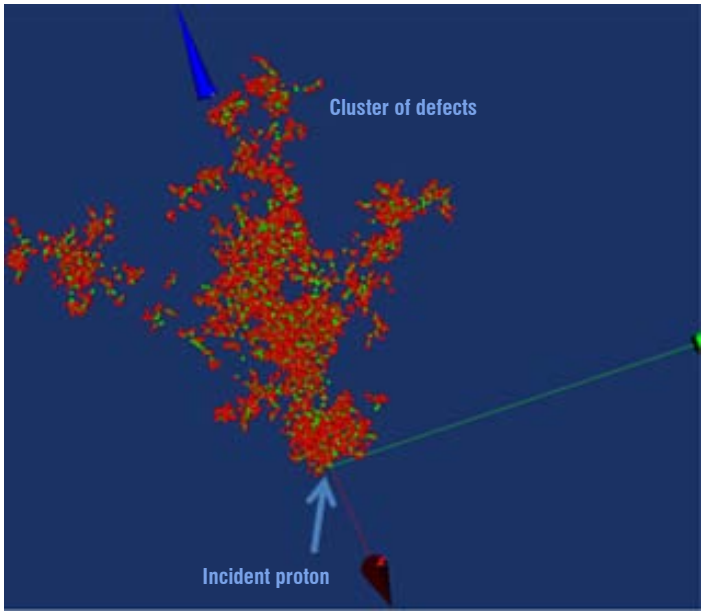


Figure 10 – Calculation with AC-DC changes in the amount of volume (left) and interface (right) charges depending on irradiation temperature and for three dose rates.

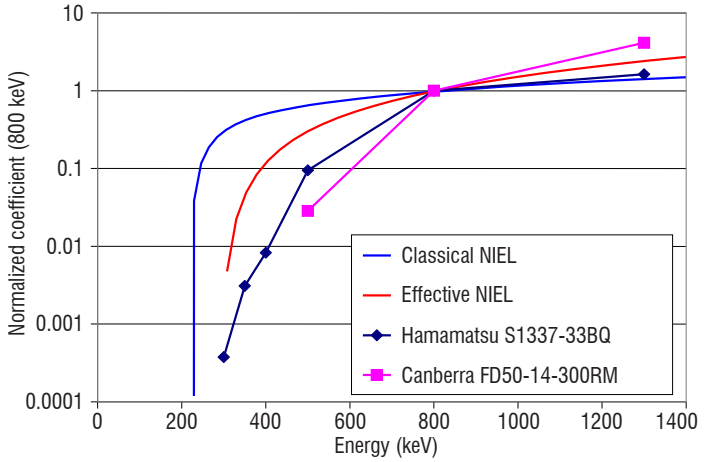


Figure 11 – Calculations based on effective or classical NIEL concepts compared with measurements.

At DESP, the existing experimental facilities (particle accelerators, characterization tools as DLTS equipment, test bench, etc.) and the numerical modeling capabilities (radiation-matter interaction tool) allowed for developing over years a deep background knowledge on physical mechanisms and thereby adequate and representative degradation models.
For instance, an effective NIEL parameter has been proposed as an improvement of NIEL calculations for predicting mean degradations [Ref. 11, Ref. 12]. In Figure 11, the deviation between calculations with classical NIEL (blue line) and experimental observations (symbols) can be observed. These deviations may come from experimental uncertainties, but more from the rough estimate of the NIEL values. Based on molecular dynamics data showing how the formation of amorphous zones (localized melting of the target matter) can modify the quantity of radiation-induced defects, the concept of effective NIEL was developed and successfully compared with existing data (red line).

These data illustrate the benefit of modeling here to provide a better description of the device response (life prediction) and to help in the selection of appropriate beam parameters for final testing.

Conclusions and perspectives

Over the last decades, the concern for radiation effects has extended to the fields of avionics, automotive, large computers and networking, etc., and to an ever wider range of technologies (CMOS, FDSOI and FinFET devices, as well as power components, SiC technology, etc.) and deeper integration. With the growing need for on-board computing, criticality increases and new mechanisms emerge (direct ionization from protons, neutrons-and muons, lately electron-induced anomalies, and now synergy between ageing and the radiation –reliability concern, etc.).

Nowadays, space missions are more diversified, leading to the definition of new environmental conditions (use of electrical orbit raising EOR, scientific missions with extreme conditions, mega-constellations, etc.). The "materials" topic is also concerned, for instance, with the EOR orbits inducing a higher electrons flux and higher doses (enhanced ageing and charging effects).

Thus, coupling modeling and experimentation begins inevitably with the anticipation of emerging phenomena in a constantly evolving context. However, access to the technology is now vital and development times are often too long for project timeframes. A realistic and pragmatic approach is therefore mandatory for future generations of tools (modularity, application-oriented, validity domain, etc) ■

References

[1] G. HUBERT, S. DUZELLIER, C. INGUIMBERT, C. BOATELLA-POLO, F. BEZERRA, and R. ECOFFET - *Operational ser Calculations on the Sac-C Orbit Using the Multi-Scales Single Event Phenomena Predictive Platform (MUSCA SEP3)*. IEEE Trans. Nucl. Sci., vol. 56, no. 6, pp. 3032-3042, Dec. 2009.

[2] G. HUBERT and L. ARTOLA - *Single-Event Transient Modeling in a 65-nm Bulk CMOS Technology Based-on Multi-Physical Approach and Electrical Simulations*. IEEE Trans. Nucl. Sci., Vol. 60, No. 6, pp. 4421-4429, Dec. 2013.

[3] G. HUBERT, S. BOURDARIE, S. DUZELLIER, R. ECOFFET - *Impact of the Solar Flares on the SER dynamics on Micro and Nanometric Technologies in the Geostationary Orbit*. IEEE Trans. Nucl. Sci., Vol. 57, No. 6, pp. 3127-3134, Dec. 2010.

[4] M. RAINE (CEA-DAM), G. HUBERT, P. PAILLET (CEA-DAM), M. GAILLARDIN (CEA-DAM), A. BOURNEL (IEF-CNRS) - *Implementing Realistic Heavy Ion Tracks in a See Prediction Tool: Comparison Between Different Approaches*. IEEE Transactions on Nuclear Science Vol. 59, No. 4, pp. 950-957.

[5] L. ARTOLA, G. HUBERT, O. GILARD, S. DUCRET, F. PERRIER, M. BOUTILLIER, P. GARCIA, G. VIGNON, B. BARADAT, N. RICARD - *Single Event Upset Sensitivity of D-Flip Flop of Infrared Image Sensors for Low Temperature Applications Down to 77 K*. IEEE Trans. Nucl. Sci., Vol. 62, No. 6, pp. 2979-2987, dec. 2015.

[6] J.-L. LERAY - *Total Dose Effects: Modeling for Present and Futur*. IEEE Trans. Nucl. Sci., NS-46, p. 1574, Dec. 1999.

[7] R.J. GRAVES, C.R. CIRBA, R.D. SCHRIMPF, R.J. MILANOWSKI, A. MICHEZ, D.M. FLEETWOOD, S.C. WITZAK and F. SAIGNE - *Modeling Low-Dose-Rate Effects in Irradiated Bipolar-Base Oxide*.

[8] R. DURAND, J.-P. DAVID, A. GAUFFIER and E. LORFEVRE - *Modeling Temperature Effects of Thick Oxide 60Co Induced Degradation Under Low Electric Field. Hardness Assurance Implications*. 9th European Conference on Radiation and Its Effects on Components and Systems, 2007.

[9] S. N. RASHKEEV, D. M. FLEETWOOD, FELLOW, IEEE, R. D. SCHRIMPF, FELLOW, IEEE, and S. T. PANTELIDES - *Effects of Hydrogen Motion on Interface Trap Formation and Annealing*. IEEE Trans. Nucl. Sci., Vol. 51, No. 6, Dec. 2004.

[10] P. J. MCWHORTER, S.L. MILLER and W.M. MILLER - *Modeling the Anneal of Radiation Induced Trapped Holes in a Varying Thermal Environment*. IEEE Trans. Nucl. Sci., VOL. 37, P. 1682-1689.

[11] C. INGUIMBERT, S. MESSENGER - *Equivalent Displacement Damage Dose For On-Orbit Space Applications*. IEEE Trans. Nucl. Sci., Vol. 59, No. 6, pp. 3117-3125, Dec. 2012.

[12] C. INGUIMBERT, P. ARNOLDA, T. NUNS, G. ROLLAND - "*Effective Niel*" *in Silicon: Calculation Using Molecular Dynamics Simulation Results*, IEEE Trans. Nucl. Sci., Vol. 57, No. 4, pp. 1915-1923, Aug. 2010.

Acronyms

AC-DC	(Analytical Computing of Dose-induced Charges (code))
BEOL	(Back-End Of Line)
CMOS, FDSOI and FinFET devices	
DFF	(D Flip-Flop cell)
ELDRS	(Enhanced Low-Dose Rate Sensitivity)
EOR	(Electric Orbit Raising)
FEOL	(Front-End Of Line)
ITRS	(International Technology Roadmap for Semiconductors)
MOS	(Metal-Oxide-Semiconductor (structure))
MUSCA SEP3	(MUlti-SCAles Single Event Phenomena Predictive Platform)
NIEL	(Non-Ionizing Energy Loss)
SEE	(Single Event Effect)
SET	(Single Event Transient)
SEU	(Single Event Upset)
TCAD	(Technology Computer-Aided Design (tool))
TiD	(Total Ionizing Dose)
TniD	(Total Non-Ionizing Dose)

Sophie Duzellier, Graduated in 1986 from the INSA (*Institut National des Sciences Appliquées*) Engineering school, and received a Ph.D. degree in microelectronics from Toulouse Univ. in 1989. She joined ONERA in 1989 as a research engineer in the Space Environment department. Her first activities addressed radiation effects on electronic devices through ground simulation and experimental work (testing methods and prediction tools). Head of the ECM group at DESP since 2005, she still carries out research activities on the topic of space materials, dealing with radiation ageing (thermal coatings, optics, etc.), and develops on-board experiments (PI of MEDET on-board ISS 2009-2010).

Guillaume Hubert, Graduated with a M.Sc. degree in theoretical physics from the Univ. of Pierre & Marie Curie (Paris VI) in 1998, and received a Ph.D. degree in electrical engineering from Montpellier Univ. in 2001. Then, he worked for five years at the European Aeronautic Defense and Space Company (EADS) on the effects of radiation on electronics. Since 2007, he is a researcher engineer at ONERA (Space Environment Department). His research activities are in the field of multi-physics / multi-scale SEE modeling for space, terrestrial and avionic environments and radiation environment characterization. He is also member of ONERA's scientific council of and the Scientific Management Board of the LSBB (Low Noise Underground Laboratory) of Rustrel.

Laurent Artola, PhD, senior research engineer at ONERA-DESP since 2012. He graduated with a M.Sc. degree from the University of Montpellier in 2007, specializing in the reliability of electronic devices subjected to radiation (in space and avionic environments). After a short experience at NXP Semiconductor IMEC (Belgium), he started his PhD program in Toulouse, supported by ONERA and the CNES (the French space agency) in the prediction of Single Event Effects (SEE) in electronic devices for space applications (PhD in 2011). He is now in charge of electrical and physical modeling of semiconductors for the prediction of radiation-induced Single Event Effects in VLSI.

Jean-Pierre David, Graduated in 1980 from the INSA (*Institut National des Sciences Appliquées*) Engineering school in solid state physics, he was hired at ONERA in 1982. His early work mainly concerned the characterization of III-V materials, and then, in the 1990s, he particularly studied ionizing dose effects and hardening assurance issues on electronic circuits. He is now the Deputy Director of the Space Environment Department, but is still in charge of studies on the topic of atomic displacements in photonics and opto-electronics (solar cells).

Thierry Nuns, Graduated from the SupElec (*Ecole Supérieure d'Electricité*) Engineering school in 1993, he joined ONERA in 1995 as a research engineer in the Space Environment department and defended his PhD in 2002 on radiation effects in commercial CCDs (ionizing and non-ionizing dose effects). In the 90s, he first led the development of software/hardware for the radiation testing of devices and flight experiments in collaboration with the CNES. He is now in charge of optoelectronics activities at the ECM group (degradation models due to ionizing and non-ionizing doses) and is developing radiation monitors based on imagers.

Christophe Inguimbert, Graduated in 1995 from Paul Sabatier University in Solid State Physics, then PhD degrees at ISAE in space electronics in 1998. After his Ph.D, he joined the space environment department of the ONERA (ECM group), where he has taken over the radiation transport-in-matter activities. He is in charge of developing the GEANT4-based codes to address the topic of radiation-induced degradation in electronics and materials (atomic displacements in optoelectronic devices, particle-semiconductor interactions leading to Single Event Effects, secondary electrons emission in materials, etc.).

P. Bidaud, L. Burlion,
H. De Plinval, T. Loquen,
J. Marzat, C. Pralet
(ONERA)

E-mail: philippe.bidaud@onera.fr

DOI : 10.12762/2016.AL12-13

Dealing with Complexity through Advanced Control Techniques

The control of complex dynamic systems, both in their behavior and in their mission, goes through the implementation of multi-loop control architectures based on information about the system internal state and from the environment, as well as on the mission plan state. This results in systems that are becoming increasingly autonomous, for which requirements in terms of safety and reliability, as well as expected performance, are increasingly high. Research works developed at ONERA in the field of control for autonomous systems cover all levels of the control architectures, which are basically structured with respect to temporal aspects, as well as the level of abstraction that they entail for the system dynamic. We will consider them in this paper by increasing level. We will discuss the advances achieved recently in the robust control techniques of uncertain dynamic systems generally implemented at the lower control level and we will discuss their extensions to consider input and output constraints, as well as the hybrid nature of most of the systems considered. To design "task" level control primitives that take place just above the previous control loops, we will introduce sensor-based robust and non-linear control techniques. These are based on information on the environment extracted from exteroceptive sensors, to adapt system behavior to uncertainties and perturbations. Multi-sensor and/or multi-objective controls will be discussed in this particular context. We will also present several recent results in the field of trajectory tracking based on visual navigation techniques in complex environments, which combine objectives and constraints within the same control architecture. We will discuss how model predictive control (MPC) techniques and advanced optimization techniques can be used for solving the resulting control problems. In addition, we will discuss several ongoing developments of these methods by exploiting distributed model predictive control techniques (DMPC) and predictive control of hybrid systems. Finally, integration with the control architectures at the upper level of reactive, predictive and distributed planning capabilities will be proposed to accommodate time constraints and uncertainties in decision.

Introduction

The issues underlying a large number of research activities in the departments for Information Processing and Systems at ONERA deal with control of systems with complex dynamic behavior, such as aircraft, spacecraft, robotics systems etc., and this in the presence of uncertainties and various constraints related to physical systems and the environments in which they evolve. Consideration of these phenomena is of major importance in the control of real systems.

In order to cope with the increasing complexity of technological systems in general and with the requirements in terms of performance and adaptability, for which there are increasing demands, a trend that can be observed since over two decades ago is to abstract the complexity of systems by integrating the control means responsible for the system behavior and their adaptation to the various "tasks" defined at a high-level; tasks which may require coordination between multiple

entities. In the implementation of these systems, which thereby become "robotic systems", safety and security issues are and remain of major importance.

The various components of the control software are designed to be integrated into an overall architecture whose structure generally separates different levels that may or may not be implemented, depending on the degree of "autonomy" to be achieved.

- The lower level is in charge of multi-loop actuator servo modules, generating control inputs and perception modules that are often used in sensor-based control loops. This level allows elementary actions of the systems to be controlled and allows perception information on the environment, as well as on the system state, to be accessed.

- The executive level operates the execution control of elementary functions (*i.e.*, the tasks defined through low-level components carried out by the entities involved, their organization and their interactions defined as a plan) and the evaluation of a number of functional and temporal properties to be fulfilled at runtime.

- The decision-making level is in charge of scheduling algorithms (*i.e.*, creation and supervision plans). It controls how the robot will bind tasks to achieve each objective of its mission. Higher-level decision-making mechanisms, known traditionally as deliberative mechanisms, are the central mechanisms in these architectures [1].

Notice that, beyond these control methods and algorithms, ONERA jointly carries out important work on formal and semi-formal specification and verification software of the underpinned critical software architectures, taking into account their hardware implementation [2].

One of the major limitations in these control structures, which combine reactive and deliberative mechanisms around an execution controller, is that complex behaviors are not anticipated, nor are the constraints on these behaviors. Also, we frequently seek to mix the principles of predictive control and stochastic optimization, aiming towards "goal-driven control architectures". This concern, as we shall see, is strongly present in the current research developments at ONERA and is expected to increase in the future.

Figure 1 – Example of a robotics system composed of several autonomous entities deployed to perform high-level tasks, such as the surveillance of an area

In this paper, we review and discuss the potential offered by current and future developments carried out at ONERA in the field of control engineering for the autonomy and reliability of complex systems. On this occasion, we will introduce, in the various sections concerned, several ONERA-toolboxes (*SMAC*, *COPERNIC*, *InCELL*, etc.).

We will start with robust control tools and their ongoing enhancements to consider hybrid systems and sub-system constraints (all physical systems have inputs and outputs, which are limited in size due to safety or physical constraints). Robust control laws are designed to achieve the desired behavior of the controlled system and maintain this behavior when faced with disturbances and hazards that affect the system during operation. Achieving a high performance and ensuring the safe behavior of complex dynamic systems, despite the parametric uncertainties and/or failures, has been for several years a central theme in control engineering activities at ONERA. These issues have led to the development of a consistent set of estimation methods for robust control synthesis, analysis and validation of control laws, in particular to limit the costs of the certification process. They provide a unified framework for control from modeling to time-frequency analysis and synthesis of robust control laws. We will discuss more specifically here the extensions of this work, first to consider input/output constraints through model predictive control techniques and anti-windup compensation and then to exploit hybrid control techniques.

At a higher level of control, exteroceptive-sensor based control loops enable "task-oriented" functions to be set by controlling a system, not in the system configuration space, but rather in the "sensor measurement" space in the form of a local relationship between the system and its environment. These output feedback control loops are called "sensor-based control". They permit the need for an accurate model of the system and errors to be avoided by setting control loops based on the information obtained by the sensors (vision sensors, force sensors, proximetric sensors, etc.). The use of visual feedback to perform closed-loop control on geometric primitives extracted from features observed by a camera embedded on aerial robots have been specifically developed at ONERA to deal with complex behaviors. We have developed methods of analysis and synthesis of "advanced" 2D1/2 visual feedback, taking into account all of the constraints (visibility, saturation of actuators, 3D constraints while driving, etc.), uncertainties related to the calibration of cameras and the non-modeled dynamic environment like the aerologic perturbations, for example. We will discuss how these techniques enable the problem of "hybrid tasks" to be solved rigorously, combining heterogeneous sensor data in the low-level servo controller.

In addition, we are also developing vision-based state estimation and trajectory control approaches (monocular-, stereo-, and RGB/D camera-based) that exploit the principles of SLAM (Simultaneous Localization and Mapping). We will not discuss SLAM techniques in this paper and the reader can refer to several recent references from ONERA in this field [3] [4]. However, we will propose a generalization of navigation techniques with steady and dynamic obstacle avoidance, as well as cooperative guidance based on predictive models of the behavior of the system. We will, more particularly, put into perspective the potential of Distributed Model Predictive Control (DMPC) methods and hybrid methods for dealing with continuous and discrete variables.

At the higher level, to achieve a complex mission, action/perception tasks are then structured into a plan generated by algorithms (planners), which relies on both a representation of the dynamic properties of the system and of the tasks, as well as on an integrative level of

information to obtain a certain representation of the world in which the system operates. The mission management is a high-level control, which determines, depending on the operating environment and on the overall goal, the elementary tasks to be carried out by the perception and action entities. The methods studied at ONERA in this field rely on formal models of the considered system (or subsystem), a formal model of the possible changes in the environment in which the system operates, and a formal model of the goals to be achieved. We thus use general mathematical modeling frameworks, such as Constraint Satisfaction Problems (CSP), discrete event systems (Petri Nets, Temporal Networks , etc.) or Markov Decision Processes (MDP), to make decisions based on uncertain data. We will discuss several directions that we are exploring to increase the system autonomy at the planning level and to make planners more reactive and proactive, such as hierarchical deliberations, distributed planning and the use of stochastic optimization and predictive models.

The following sections consider these ongoing developments, or those considered on these different control levels, from the point of view that we have of the evolution of aerospace and defense/security systems autonomy.

Achieving robust constrained control of dynamical systems

The development of robust control techniques for sequenced and multi-objectives systems is an important area of research at ONERA. Robustness to uncertainties for stationary systems (LTI) has been the purpose of many developments based on the LFT formalism and a large set of appropriate tools for the construction of this type of model and for robustness analysis are today available in the SMAC toolbox w3.onera.fr/smac/.

Once an LFT model of the closed loop is available, validation of control laws may be considered by various analytical techniques for robust stability and performance, using analysis techniques [5] particularly suitable for LTI systems. Current research at ONERA in this area is aimed at:

- better controlling the trade-off complexity/precision [6, 7, 8],
- dealing with higher order systems (flexible modes, finer modeling calculators) [9],
- controlling precision and reducing conservatism (less pessimistic margins) [10].

Stationary uncertainties represent only a relatively small part of all of the uncertainties encountered in practice. Most effects vary over time and systems being studied become unsteady. Consequently, analysis techniques are no longer applicable and the control law validation issue becomes numerically much more complex. Two large families of techniques exist to address this kind of uncertainty. The first one, inspired by analysis techniques, is based on the optimization of matrix "scaling" that are constant (for LTV uncertainties) and dynamic (for LTI uncertainties) in the frequency domain [11, 12]. This approach relies on special techniques based on the IQC formalism (Integral quadratic constraints). The second one makes use of dependent Lyapunov parameters. For the cost of higher computational complexity, these techniques can be considered to be bounds on speed parameter variation. The specific contribution that we made to these techniques is the reduction of their numerical complexity to make them competitive over intensive simulations [13].

The robustness analysis in the presence of nonlinearities such as input/output constraints remains an open research area, which motivates various extensions for tools developed as part of the analysis of robustness LTI/LTV. Inputs and outputs constraints arise in most control engineering applications. This topic is crucial since ignoring these constraints can lead to a dramatic loss of stability and performance. It is therefore not surprising that the topics of model predictive control [14, 15] and anti-windup compensation [16, 17] have been intensively studied for many years.

Model predictive control

Model Predictive Control (MPC) is a branch of control design techniques that take advantage of the knowledge of an explicit model to predict the future response of the plant over a sufficiently large time interval (i.e., the selected time horizon exceeds the expected transient time of the control system). Using such a prediction, the control problems are then formulated as optimizing on line the system behavior under its input and output constraints. Although this technique is very popular and effective when the model is linear, low dimensional and perfectly known, it is sometimes criticized because of its lack of intuition in terms of tuning, thus designing a controller "from scratch" by an "obscure" optimization and because of its computational load or conservativeness when the model is uncertain or even nonlinear. Recent research has already started addressing such issues [18, 19, 20] and model predictive control deserves further consideration in aerospace control where the models are often uncertain and high dimensional. We will consider the use of MPC techniques in Section 5. Another limitation in use of the MPC is the computational cost compared to anti-windup techniques, which are discussed below.

Anti-windup compensation

Anti-windup is a very popular technique that is currently mainly used to deal with control saturation. In anti-windup compensation, a nominal controller that does not explicitly take into account controller saturations is first designed using the best control design tools. Then, after that, an anti-windup compensator is combined with this controller to both ensure the stability (at least in some region near the origin when the open loop system is unstable) and to avoid performance degradation when saturations are active. It is worth noting that such a technique is appealing because the desired response of the nominal control law is recovered when the saturation limits are not exceeded.

Nominal control law design

As a main drawback for practitioners, the nominal control design (i.e., the unconstrained control law design) specifications of many aerospace applications require either a large amount of tuning or mastering advanced control design tools.

Most of the time, the nominal control law must be designed using powerful control design techniques that account for uncertain and high dimensional linear models. Another possibility in which we are deeply involved is the use of specialized non-smooth optimization techniques dedicated to control system design [21, 22, 23].

Input constraints and anti-windup

Among many possibilities that are summarized in [16, 17], the so-called MRAW (Model Recovery Anti-Windup) is often preferred because it can be applied with any (possibly nonlinear) nominal

controller [24, 25]. First, the nominal control is designed; then it is merged with a global controller that is able to globally (or at least locally) stabilize the plant, despite the control saturations. This anti-windup solution is illustrated in Figure 2.

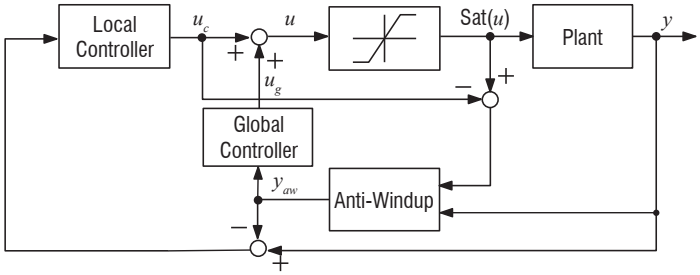


Figure 2 – Model recovery anti-windup architecture for merging the local and global controllers.

Basically, anti-windup block design involves solving some LMIs (Linear Matrix inequalities) [16, 17]. However, some papers have shown how this problem should be addressed, by using other powerful non-smooth optimization techniques [26, 27]. Recent efforts have also been concentrated on MRAW design when the plant model is uncertain (see e.g., [28, 29]).

Output constraints and anti-windup

Among several approaches summarized in [30, 31], taking output constraints into account may consist in modifying a nominal controller whenever the constraints are about to be violated. Following this line, the Output to Input Saturation Transformation (OIST) approach was recently proposed [32] to replace an output constraint by a state-dependent control saturation. Such an approach is attractive because the output constrained problem is then recast into an input constrained problem, which may in particular lead to considering the aforementioned MRAW loop design.

This OIST technique was originally proposed for the state feedback control of nonlinear systems when a minimum phase output is constrained. Guarantees on the global asymptotic stability of the closed-loop in the presence of the obtained control saturation have also been studied for a large class of linear systems [33]. The OIST technique has then been further developed in the output feedback case [34]: an interval observer was used in combination with the OIST technique, to ensure that the constraint is still not violated; this new technique was therefore called OISTeR, which means OIST extended with robustness properties (with respect to the uncertain initial state). Figure 3 illustrates this novel framework.

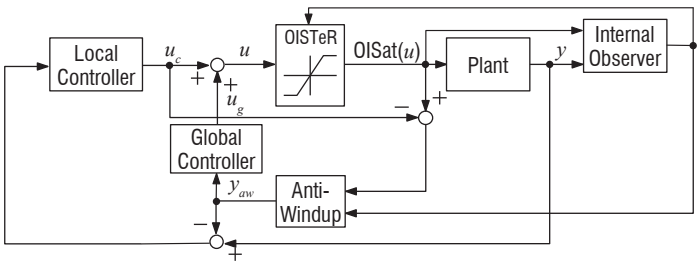


Figure 3 – Model recovery anti-windup architecture associated with OISTeR.

Anti-windup techniques can now be applied to control design problems under both input and output constraints; the major interest of using the anti-windup framework is to apply some very sophisticated robust control design techniques when we are far from the constraints. Anti-windup is now mature enough to take into account input constraints. Then, a novel Output to Input Saturation technique has recently been developed, in order to recast an output-constrained control problem into an input-constrained control problem. Such an innovative technique has been already successfully assessed for several aerospace applications, including:

- the longitudinal control of a large scale long range flexible aircraft under a wing root bending moment constraint [35],
- the problem of satellite attitude reference trajectory tracking under some angular velocity limitations [36],
- the atmospheric flight control of a flexible launch vehicle under an angle of attack constraint [37, 38],
- the obstacle avoidance of a UAV [39].

Future research may extend this technique by considering modeling uncertainties. Moreover, current research is aimed at studying local stability results when the constrained outputs are non-minimum-phase.

Hybrid system control

Aircraft flight control systems implement a gain-scheduled controller switching strategy for linear parameter-varying systems. However, the switching action tends to generate non-smooth control input, which may cause undesired behavior and even instability. For other systems, such as robotics systems, breaking down the overall control task into several simpler ones can simplify the design process; here also, the instantaneous changes in the system dynamics and the rules of the discrete switching logic may result in unexpected system behaviors, or even catastrophic failures, making the analysis and safety verification of the overall system significantly more challenging. Hence, to guarantee safety and to meet the specific performance requirements, these couplings should be properly incorporated into the mathematical representation of the system, necessitating the use of a hybrid system model incorporating discrete and continuous variables. Many of the aeronautic and aerospace vehicles or robotics systems that we are considering show continuous and discrete dynamics interactions and can be seen as hybrid systems. This is the case for:

- self/event-triggered systems, including sample and hold control, quantized control systems, etc. [40, 41, 42],
- switching systems, including systems described by a family of differential equations, combination of local and global controllers, systems with explicit discrete states (or logical modes), hybrid automata, etc. [43, 44],
- "hybridized" systems, including nonlinear models seen as sets of simpler equations, reset controllers, etc. [45, 46].

Hybrid control system methods provide a unique framework to investigate systems with such dynamical behavior. Let us briefly recall that, from control theory point of view, the term hybrid refers to combinations or compositions of continuous and discrete parts and a hybrid dynamical system (or simply a hybrid system) combines behaviors that are typical of continuous-time dynamical systems with behaviors that are typical of discrete-time dynamical systems

(see Figure 4). In this framework, a system can be represented in the following way:

$$\begin{cases} x \in C & \dot{x} = f(x) \\ x \in D & x^{+} = g(x) \end{cases}$$

This representation suggests that the state of the hybrid system, represented by x, can change according to a differential equation $\dot{x} = f(x)$ while in the set C, and it can change according to a differential equation $x^{+} = g(x)$ while in the set D. The notation $\dot{x}$ represents the time derivative of the state x, while x^{+} represents the value of the state after an instantaneous change.

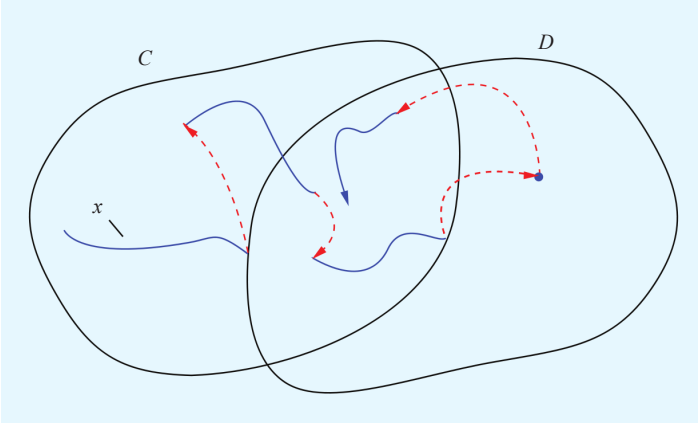


Figure 4 – Hybrid behavior: succession of continuous-time trajectories (blue line) in the set C and discrete transition (red dashed line) in the set D

The modeling and design of the control solutions for hybrid dynamical systems have been widely studied over these last years [44, 47]. They can be used to capture and better understand the behavior of control change between free and constrained movements of a robotic arm (or an autonomous robot moving in a constrained environment) where the switching between sensor-based control laws is stochastic and provided by finite state automata. Moreover, in a context of aeronautic applications, it can be useful to transform, in a more suitable framework, detailed models that include equations, lookup tables and switching logics that are excellent for simulation but not for analysis or controller design. The dynamical behavior of systems like aircraft subject to transitions between different phases of flight (with a gain-scheduled/linear parameter-varying control law and supervisory switching control). In practice, such control laws are often in block diagram form, where time-continuous processes and logical modes are nested and cannot be manually identified. A problem of interest is the development of tools that, based on such block schemes, generate automatically structured formal code like Lustre and Prelude high-level programming languages for real-time embedded applications, allowing the modeling of the overall systems as a hybrid automata: sets of differential equations and switching rules among this model.

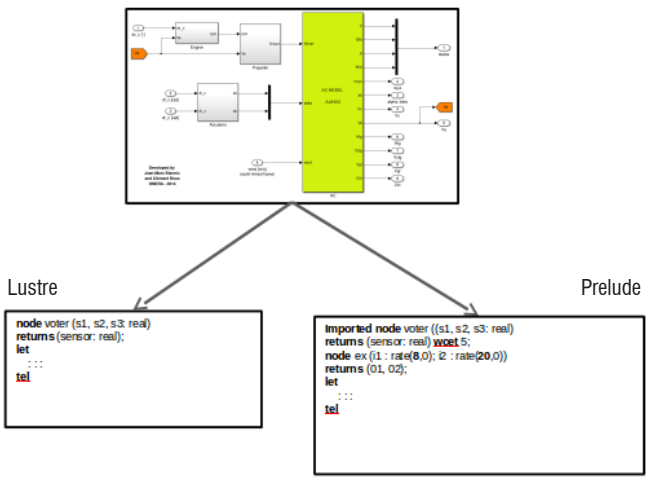


Figure 5 – From Block schemes to structured language

Regarding the analysis of hybrid dynamical systems, switching among simpler dynamical systems or integrator resets has been used successfully in practice for many decades. Recent efforts concentrate on guaranteeing properties such as reachability and stability. Reachability concerns more specifically discrete automata or Petri net states and must assess, among other things, the absence of Zeno execution (infinity of instantaneous transition in a finite delay which is unrealistic for real systems) [48, 44]. The stability of the continuous states of a hybrid system is often carried out using multiple or piecewise Lyapunov functions (see Figure 6) [49, 50]. Such functions can be seen as particular energy functions depending on continuous states of the system. By proving the monotonical decrease of this function along state trajectories (continuous behavior and instantaneous transition), the asymptotic stability of the hybrid system can be assessed.

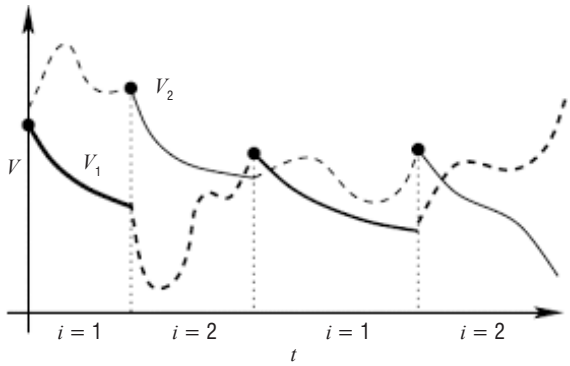


Figure 6 – Multiple Lyapunov functions: a function V_i is used in Logical Mode i

Stability of sets can be considered for systems that include timers, counters, and other discrete states that do not converge. By exploiting such approaches, a major issue for aerospace control engineering is the stability analysis of hybrid dynamical systems in the presence of uncertain inputs and uncertain parameters (see Figure 7).

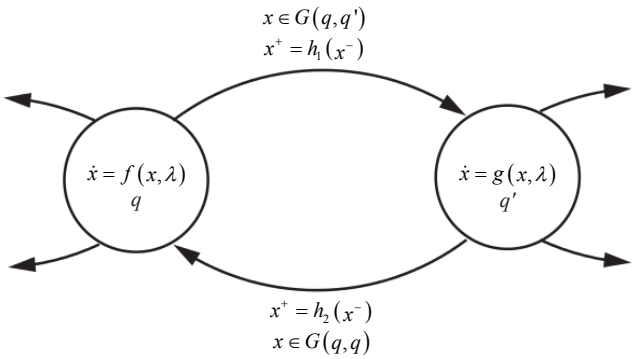


Figure 7 – Hybrid automata with uncertain time-continuous dynamics: Stochastic Parameter λ

The practical stability of such system can be studied with Monte Carlo and Quasi-Monte Carlo methods. Such time consuming sampling methods might be improved by exploiting Polynomial Chaos expansions to deal with uncertainties and properties of Multiple Lyapunov functions.

Hybrid control framework also opens a large field of investigation in terms of synthesis, both for controllers and for the corresponding switching rules. Making hybrid time-continuous controllers to improve performance is another form of investigation. This includes, among others, reset systems where a reset law enriches a nominal controller to improve controller state decrease or L_2 gain [51, 52]. By resetting, under conditions and when necessary, all or part of the controller states, such approaches enable performance improvements without modifying the nominal structure and behavior of the nominal controller [53, 54].

Moreover, the fact that a hybrid dynamical system framework enables the widest representation of systems opens up promising perspectives for Model Predictive Control (MPC). Indeed, such control methods are based on the explicit use of a model of the process (see Section 2) and the richer the model is, the more efficient the control law is (beyond the efficiency of the chosen optimization algorithm, terminal cost, stage cost and prediction horizon). A number of controller design techniques have been proposed recently based on representations relying, in particular, on piecewise-affine (PWA) or mixed logical dynamical (MLD) systems [55, 56]. They can be achieved by formulating a MPC problem and solving it on-line using a mixed-integer quadratic program (MIQP) [57] or a multi-parametric mixed-integer linear program (mp-MILP) and computing a piecewise linear (PWL) optimal control law offline [58].

Beyond the challenge of the hybrid modeling of complex systems, several other issues must be addressed, in order to consider large size problems (in terms both of continuous and discrete states), uncertainties (in terms of inputs and parameters) and computation load (critical for aircraft future configurations, for instance, see Figure 8) specifically encountered in such industrial processes.

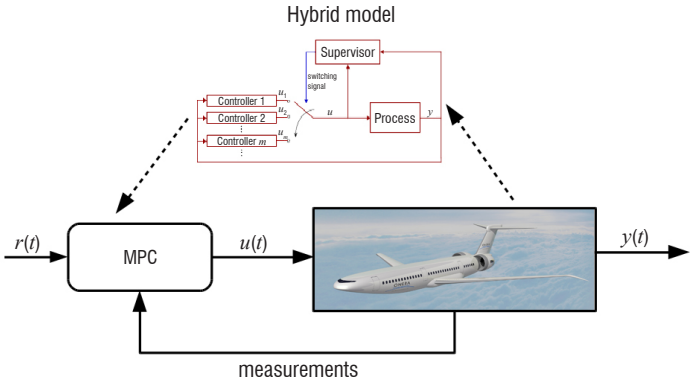


Figure 8 – Predictive control based on the evolution of the hybrid system

Multi-sensor-based control

Sensor-based control provides a framework to control the evolution of systems with embedded sensors in dynamic environments (not known a priori: disturbances, moving obstacles, etc.). Basically, this type of control directly exploits, in a closed loop, physical cues of the environment perceived by sensors to obtain a desired behavior with regard to the environment. The control is defined in the formalism introduced in [59] [60] as the control to zero steady-state error on a time horizon of a function $e(q;t)$ called "task function".

$$e(q;t) = C\left(s\left(r(q;t);t\right) - s(t)\right)$$

The vector $s(r(t))$ is a set of measurements (e.g., for a vision sensor: geometric characteristics of an object) given by a sensor whose configuration in $SE(3)$ is known from a set of state parameters q. The vector $s(t)$ contains the desired value of the features, which can be either constant in the case of a fixed goal, or varying if the task consists in following a specified trajectory. C is a matrix for taking into account a possible redundancy of information.

Vision-based control

We have used this technique recently to deal with under-actuated systems (rotary wing UAVs) equipped with a monocular visual sensor to obtain control laws providing stability and robustness properties. The latter includes local stabilization on visual landmarks [61] [62] or in large displacements for path tracking or moving object tracking [63] [64]. Their design makes no a priori assumption on the environment (with respect to the tracking object configuration, for instance, or type of tracking motion) and by taking into account physical constraints acting on the system (flight envelope, actuator saturation) which makes it particularly nonlinear. Using anti-windup techniques (see Section "anti-windup compensation"), input saturation has been incorporated into the control scheme to implicitly deal with constraints on the output variables. This approach has been used for reactive obstacle avoidance [39] and to permanently maintain the availability of visual cues in the image while performing maneuvers [65] [66].

Figure 9 – Vision-based tracking of an autonomous mobile robot by an AirMax drone equipped with a monocular camera

Multi-task control

Multi-task control (tracking a vehicle with obstacle avoidance, for example, or spatial manipulation under constraints) or control using multimodal exteroceptive sensors can be considered using extensions of sensor-based control technique [67].

Several tasks may be indeed considered simultaneously using a cascade of orthogonal projections in the null-space of tasks (expressed in the form of linear relations between the operational variables and the control variables) of higher priority using the recursive approach

proposed in [68]. Let q be the generalized coordinate vector for the considered system, the multi-task motion control can be obtained from:

$$\ddot{q}_k = \ddot{q}_k - 1 + \left(\boldsymbol{J}_k \boldsymbol{P}_{k-1}\right)\dagger\left(\ddot{x}_k - \dot{\boldsymbol{J}}_k \dot{q}_{k-1} - \boldsymbol{J}_k \ddot{q}_{k-1}\right)$$

where $\ddot{q}_0 = 0$, and where $\boldsymbol{J}_k$ is the Jacobian associated with the task k and $\boldsymbol{P}_k$ is the projector operator in the null-space of the augmented Jacobian matrix for the task k, which can be computed as:

$$\boldsymbol{P}_k = \boldsymbol{P}_{k-1} - \left(\boldsymbol{J}_k \boldsymbol{P}_{k-1}\right)\dagger \boldsymbol{J}_{k-1} \boldsymbol{P}_{k-1}$$

starting from $\boldsymbol{P}_0 = \boldsymbol{I}_d$.

Equality or inequality constraints (such as actuator saturation or obstacles) can thus be placed in a Jacobian of constraints, as proposed in [69]. The various tasks can also be a decoupled by using a consistent pseudo-inverse [70]. Using this algebraic form, inequality constraints can be verified and enforced a posteriori only, thus potentially leading to sub-optimal solutions using, for instance, the Constraint Compliant Control framework [71] [72]. However, once the number of constraints becomes high, this type of method cannot lead to control solutions that can strictly guarantee constraint satisfaction and the use of Quadratic Programming (QP) techniques may thus be considered (see the following section).

The sensor-based control paradigm can be extended to systems with several types of embedded sensors. A multi-sensor approach can either use several identical sensors and exploit their redundancy, or use different sensor types for their complementarity and reduce inaccuracies and uncertainties in the measurements. This may lead to the implementation of fault-tolerant control schemes if a decision layer for detection and reconfiguration is associated with the multi-sensor based controller.

Heterogeneous sensor data can be fed into fusion algorithms (e.g., Kalman or Bayesian methods) to provide state estimation for modeling the environment. However, since these sensors generally measure different physical phenomena, it is preferable to use them directly in the low-level servo controller rather than to apply multisensory fusion, or to design complex state machines [73]. This idea, originally proposed in the hybrid position-force control paradigm [74], can be extended to feedback from multiple sensors. This brings new challenges to controller design (e.g., related to the sensor characteristics (synchronization, hybrid control, task compatibility, etc.) or to the task representation [75]. A matrix C must be defined for this purpose to either decompose the sensor feedback in an orthogonal or reciprocal base and smoothly switching between different sensor feedback (transition between free and constrained motions for instance) or by treating the data coming from different sensors as a unique, higher-dimensional signal.

In another approach, each sensor is given a reference signal and considered as an independent sub-task in the global task function [76]. Control can thus be drawn as a hierarchy between the different sensors to build a control scheme that prevents lower subtasks from disturbing higher ones. This hierarchy can be made dynamic to avoid local minima [77], as suggested above for multi-task control schemes.

Quadratic multi-objective optimization

The control problem for a system whose behavior results from sensor-based task redundancy may be expressed as the problem of finding series of control inputs that will drive the system from an initial state towards several objectives and can be seen as a multi-objective optimization problem that can be treated with optimization techniques, such as a quadratic multi-objective optimization problem under linear constraints (system dynamics) and inequality constraints (control input and other physical constraints) where priorities between the objectives can be dealt with through a strict or soft hierarchy. Assuming a convex solution space, the optimal solution of the control problem lies at the boundary of the feasible (constraint compliant) solution space. Finding the optimal solution thus boils down to finding the active constraint set, *i.e.*, the boundary where it lies. Optimization problem solvers are designed to optimally choose this subset of constraints that should be considered when computing the optimal solution of the control problem.

Obstacle avoidance techniques included in the control law structure based on a Quadratic Program (QP) were investigated initially by [78]. Since then, control approaches relying on optimization tools such as a Linear Quadratic Problem (LQP), enable inequality constraints to be solved more easily and Hierarchical Quadratic Programming (HQP) has been introduced [79] by starting to solve a QP to obtain a solution for a higher priority task and then solving other QP for lower priority tasks, without interfering with any higher priority task.

An alternative solution is to solve a single QP for a prioritized multi-objective problem by introducing a weighting vector influenced by a decision variable [76]. Note that Pareto-optimal solutions can also be sought. Prioritized constraint satisfaction and, if needed, constraint relaxation or uncertainties can be introduced via slack variables to "soften" the constraints and can be considered through setting adequate multiparametric-MINLP or mp-MIQP problems [80]. Smooth task transitions can be easily achieved within a framework using a weighting strategy [76]. More recently, a generalized hierarchical framework that enables both soft and strict priority problems with smooth priority transitions to be described have been proposed [81] [82].

Notice that the problem can also be formulated using Model Predictive Control (MPC) to ease the integration of constraints (through soft-constraints and an exact penalty function to guarantee the problem feasibility) and to explicitly take into account the effect of uncertainty and disturbances on the future evolution of the system (see next Section).

Model Predictive Control techniques in navigation

For the evolution of vehicles in outdoor environments, absolute localization by merging the INS (Inertial Navigation System) and GPS information is often used. It is however subject to a lack of robustness, due to GPS signal masking problems and to multi-path signals encountered in urban environments, or even no signal at all. The localization can be carried out alternatively or additionally by combining measurements of proprioceptive sensors

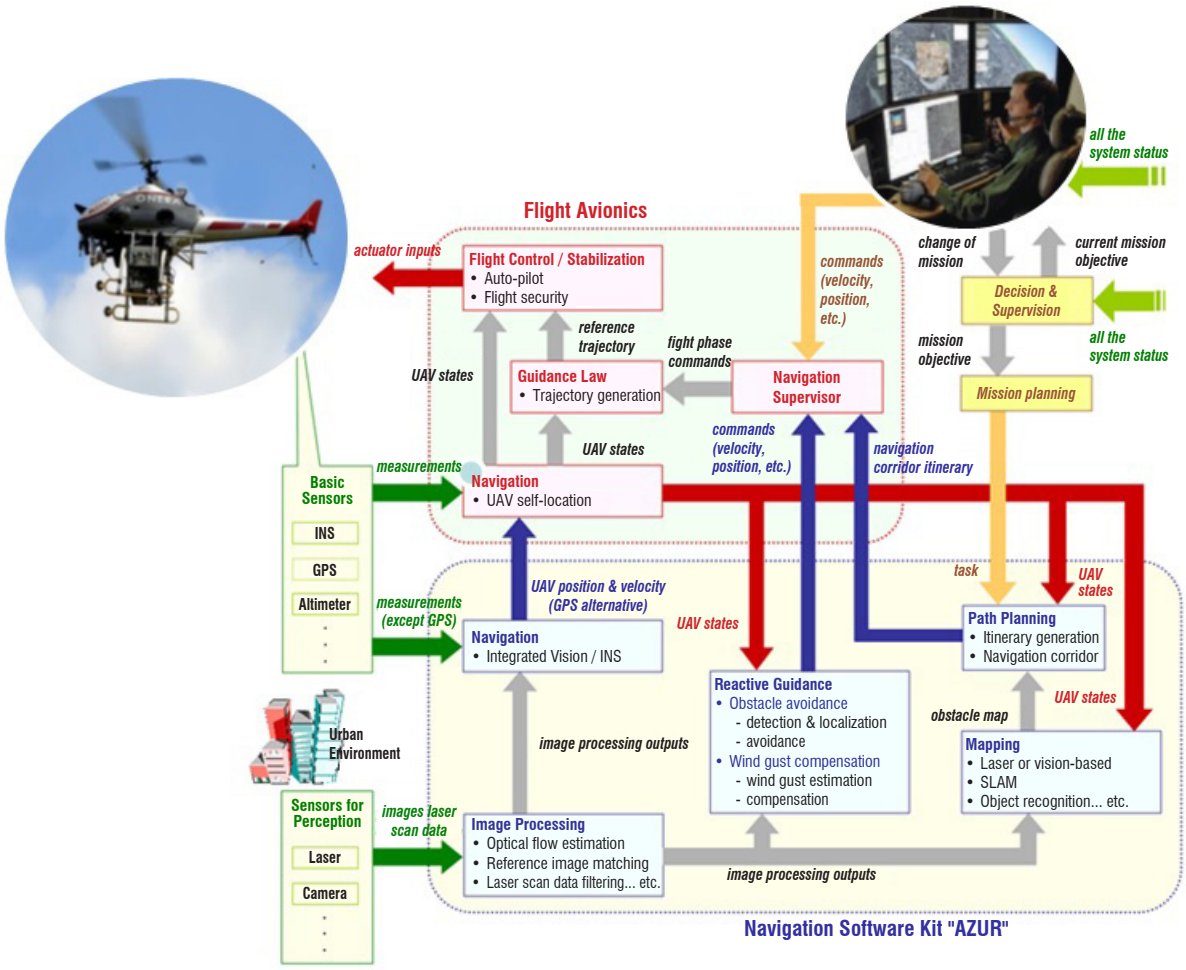


Figure 10 – "AZUR" navigation software kit

(inertial, odometry) and visual cues perceived by active sensors (lidar, camera RGBD) or passive sensors (stereo and monocular cameras). We have developed at ONERA a navigation software kit called *AZUR*, which combines in a modular way several techniques for localization and guidance with a path planner (see Figure 10), which can be said to belong to *map-based navigation techniques*, which rely on the absolute localization on a previously acquired map of the environment. For navigation in large and unknown environments, techniques that simultaneously associate safe exploration/navigation and mapping/self-localization processes automatically and on-line have been investigated for more than a decade. These techniques are grouped under the SLAM (Simultaneous Localization And Mapping) or CML (Concurrent Mapping and Localization) acronyms. They basically rely on an estimation of the mobile sensor configuration and then of the trajectory from noisy sensor data and probabilistic methods. Today, the latter experience a rapid development due to their central role in the development of autonomous vehicles.

Most advanced navigation techniques, map-based navigation and mapless navigation combine a multi-metric representation of the environment in which the robot operates with perceptual information "as they come" (optical flow, feature detection and tracking, environment appearance, and other qualitative information). Thus, with an accurate estimation of the vehicle trajectory, the information gathered on the environment over time can be aggregated into a common reference to build a 3D representation of the scene (stereo depth map or RGBD camera or cloud Lidar points) in which high level semantic information on the objects, such as the dynamic type of the object or object class (car, pedestrian, tree, etc.) can be added. Figure 11 shows the result of a 3D modeling chain developed at ONERA for the real-time environment modeling from embedded vision sensors of autonomous vehicles, which have to be equipped with safe navigation algorithms (obstacle detection and path planning).

Indeed, while the vehicle state is estimated and the environment modeled, navigation algorithms must run to plan and execute optimal paths to fulfill a mission, which can, for instance, be to explore the environment while avoiding obstacles (*i.e.*, to perform an active vision function). The mission objectives can be translated for this into various cost functions expressed in terms of control variables. Based on these principles, we have developed a whole set of autonomous safe navigation algorithms in dynamic environments which have been grouped in a toolbox call COPERNIC (w3.onera.fr/copernic). Many of the techniques take advantage of Model Predictive Control principles.

MPC principles as a general framework for navigation

MPC is an effective means to deal with a multi-variable constrained control strategy for which the issues of real-time implementation, stability and performance are well understood for linear systems. Moreover, much progress have been made over the last decade with regard to non-linear systems, with model uncertainties [83], as well as with regard to systems that are subject to a large set of constraints [84].

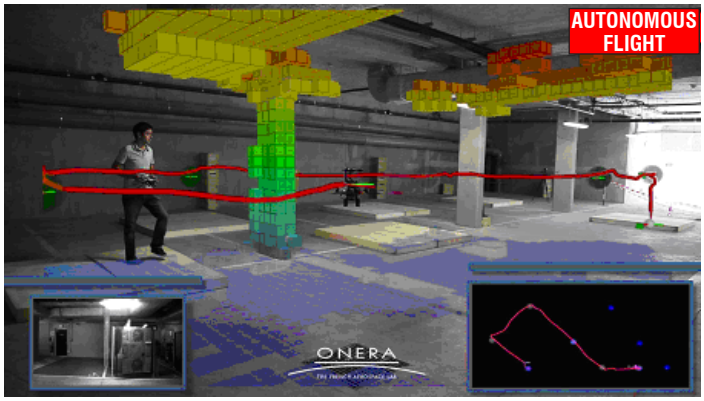


Figure 11 – 3D modeling of an indoor environment based on stereo images where the voxel color code displays the height information

The MPC problem for a discrete time model of a LTI system can be formalized as:

$$x(k+1) = A_x(k) + Bu(k)$$

where $x(k) \in R^n$ et $u(k) \in R^m$ denote the state and control input respectively [83]. The predictive control feedback law is computed by minimizing a predicted performance cost, which is defined in terms of the predicted sequences of state and input predictions over prediction horizons. This predicted cost can be set in the following quadratic form:

$$J_{(p,m)} = \underset{u(.)}{\text{minimize}} \left[x(p)^t \boldsymbol{P}_0 x(p) + \sum_{i=0,(p-1)} x(i) \boldsymbol{Q} x(i) + \sum_{i=1,(m-1)} \boldsymbol{R} u(i) \right]$$

subject to

$$Ex + Fu \leq \Phi$$

where p and m denote respectively the prediction and the control horizons with $p \geq m$. $\boldsymbol{P}$, $\boldsymbol{Q}$, $\boldsymbol{R}$ are positive definite matrices ($\boldsymbol{Q}$ may be positive semi-definite). The origin of the problem ($x = 0$, $u = 0$) is assumed to be in a feasible region. Once the minimizing control sequence is determined, a receding horizon scheme proceeds by implementing the first control $u^*_{(p,m)}(0|x(k))$ to obtain $Ax(k) + Bu^*_{(p,m)}$, the rest of the control sequence being used to update the problem. When the constraints are linear, the convex optimization problem of the objective function under constraints, which therefore has a unique solution, requires Quadratic Program (QP) to be solved.

Fast and reliable solvers based on interior point or active set methods are today available for solving QP problems in real-time, and parallelizable forms exists (PQP) that can readily exploit the full parallelism of multiprocessor machines, including multi-core, SIMD, and GPU [85]. MPC may be used in various control problems as discussed in the previous sections. In particular, we used this problem formulation in navigation algorithms. Based on a prediction of the system behavior, a multi-objective performance criterion is optimized at each time-step for computing control inputs that achieve the required goals. Unlike most path planning methods, this scheme allows the dynamics of the system to be adequately described using model reduction techniques, for instance, as well as environment changes, since new control inputs are computed on the basis of measurements acquired in real time. Moreover, MPC translates the preview over a future horizon of the consequences of the control on the system state in the sense of a certain cost while satisfying the constraints for producing emergent behaviors.

Example of emergent behaviors

Using this MPC framework, a first contribution has been made to 3D trajectory tracking with obstacle avoidance for UAVs [86] defining a multi-objective control problem. The latter consists in finding, under constraints, the desired sub-task parameters (a set of gains of PD-feedback elementary controllers) over a preview horizon N that minimizes the sub-task errors; *i.e.*, $\|\hat{x} - \hat{x}^r\|$ according to the measured state $(x_k, \dot{x}_k)$, at the time period k, the optimization problem:

$$\underset{K_{k+N|k}}{\text{minimize}} \sum_{i=1,N} \omega_m \left\| \hat{x}_{k+i|k} - \hat{x}^r_{k+i|k} \right\|^2 + r^m_{k+i|k}$$

while considering the system dynamics to satisfy which can be expressed as equality constraints. $K_{k+N|k}$ denotes the horizon of the task parameters. The terms r^m are regularization costs introduced to limit the variations of the optimization variables.

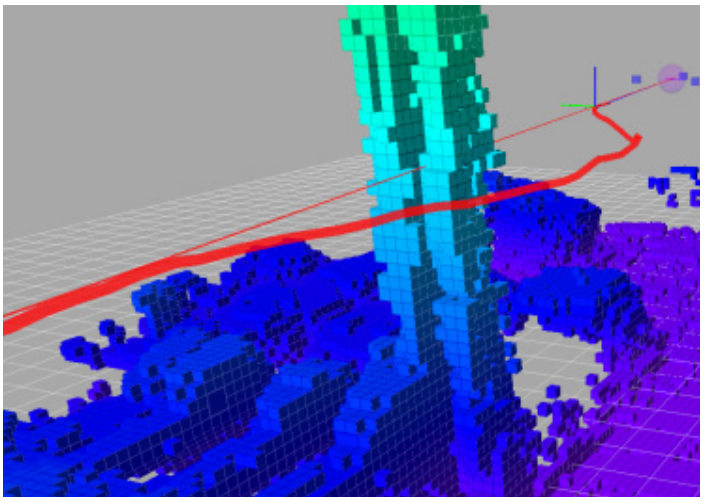


Figure 12 – Obstacle avoidance trajectory by a UAV from a real-time reconstruction of the 3D environment using a predictive model of the vehicle and a multi-objective optimization for its motion control

Several types of behaviors have been explored for a fleet of autonomous robots equipped in the experimental setup with stereo vision sensors by using this general framework [87].

- Exploration strategies of a finite environment have been experimented on using a weighted cost function combining a cost of navigation, which allows the regulation of the speed of vehicles and the control of their travel to waypoints to be minimized, a security cost whose cost minimization allows obstacle and collision avoidance between vehicles and a cost on the energy consumption [87].

- Another contribution [88, 89] relies on the resolution of the optimization problem (non-linear and constrained) by a discretization and exploration of the control space to complete the exploration of a defined area, considering that each point of the zone was visited by at least one of the vehicles in the fleet. The algorithm has the advantage of ensuring a constant computation time at each iteration, to be potentially less sensitive to local minima and to optimizer initialization. It has been successfully implemented on mobile robots and UAVs for safe exploration of a cluttered environment [90].

- Source localization strategies within a pre-defined area by a swarm of robotic systems have been designed in a distributive way and have shown high efficiency in terms of time spent and robustness to vehicle and communication failures [131].

Other emergent behaviors for swarms of autonomous vehicles can be explored on the basis of Distributed Model Predictive Control (DMPC) [91], this by exploiting MPCs on local sub-systems and exchanging predictions so as to coordinate with each other, the whole system convergence and stability being ensured by consensus mechanisms [92, 93, 94] or by using structured optimization techniques (see the next section).

The potential offered by the various DMPC strategies is quite broad and could constitute a general framework for implementing agile autonomous sensor networks (embedded on autonomous robots) whose implementation would face a set of practical difficulties, such as the existence of possible communications between the distributed sensors, uncertainty on the measurements collected and the environment, and the possible loss of observability depending on trajectories, taking into account strapdown constraints [95, 96]. Based on such a framework, we can, for instance, benefit from the information redundancy to increase the localization robustness by fusing individual localizations or, in a more integrated way, by merging information collected on the environment from various sensors in a distributed filter architecture, a problem which is known as: distributed localization and mapping [97, 98, 99]. Collaborative SLAM (CoSLAM) [100, 88, 95], considering non-coordinated vehicles sharing their SLAM information under communication constraints can also be investigated on the basis of such a framework.

Structured Distributed Control

It is possible to go further in the integration of subsystem MPCs, while controlling the overall system performance/stability by the use of structured optimization methods to compute a distributed control law. Distributed control is introduced as an alternative architecture to centralized and decentralized control [91]. Distributed approaches suggest breaking down the system into autonomous sub-agents organized into an information exchange network, in order to reduce the drawbacks of a centralized architecture by taking advantage of the sparsity of large-scale systems where the interaction between subsystems can generally be reduced to direct neighbors. Setting up distribution techniques nevertheless requires the coordination problem of the several local controllers to be addressed. In this perspective, Model Predictive Control is particularly relevant since it provides a temporal window – the preview horizon – to exploit in order to establish a coordination strategy.

Distributed architectures rely on a decomposition of the system into subsystems of lower dimension, and single out from decentralized approaches by the coordination of sub-agents. The table in Figure 13 presents an optimal control architecture nomenclature in the case of two coupled sub-systems [101]. This coordination is aimed at bringing the distributed optimization toward a collective optimum (in the sense of Pareto), in the case of cooperative algorithms. It may be viewed as a class of sub-optimal control. Conversely, non-cooperative algorithms imply the respective optimization from each sub-agent of a local cost function, whose evolution is also subject to known actions from the other sub-systems. As a result, actions are taken individually, with each controller being aimed at accepting a change in a variable under the sole condition of a local benefit, and thus tending to produce solutions drawing the system towards a Nash equilibrium. This equilibrium is a stable state around which one or all system(s) would have a handicap with respect to its or their respective local objective.

Input variables	$\boldsymbol{u}^{(1)} \in K_u^{(1)}$	$\boldsymbol{u}^{(2)} \in K_u^{(2)}$
Cost functions	$\overline{g}^{(1)}(\boldsymbol{u}^{(1)}), g^{(1)}(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)}), \overline{g}^{(2)}(\boldsymbol{u}^{(2)}), g^{(2)}(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)})$ $g(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)}) \triangleq \omega_1 g^{(1)}(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)}) + \omega_2 g^{(2)}(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)})$	
Decentralized	$\min_{\boldsymbol{u}^{(1)} \in K_u^{(1)}} \tilde{g}^{(1)}(\boldsymbol{u}^{(1)})$	$\min_{\boldsymbol{u}^{(2)} \in K_u^{(2)}} \tilde{g}^{(2)}(\boldsymbol{u}^{(2)})$
Non-cooperative distributed (Nash)	$\min_{\boldsymbol{u}^{(1)} \in K_u^{(1)}} g^{(1)}(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)})$	$\min_{\boldsymbol{u}^{(2)} \in K_u^{(2)}} g^{(2)}(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)})$
Cooperative distributed (Pareto)	$\min_{\boldsymbol{u}^{(1)} \in K_u^{(1)}} g(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)})$	$\min_{\boldsymbol{u}^{(2)} \in K_u^{(2)}} g(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)})$
Centralized	$\min_{(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)}) \in K_u^{(1)} \times K_u^{(2)}} g(\boldsymbol{u}^{(1)},\boldsymbol{u}^{(2)})$	

Figure 13 – Distributed control architecture nomenclature for two coupled sub-systems (from [102])

Problems for which decomposition (in which variables can be partitioned into sub-vectors and each constraint involves only variables from one of the sub-vectors) lead evidently to the ability of solving each problem separately (and in parallel), and then re-assembling the solution. A more interesting situation occurs when there is some coupling or interaction between the sub-vectors, so the problems cannot be solved independently. For these cases, such decomposition techniques can be used to solve the overall problem by iteratively solving a sequence of smaller problems, including the dual decomposition technique [103].

This approach makes it possible to strictly decompose the multi-objective MPC problem into coupled and conflicting sub-problems. The coordination problem is translated into a non-cooperative game problem, providing the control architecture with a gain in modularity that allows approximations on the couplings between sub-systems to be naturally enforced. A parallel algorithm can then be used to solve the resulting distributed control problem as a set of sub-problems of various time scales and approximation levels, able to successfully solve conflicts between objectives [102].

DMPC also offers a set of possibilities for the implementation of Fault-Tolerant Control (FTC) laws in case of sensor and actuator fault detection, in order to maintain the desired closed-loop performance [104]. In [105], a data-based monitoring and reconfiguration system was developed for a distributed model predictive control system in the presence of control actuator faults. In addition to a monitoring method, appropriate DMPC reconfiguration (fault-tolerant control) strategies were designed to handle the actuator faults and maintain the closed-loop system state within a desired operating region.

Coordination strategy with mixed-integer-continuous variables

Coordination strategies between autonomous vehicles can also involve discrete and continuous variables. The use of mixed-integer variables allows discrete actions to be described in the model and non-convex admissible domains to be handled. The mixed-integer MPC formulation thus allows discrete actions to be drawn at the control layer, which are commonly supported by the decision layer, and thus enables a more consistent coordination of discrete events. Mixed-Integer Programming (MIP) [57] indeed proposes an ideal framework to account for logic events and non-convex admissible domains. MIP approaches generally use integer variables to describe logic and combinatorial systems, with these variables being used either as discrete state/input values or as triggers to activate and deactivate constraints in the optimization problem. The latter use offers opportunities to regard non-convex admissible domains as an arrangement of convex regions, an integer variable specifying in which convex region the now convex problem is currently being considered. While integer programs are NP-hard, efficient algorithms and solvers are available for specific classes of programs, exploiting the form of the optimization problem or employing heuristics. For example, quadratic MIPs (MIQPs) can be solved using branch and bound algorithms which solve a graph of QP problems resulting from the con-sideration of integer variables as real variables [106]. Given that real-valued QP problems are convex, the resolution of these relaxed sub-problems is computationally efficient.

Dynamic Adaptive Planning

At the upper level of the control hierarchy, the role of the decision layer is to produce high-level activity plans composed of sequences of elementary tasks to be fulfilled by a system whose control would be performed at a task-level by making use of some of the techniques discussed before. Such plans are especially useful for complex missions requiring high-level goals to be achieved, which could be, for example, the "long-term surveillance of critical areas". High-level goals of this type are hardly satisfied by using only low-level reactive strategies, which is the reason why a deliberative planner comes into play [107]. Several toolboxes have been developed at ONERA over this last decade to solve planning and scheduling problems, such as *InCELL* [108], *HiPOP* [109] or *CPT* [110].

As an input, the planner considers a symbolic representation of the current state of the system (e.g., a Boolean component failure status), a symbolic representation of the environment (e.g., a waypoint graph for modeling trajectories on a complex terrain), goals that must be fulfilled (tasks to be performed, states to be reached, etc.), constraints that must be met (temporal constraints, resource constraints, etc.), optimization criteria (mission duration, action costs, etc.), and simplified equations of the dynamic behavior (e.g., navigation durations modeled as arc weights in the waypoint graph).

The planner uses these inputs to produce a plan. In practice, it must also be able to cope with the uncertainty about the current state of the system and about possible future evolutions of the environment, and with dynamic missions in which the set of high-level goals provided can be updated following new event detections or new operator requests. Two main approaches are classically used in such contexts:

- *Reactive planning*, which consists in building an initial plan based on arbitrary deterministic assumptions and in performing online replanning or online plan repairs each time that a new relevant item of data coming from the lower levels of control or from the perception modules is received [111]. The integration of this planning level with the task control level is tricky. The necessary dialog between action and cognition is a particularly complex issue. Receding horizon optimization methods, such as those discussed in the previous section, may help to anticipate the effects of actions and produce local adaptations, for instance by using Mixed-Integer Linear Programming (MILP) to incorporate kino-dynamic, obstacle avoidance and collision avoidance constraints, as proposed in [112] or Mixed-Integer Quadratic Programming (MIQP) to reach complex adaptive behavior [113]. This dialog must be supported by a common representational medium, such as that proposed by the Theory of Event Coding (TEC), in which perceptual contents and action plans are coded for an adequate theoretical treatment of perception and action planning [114].

- *Proactive planning*, which consists in producing offline a decision strategy covering several situations; such a strategy might be represented as a *decision policy* (mapping from observable state values to actions), as a *conditional plan* (a plan with branches), or as a *robust plan* by explicitly taking into account the effect of uncertainty and *constraint satisfaction* using for instance adaptations of the MPC formulation [115].

See Figure 14 for an illustration of these concepts. For more detailed examples of reactive planning, see [116] for an application to a multi-robot area surveillance mission and [117] for an application to an autonomous satellite surveillance mission. On the proactive side, see [118] for the production of decision policies for a UAV and [119] for the production of temporally flexible plans for multi-robot deployment.

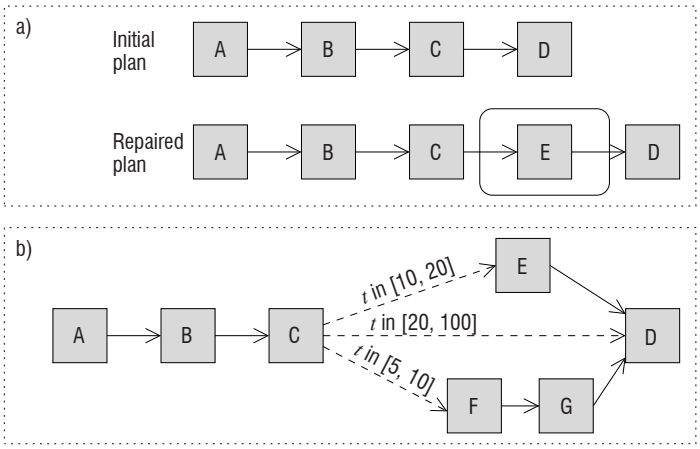


Figure 14 – (a) Reactive planning transforming an online plan containing a sequence of four tasks into a new plan containing one more task. (b) Proactive conditional plan (dashed lines correspond to conditional branches depending on an observable state parameter t and affecting the performance of tasks E, F, G)

Reactive and proactive planners can be based on various resolution methods, such as PDDL-like deterministic and non-deterministic planning algorithms [107], scheduling algorithms [120], constraint programming [121], Markov Decision Processes [122], Simple Temporal Networks with Uncertainty [123], etc. In each case, complete or approximate search techniques can be used (exhaustive tree search, local search, global optimization, dynamic programming, etc.).

In the following, we describe research directions deserving to be explored to achieve increasing system autonomy at the planning level. These directions are summarized in Figure 15.

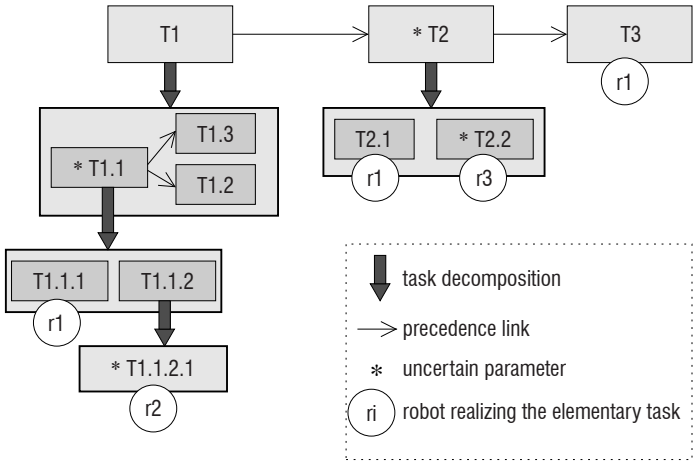


Figure 15 – Hierarchical distributed plans with uncertain parameters and temporal constraints (other constraints, such as resource constraints, are not represented, nor the planning criteria)

Hierarchical deliberations

In the case of reactive planning, the first idea is to push *hierarchical deliberations* further, which produce plans containing high-level tasks decomposed into lower level tasks, such as in *Hierarchical Task Networks* (HTN [124]). With regard to HTN, a major objective is to mix hierarchical representations with scheduling models containing complex temporal constraints among tasks, complex resource constraints, and complex criteria. Another objective is to refine physical models used at the lower levels of the hierarchy (e.g., navigation models). These lower levels could for instance embed Model Predictive Control techniques to obtain a more accurate view of the impact of high-level decisions. Such an integration with MPC would be quite natural since reactive planning and MPC are based on similar principles (search for a control strategy over a finite horizon using deterministic modeling of future evolutions). Note also that using hierarchical representations enables for more modularity in mission modeling, for shorter computation times, and for more readability from an operational point of view [109].

Real-time planning strategies

From the online reasoning side, a second goal is to explore *real-time planning* strategies more in depth. On this point, hierarchical deliberations pave the way for obtaining a continuum between conservative high-level coarse-grain reasoning when quick deliberations are required, and fine-grain reasoning at lower levels of the hierarchy when more computation time is available. Hierarchical deliberations can also help to make plan repairs more local, leading to increased plan stability [125]. Another advantage of hierarchical representations is that they can help to handle the tricky issues classically encountered when combining online reactive planning and real-time execution [126] (action interruption issues, state projection issues, plan concatenation issues, etc.). For example, reactive planners need an initial state as an input and, with current approaches, this initial state is very often either the state obtained by interrupting all pre-emptible actions, or the state obtained by finishing all ongoing actions. Having low-level tasks embedding MPC techniques could help in considering intermediate states obtained in the course of actions, which could lead to more fluidity in the real-time control of the system.

Distributed plans

In another direction, a third objective is to push further *distributed plans*, which are needed for missions involving multiple agents that must be coordinated for achieving common high-level goals, such as robots performing acquisition tasks and robots deployed to establish an *ad hoc* communication network. Concerning distributed plans, our ambition is to develop generic strategies combining centralized reasoning for high-level decisions, like task allocation among agents, and decentralized plan repair for low-level decisions referring to a single agent or team of agents. When using distributed plans, one major difficulty is also the outage of communication links between agents, which is why the plans built in a centralized way must ideally be robust to the absence of communication links (*multi-agent dynamic controllability* issues [127]). Another difficulty is that operators supervising the mission must have some global situation awareness, which is why the generic schemes developed should explicitly describe by which agent a particular decision can be made. As an illustration, see [128] for a space mission where robust plans sent to satellites are built at a centralized mission center, while opportunistic reactive planning is performed by decentralized satellite on-board reasoning. Note that since we target missions in which agents all share the same high-level goals, using multiagent negotiation schemes is less relevant.

Uncertainty management

Lastly, some of the planners that we develop are built upon constraint-based optimization models solved using local search and global optimization strategies [129, 108]. Using such planners for reactive planning is quite natural, but there is still some work to be done before an efficient adaptation to proactive planning can be obtained, e.g., for producing plans maximizing an expected reward or minimizing a risk-level as in *Online Combinatorial Stochastic Optimization* (OSCO [130]). For this, we would like to explore new combinations between constraint-based scheduling on the one hand and stochastic uncertainty reasoning or set uncertainty reasoning on the other hand. Finally, uncertainty management should be combined with all of the hierarchical and distributed aspects mentioned above.

Conclusion

Functional and decisional autonomy in automated systems is at the core of research activities for designing future aircraft and spacecraft, as well as all systems for mobility in the fields of transportation,

the environment, security, defense, etc. The control architectures of these autonomous systems must meet requirements in terms of performance that are different in nature (robustness to disturbances, adaptation to non-deterministic and varied functions, intersystem and man/system interaction capabilities, etc.) while satisfying security constraints and the reliability required for their implementation.

Several advances, at different levels of these control architectures, are discussed in this paper. Robust control techniques should allow the various physical constraints imposed on scalable systems to be addressed as well as dealing with sequences of control modes. In addition, we have considered evolutions on sensor-control techniques for performing hybrid tasks (combining several types of sensors), or multiple tasks (multiple functional objectives). Model-based control techniques can contribute greatly to complex control problems for guidance/navigation and coordination of dynamic systems under constraints, especially in view of a real-time implementation of optimization methods on the novel hardware architectures of embedded computers. Finally, considering that the reactive and deliberative control levels have to be more integrated in these control architecture, we have discussed the need for hierarchical and/or distributed decision in the context of real-time planning with uncertainties ■

References

[1] A. TATE - *Intelligible AI Planning - Generating Plans Represented as a Set of Constraints*. Springer London, London, p. 3-13, 2001.

[2] V. WIELS, R. DELMAS, D. DOOSE, P.-L. GAROCHE, J. CAZIN, and G. DURRIEU - *Formal Verification of Critical Aerospace Software*. AerospaceLab Journal Issue 4, p. 526-531, 2013.

[3] M. SANFOURCHE, V. VITTORI, and G. LE BESNERAIS - *EVO: A realtime Embedded Stereo Odometry for MAV Applications*. 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, p. 2107-2114, Nov 2013.

[4] M. SANFOURCHE, B. LE SAUX, A. PLYER, and G. LE BESNERAIS - *Environment Mapping Amp; Interpretation by Drone*. 2015 Joint Urban Remote Sensing Event (JURSE), p. 1-4, March 2015.

[5] G. FERRERES and J. M. BIANNIC - *Smt : The Skew μ Toolbox (a Matlab Toolbox for Robustness Analysis)*. SMT: The skew Mu Toolbox, http://www.onera.fr/staff/jean-marc-biannic, 2004.

[6] J.M. BIANNIC and C. ROOS - *Flight Control Laws : Recent Advances in the Evaluation of their Robustness Properties*. AerospaceLab Journal, 4, 2012.

[7] C. ROOS and J.M. BIANNIC - *Efficient Computation of a Guaranteed Stability Domain for a Figh-Order Parameter Dependent Plan*. American Control Conferencee, Baltimore, Maryland, USA, 30 06(2 07):3895-3900, 2010.

[8] C. ROOS, F. LESCHER, J.-M. BIANNIC, C. DLL and G. FERRERES - *A Set of Analysis Based Tools to Evaluate the Robustness Properties of High-Dimensional Uncertain Systems*. 2011 IEEE International Symposium on Computer- Aided Control System Design (CACSD), p. 644-649, Sept 2011.

[9] J.-M. BIANNIC and C. ROOS - *Enhanced Mu-Analysis Techniques for Clearance of Flight Control Laws*. Final COFCLUO Workshop. Also published in a Springer Book, LNCIS series, 416:131-147, 2010.

[10] C. ROOS - *A Practical Approach to Worst-Case h x221e; Performance Computation*. 2010 IEEE International Symposium on Computer-Aided Control System Design, p. 380-385, Sept 2010.

[11] C. ROOS and J.-M. BIANNIC - *Robustness Analysis Versus Mixed LTI/LTV Uncertainties for On-Ground Aircraft*. Springer Berlin Heidelberg, Berlin, Heidelberg, p. 175-194, 2007.

[12] G. FERRERES and C. ROOS - *Robust Feedforward Design in the Presence of LTI/LTV Uncertainties*. International Journal of Robust and Nonlinear Control, 17(14):12781293, 2007.

[13] J.-M. BIANNIC, C. PITTET, L. LAFOURCAD, and C. ROOS - *LPV Analysis of Switched Controllers in Satellite Attitude Control Systems*. AIAA Guidance, Navigation, and Control Conference, Toronto, Canada, 4, 2010.

[14] J. MACIEJOWSKI - *Predictive Control with Constraints*. Prentice Hall, 2001.

[15] S. J. QIN and T. A. BADGWELL - *A Survey of Industrial Model Predictive Control Technology*. Control Engineering Practice, 11(7):733 - 764, 2003.

[16] S. GALEANI, S. TARBOURIECH, M. TURNER, and L. ZACCARIAN - *A Tutorial on Modern Anti-Windup Design*. European Journal of Control, 15(34):418-440, 2009.

[17] S. TARBOURIECH and M. TURNER - *Anti-Windup Design: an Overview of Some Recent Advances and Open Problems*. Control Theory Applications, IET, 3(1):1-19, 2009.

[18] D. Q. MAYNE, S. V. RAKOVI, R. FINDEISEN, and F. ALLGWER - *Robust Output Feedback Model Predictive Control of Constrained Linear Systems*. Automatic, 42(7):1217-1222, 2006.

[19] R. FINDEISEN, R. ALLGOWER, and L. BIEGLER - *Assessment and Future Directions of Nonlinear Model Predictive Control*. Lecture Notes in Control and Information Sciences, Springer Verlag, Berlin, 2008.

[20] M. ALAMIR - *A Framework for Real-Time Implementation of Low-Dimensional Parameterized NMPC*. Automatica, 48(1):198-204, 2012.

[21] P. APKARIAN and D. NOLL - *Nonsmooth H-infinity Synthesis*. IEEE Transactions on Automatic Control, 51(1):71-86, Jan 2006.

[22] P. APKARIAN, M. N. DAO, and D. NOLL - *Parametric Robust Structured Control Design*. IEEE Transactions on Automatic Control, 60(7):1857-1869, July 2015.

[23] J. LESPRIER, J.-M. BIANNIC, and C. ROOS - *Nonlinear Structured H-Infinity Controllers for Parameter-Dependent Uncertain Systems with Application to Aircraft Landing*. 2014 IEEE Conference on Control Applications (CCA), p. 433-438, Oct 2014.

[24] A. R. TEEL and N. KAPOOR - *Uniting Local and Global Controllers*. European Control Conference, Brussels, Belgium, July 1997.

[25] A. R. TEEL and N. KAPOOR - *The l2 Anti-Windup Problem: Its Definition and Solution*. Control Conference (ECC), 1997 European, p. 1897-1902, July 1997.

[26] J.-M. BIANNIC and P. APKARIAN - *Anti-Windup Design via Nonsmooth Multiobjective H-Infinity Optimization*. Proceedings of the 2011 American Control Conference, p. 4457-4462, June 2011.

[27] J.-M. BIANNIC - *Limit-Cycles Prevention via Multiple H-Infinity Constraints with an Application to Anti-Windup Design*. 9th IFAC Symposium on Nonlinear Control Systems, IFAC Proceedings Volumes, 46(23):640-645, 2013.

[28] S. FORMENTIN, F. DABBENE, R. TEMPO, L. ZACCARIAN, and S. M. SAVARESI - *Robust Linear Static Anti-Windup with Probabilistic Certificates*. To appear in IEEE Transactions on Automatic Control.

[29] C. ROOS and J.-M. BIANNIC - *Aircraft-On-Ground Lateral Control by an Adaptive LFT-Based Anti-Windup Approach.* IEEE Conference on Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control, p. 2207-2212, Oct 2006.

[30] G. C. GOODWIN, M. M. SERONK, and J. A. DE DONA - *Constrained Control and Estimation: an Optimization Approach*. Springer, 2004.

[31] A. H. GLATTFELDER and W. SCHAUFELBERGER - *Control Systems with Input and Output Constraints*. Springer, 2003.

[32] L. BURLION - *A New Saturation Function to Convert an Output Constraint into an Input Constraint*. Control Automation (MED), 2012 20th Mediterranean Conference on, p. 1217-1222, 2012.

[33] E. CHAMBON, L. BURLION, and P. APKARIAN - *Time-Response Shaping Using Output to Input Saturation Transformation*. International Journal of Control, 2017.

[34] E. CHAMBON, L. BURLION, and P. APKARIAN - *Robust Output Interval Constraint Using O/I Saturation Transformation with Application to Uncertain Linear Launch Vehicle*. Control Conference (ECC), 2015 European, p. 1802-1807, 2015.

[35] L. BURLION, C. POUSSOT-VASSAL, P. VUILLEMIN, M. LEITNER, and T. KIER - *Longitudinal Manoeuvre Load Control of a Flexible Large-Scale Aircraft*. 19th IFAC World Congress, IFAC Proceedings Volumes, 47(3):3413-3418, 2014.

[36] L. BURLION, J.M. BIANNIC, and T. AHMED-ALI - *Interval Based Robust Control of a Flexible Spacecraft Under Angular Velocity Constraints*. Submitted.

[37] E. CHAMBON, L. BURLION, and P. APKARIAN - *Robust Output Interval Constraint Using O/I Saturation Transformation with Application to Uncertain Linear Launch Vehicle*. Control Conference (ECC), 2015 European, p. 1802-1807, 2015.

[38] E. CHAMBON - *Frequency- and Time-Domain Constrained Control of Linear Systems Application to a Flexible Launch Vehicle*. PhD thesis, University of Toulouse, 2016.

[39] C. CHAUFFAUT, F. DEFAY, L. BURLION, and H. DE PLINVAL - *UAV Obstacle Avoidance Scheme Using an Output to Input Saturation Transformation Technique*. 2016 International Conference on Unmanned Aircraft Systems (ICUAS), p. 227-234, June 2016.

[40] R. POSTOYAN, P. TABUADA, D. NESIC, and A. ANTA - *A Framework for the Event-Triggered Stabilization of Nonlinear Systems*. IEEE Transactions on Automatic Control, 60(4):982-996, 2015.

[41] A. SEURET, C. PRIEUR, and L. L. ZACCARIAN - *LQ-Based Event-Triggered Controller Co-Design for Saturated Linear Systems*. Working paper or preprint, 2016.

[42] D. LIBERZON - *Hybrid Feedback Stabilization of Systems with Quantized Signals*. Automatica, 39(9):1543-1554, 2003.

[43] J. LYGEROS, K. H. JOHANSSON, S. N. SIMIC, J. ZHANG, and S. S. SASTRY - *Dynamical Properties of Hybrid Automata*. IEEE Transactions on Automatic Control, 48(1):2-17, 2003.

[44] R. GOEBEL, R. G. SANFELICE, and A. R. TEEL - *Hybrid Dynamical Systems*. IEEE Control Systems Magazine, 29(2):28-93, April 2009.

[45] O. BEKER, C. V. HOLLOT, and Y. CHAIT - *Fundamental Properties of Reset Control System*. Automatica, 40:905-915, 2004.

[46] D. NESIC, L. ZACCARIAN, and A. R. TEEL - *Stability Properties of Reset Systems*. Automatica, 44, 8:2019-2026, August 2008.

[47] R. GOEBEL, J. HESPANHA, A. R. TEEL, C. CAI, and R. G. SANFELICE - *Hybrid Systems: Generalized Solutions and Robust Stability*. Proc. IF AC: Symp on Nonlinear Control Systems, Stuttgart, Germany, p. 1-12, September 2004.

[48] K. H. JOHANSSON, J. LYGEROS, S. SASTRY, and M. EGERSTEDT - *Simulation of Zeno Hybrid Automata*. Proceedings of the 38th IEEE Conference on Decision and Control, volume 4, p. 3538-3543, December 1999.

[49] M. JOHANSSON and A. RANTZER - *Computation of Piecewise Quadratic Lyapunov Functions for Hybrid Systems*. IEEE Transactions on Automatic Control, 43(4):555-559, April 1998.

[50] M. S. Branicky - *Multiple Lyapunov Functions and Other Analysis Tools for Switched and Hybrid Systems.* IEEE Transactions on Automatic Control, 43(4):475-482, Apr 1998.

[51] T. LOQUEN, S. TARBOURIECH, and C. PRIEUR - *Stability of Reset Control Systems with Nonzero Reference*. Proceedings of the 47th IEEE Conference on Decision and Control, Cancun, Mexico, p. 3386-3391, December 2008.

[52] G. WITVOET, W. H. T. M. AANGENENT, W. P. M. H. HEEMELS, M. J. G. VAN DE MOLENGRAFT, and M. STEINBUCH - *H2 Performance Analysis of Reset Control Systems*. Proceedings of the 46th IEEE Conference on Decision and Control, New Orleans, LA, USA, p. 3278-3284, December 2007.

[53] J. CARRASCO, A. BANOS, and A. BARREIRO - *Stability of Reset Control Systems with Inputs*. 16th Mediterranean Conference on Control and Automation, p. 1496-1501, June 2008.

[54] J. BAKKEHEIM, T. A. JOHANSEN, O. N. SMOGELI, and A. J. SORENSEN - *Lyapunov-Based Integrator Resetting with Application to Marine Thruster Control*. IEEE Transactions on Control Systems Technology, 16(5):908-917, September 2008.

[55] E. F. CAMACHO, D. R. RAMIREZ, D. LIMON, D. MUOZ DE LA PEA, and T. ALAMO - *Model Predictive Control Techniques for Hybrid Systems*. Annual Reviews in Control, 34(1):21-31, 2010.

[56] A. BEMPORAD and S. DI CAIRANO - *Model-Predictive Control of Discrete Hybrid Stochastic Automata*. IEEE Transactions on Automatic Control, 56(6):1307-1321, 2011.

[57] A. BEMPORAD and M. MORARI - *Control of Systems Integrating Logic, Dynamics, and Constraints*. Automatica, 35(3):407-427, 1999.

[58] A. BEMPORAD, F. BORRELLI, and M. MORARI - *Piecewise Linear Optimal Controllers for Hybrid Systems.* American Control Conference, 2000, Proceedings of the 2000, volume 2, p. 1190-1194 vol. 2, 2000.

[59] C. SAMSON, B. ESPIAU, and M. LE BORGNE - *Robot Control: the Task Function Approach*. 1991.

[60] B. ESPIAU, F. CHAUMETTE, and P. RIVES - *A New Approach to Visual Servoing in Robotics*. IEEE Transactions on Robotics and Automation, 8(3):313-326, Jun 1992.

[61] H. DE PLINVAL, P. MORIN, P. MOUYON, and T. HAMEL - *Visual Servoing for Underactuated VTOL UAVS: a Linear, Homography-Based Approach*. Robotics and Automation (ICRA), 2011 IEEE International Conference on, p. 3004-3010, May 2011.

[62] H. DE PLINVAL, P. MORIN, P. MOUYON, and T. HAMEL - *Visual Servoing for Underactuated VTOL UAVs: a Linear, Homography-Based Framework*. International Journal of Robust and Nonlinear Control, Wiley - 24 (16), 2014.

[63] H. DE PLINVAL, P. MORIN, and P. MOUYON - *Nonlinear Control of Underactuated Vehicles with Uncertain Position Measurements and Application to Visual Servoing*. 2012 American Control Conference (ACC), p. 3253-3259, June 2012.

[64] P. MORIN, H. DE PLINVAL, and P. MOUYON - *Stabilization of a Class of Underactuated Vehicles with Uncertain Position Measurements and Application to Visual Servoing*. Automatica, p. 155-169, 2017.

[65] L. BURLION and H. DE PLINVAL - *Keeping a Ground Point in the Camera Field of View of a Landing UAV*. Robotics and Automation (ICRA), 2013 IEEE International Conference on, p. 5763-5768, May 2013.

[66] H. DE PLINVAL and L. BURLION - *Nonlinear Visual Servoing Control for VTOL UAVs with Field of View Constraint.* Advances in Aerospace Guidance Navigation and Control - Springer, p. 531-548, 2015.

[67] A. CHERUBINI and F. CHAUMETTE - *A Redundancy-Based Approach for Obstacle Avoidance in Mobile Robot Navigation*. Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on, p. 5700-5705, Oct 2010.

[68] B. SICILIANO and J. J. E. SLOTINE - *A General Framework for Managing Multiple Tasks in Highly Redundant Robotic Systems*. Advanced Robotics, 1991. "Robots in Unstructured Environments", 91 ICAR., Fifth International Conference on, p. 1211-1216 vol.2, June 1991.

[69] S. RUBRECHT - *Contributions to the Control of Constrained Robots*. PhD Thesis Department of Engineering, University Paris 6, UPMC, 2011.

[70] L. SENTIS and O. KHATIB - *Synthesis of Whole-Body Behaviors through Hierarchical Control of Behavioral Primitives*. International Journal of Humanoid Robotics, 02(04):505-518, 2005.

[71] S. RUBRECHT, V. PADOIS, P. BIDAUD, M. DE BROISSIA, and M. DA SILVA SIMOES - *Motion Safety and Constraints Compatibility for Multibody Robots*. Autonomous Robots, 32(3):333-349, 2012.

[72] S. RUBRECHT, V. PADOIS, P. BIDAUD, and M. DE BROISSIA - *Constraints Compliant Control: Constraints Compatibility and the Displaced Configuration Approach*. Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on, p. 677-684, Oct 2010.

[73] O. KERMORGANT and F. CHAUMETTE - *Multi-Sensor Data Fusion in Sensor-Based Control: Application to Multi-Camera Visual Servoing*. Robotics and Automation (ICRA), 2011 IEEE International Conference on, p. 4518-4523, May 2011.

[74] J. J. CRAIG and M. H. RAIBERT - *A Systematic Method of Hybrid Position/Force Control of a Manipulator.* Computer Software and Applications Conference, 1979. Proceedings. COMPSAC 79. The IEEE Computer Society's Third International, p. 446-451, 1979.

[75] T. YOSHIKAWA and A. SUDOU - *Dynamic Hybrid Position/Force Control of Robot Manipulators: On-Line Estimation of Unknown Constraint*. Springer Berlin Heidelberg, Berlin, Heidelberg, p. 116-134, 1990.

[76] J. SALINI, V. PADOIS, and P. BIDAUD - *Synthesis of Complex Humanoid Whole-Body Behavior: a Focus on Sequencing and Tasks Transitions*. Robotics and Automation (ICRA), 2011 IEEE International Conference on, p.1283-1290, May 2011.

[77] A. ESCANDE, N. MANSARD, and P. B. WIEBER - *A Framework for the Event-Triggered Stabilization of Nonlinear Systems*. International Journal of Robotics Research, 33(7):1006-1028, 2015.

[78] B. FAVERJON and P. TOURNASSOUD - *A Local Based Approach for Path Planning of Manipulators with a High Number of Degrees of Freedom*. Robotics and Automation. Proceedings. 1987 IEEE International Conference on, volume 4, p. 1152-1159, Mar 1987.

[79] O. KANOUN, F. LAMIRAUX, P. B. WIEBER, F. KANEHIRO, E. YOSHIDA, and J. P. LAUMOND. *Prioritizing Linear Equality and Inequality Systems: Application to Local Motion Planning for Redundant Robots.* Robotics and Automation, 2009. ICRA'09. IEEE International Conference on, p. 2939-2944, May 2009.

[80] D. A. C. NARCISO, V. DUA, and E. N. PISTIKOPOULOS - *Multiparametric Mixed-Integer Quadratic and Nonlinear Programming*. Wiley-VCH Verlag GmbH Co. KGaA, p. 73-97, 2007.

[81] R. LOBER, V. PADOIS, and O. SIGAUD - *Multiple Task Optimization Using Dynamical Movement Primitives for Whole-Body Reactive Control*. 2014 IEEE-RAS International Conference on Humanoid Robots, p. 193198, Nov 2014.

[82] R. LOBER, V. PADOIS, and O. SIGAUD - *Variance Modulated Task Prioritization in Whole-Body Control*. Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on, p. 3944-3949, Sept 2015.

[83] M. MORARI and J. H. LEE - *Model Predictive Control: Past, Present and Future*. Computers Chemical Engineering, 23(45):667-682, 1999.

[84] E. C. KERRIGAN, A. BEMPORAD, D. MIGNONE, M. MORARI, and J. M. MACIEJOWSKI - *Multi-Objective Prioritisation and Reconfiguration for the Control of Constrained Hybrid Systems*. American Control Conference, 2000. Proceedings of the 2000, vol.3, p. 1694-1698, 2000.

[85] M. BRAND, V. SHILPIEKANDULA, C. YAO, and S. A. BORTOFF - *A Parallel Quadratic Programming Algorithm for Model Predictive Control*. World Congress of the International Federation of Automatic Control (IFAC), June 2011.

[86] M. SANFOURCHE, J. DELAUNE, G. LE BESNERAIS, H. DE PLINVAL, J. ISRAEL, P.-H. CORNIC, A. TREIL, Y. WATANABE, and A. PLYER - *Perception for UAV: Vision-Based Navigation and Environment Modeling*. AerospaceLab Journal, 4, 2012.

[87] S. BERTRAND, J. MARZAT, H. PIET-LAHANIER, A. KHAN, and Y. ROCHEFORD - *Model Predictive Control Strategies for Cooperative Guidance of Autonomous Vehicles*. Aerospace Lab Journal, 8, 2014.

[88] Y. ROCHEFORT, H. PIET-LAHANIER, S. BERTRAND, D. BEAUVOIS, and D. DUMUR - *Model Predictive Control of Cooperative Vehicles Using Systematic Search Approach*. Control Engineering Practice, 32:204-217, 2014.

[89] J. MARZAT, H. PIET LAHANIER, and A. KAHN - *Cooperative Guidance of Lego Mindstorms NXT Mobile Robots*. 11th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2014), pages 667 - 682, 2014.

[90] H. ROGGEMAN, J. MARZAT, M. SANFOURCHE, and A. PLYER - *Embedded Vision-Based Localization and Model Predictive Control for Autonomous Exploration*. IROS Workshop on Visual Control of Mobile Robots (ViCoMoR 2014), Chicago, 2014.

[91] P. D. CHRISTOFIDES, R. SCATTOLINI, D. MUOZ DE LA PEA, and J. LIU - *Distributed Model Predictive Control: a Tutorial Review and Future Research Directions*. Computers Chemical Engineering, 51:21-41, 2013. {CPC} {VIII}.

[92] W. REN and R. BEARD - *Distributed Consensus in Multi-Vehicle Cooperative Control : Theory and Applications*. Springer-Verlag London, 2008.

[93] G. LAFFERRIERE, A. WILLIAMS, J. CAUGHMAN, and J. J. P. VEERMAN - *Decentralized Control of VehicleFormations*. Systems Control Letters, 54(9):899-910, 2005.

[94] C. C. CHEAH, S. P. HOU, and J. J. E. SLOTINE - *Region-Based Shape Control for a Swarm of Robots*. Automatica, 45(10):2406-2411, 2009.

[95] C. FORSTER, S. LYNEN, L. KNEIP, and D. SCARAMUZZA - *Collaborative Monocular Slam with Multiple Micro Aerial Vehicles*. 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, p. 3962-3970, Nov 2013.

[96] G. BRESSON, R. AUFRRE, and R. CHAPUIS - *Consistent Multi-Robot Decentralized Slam with Unknown Initial Positions*. Information Fusion (FUSION), 2013 16th International Conference on, p. 372-379, July 2013.

[97] I. V. MELNYK, J. A. HESCH, and S. I. ROUMELIOTIS - *Cooperative Vision-Aided Inertial Navigation Using Overlapping Views*. Robotics and Automation (ICRA), 2012 IEEE International Conference on, p. 936-943, May 2012.

[98] M. W. ACHTELIK, S. WEISS, M. CHLI, F. DELLAERTY, and R. SIEGWART - *Collaborative stereo.* 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, p. 2242-2248, Sept 2011.

[99] N. PIASCO, J. MARZAT, and M. SANFOURCHE. *Collaborative Localization and Formation Flying Using Distributed Stereo-Vision*. 2016 IEEE International Conference on Robotics and Automation (ICRA), p. 1202-1207, May 2016.

[100] D. ZOU and P. TAN. *Coslam: Collaborative Visual Slam in Dynamic Environments*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(2):354-366, Feb 2013.

[101] A. N. VENKAT - *Distributed Model Predictive Control: Theory and Applications*. PhD Thesis University of Wisconsin Madison, June 2006.

[102] A. IBANEZ - *Emergence of Complex Behaviors from Coordinated Predictive Control in Humanoid Robotics*. PhD Thesis University Pierre and Marie Curie, Paris, France, 2015.

[103] M. D. DOAN, T. KEVICZKY, and B. DE SCHUTTER - *A Dual Decomposition-Based Optimization Method with Guaranteed Primal Feasibility for Hierarchical MPC Problems*. 18th IFAC World Congress. IFAC Proceedings Volumes, 44(1):392-397, 2011.

[104] M. BLANKE, M. STAROSWIECKI, and N. E. WU - *Concepts and Methods in Fault-Tolerant Control*. American Control Conference, 2001. Proceedings of the 2001, volume 4, p. 2606-2620, 2001.

[105] D. CHILIN, J. LIU, J. F. DAVIS, and P. D. CHRISTOFIDES - *Data-Based Monitoring and Reconfiguration of a Distributed Model Predictive Control System*. Proceedings of the 2011 American Control Conference, p. 3158-3165, June 2011.

[106] D. AXEHILL - *Applications of Integer Quadratic Programming in Control and Communication*. PhD thesis, Department of Electrical Engineering Linkpings universitet, Sweden, 2005.

[107] M. GHALLAB, D. NAU, and P. TRAVERSO - *Automated Planning: Theory and Practice*. Morgan Kaufmann, 2004.

[108] C. PRALET and G. VERFAILLIE - *Dynamic Online Planning and Scheduling Using a Static Invariant-Based Evaluation Model*. Proceedings of the International Conference on Automated Planning Scheduling (ICAPS- 13), 2013.

[109] P. BECHON, M. BARBIER, G. INFANTES, C. LESIRE, and V. VIDAL - *HiPOP: Hierarchical Partial-Order Planning).* Proc. of STAIRS-14, 2014.

[110] V. VIDAL and H. GEFFNER. *Branching and Pruning: an Optimal Temporal POCL Planner Based On Constraint Programming*. Artificial Intelligence, 170(3):298-335, 2006.

[111] S. CHIEN, R. KNIGHT, A. STECHERT, R. SHERWOOD, and G. RABIDEAU - *Using Iterative Repair to Improve the Responsiveness of Planning and Scheduling*. 2000.

[112] T. SCHOUWENAARS - *Safe Trajectory Planning of Autonomous Vehicles*. PhD Thesis Department of Aeronautics and Astronautics, Katholieke Universiteit Leuven, 2001.

[113] A. IBANEZ, P. BIDAUD, and V. PADOIS - *Emergence of Humanoid Walking Behaviors from Mixed-Integer Model Predictive Control*. 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, p. 40144021, Sept 2014.

[114] B. HOMMEL, J. MSSELER, G. ASCHERSLEBEN, and W. PRINZ. *The Theory of Event Coding (Tec): a Framework for Perception and Action Planning*. Behavioral and Brain Sciences, 24(5):849-878.

[115] A. G. RICHARDS - *Robust Constrained Model Predictive Control*. PhD Thesis Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, 2005.

[116] P. BECHON, M. BARBIER, C. LESIRE, G. INFANTES, and V. VIDAL - *Using Hybrid Planning for Plan Reparation*. Proc. of ECMR-15 (European Conference on Mobile Robots), 2015.

[117] C. PRALET, G. INFANTES, and G. VERFAILLIE - *Application Showcase: a Generic Constraint-Based Local Search Library for the Management of an Electromagnetic Surveillance Space Mission.* Proc. of the ICAPS-13 Application showcase, 2013.

[118] C. PONZONI CARVALHO CHANEL, C. LESIRE, and F. TEICHTEIL-KONIGSBUCH - *A Robotic Execution Framework for Online Probabilistic (Re)Planning*. Proc. of the 24th International Conference on Automated Planning and Scheduling (ICAPS-14), 2014.

[119] C. PRALET and C. LESIRE - *Deployment of Mobile Wireless Sensor Networks for Crisis Management: a Constraint-Based Local Search Approach.* 20th International Conference on Principles and Practice of Constraint Programming (CP-14), 2014.

[120] M. PINEDO - *Scheduling: Theory, Algorithms, and Systems*. Springer, 2012.

[121] F. ROSSI - *Handbook of Constraint Programming.* Elsevier, 2006.

[122] M. PUTERMAN - *Markov Decision Processes, Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 1994.

[123] P. MORRIS, N. MUSCETTOLA, and T. VIDAL - *Dynamic Control of Plans with Temporal Uncertainty*. 2001.

[124] K. EROL, J. HENDLER, and D. S. NAU. *HTN Planning: Complexity and Expressivity*. Proc. of the 12th National Conference on Artificial Intelligence (AAAI-94), 1994.

[125] M. FOX, A. GEREVINI, D. LONG, and I. SERINA - *Plan Stability: Replanning Versus Plan Repair*. Proceedings of The International Conference on Automated Planning Scheduling (ICAPS-06), 2006.

[126] S. LEMAI and F. INGRAND - Interleaving temporal planning and execution in robotics domains. AAAI'04 Proceedings of the 19th national conference on Artifical intelligence (AAAI04P), pages 617-622, 2004.

[127] G. CASANOVA, C. PRALET, C. LESIRE, and T. VIDAL - *Solving Dynamic Controllability Problem of Multi-Agent Plans with Uncertainty Using Mixed Integer Linear Programming.* Proc. of the 22nd European Conference on Artificial Intelligence (ECAI-16), 2016.

[128] A. MAILLARD, C. PRALET, J. JAUBERT, I. SEBBAG, F. FONTANARI, and J. L'HERMITTE - *Ground and Bboard Decision-Making on Data Downloads*. 25th International Conference on Automated Planning and Scheduling (ICAPS-15), 2015.

[129] P. VAN HENTENRYCK and L. MICHEL - *Constraint-Based Local Search*. The MIT Press, 2005.

[130] P. VAN HENTENRYCK and R. BENT - *Online Stochastic Combinatorial Optimization*. MIT Press, 2006.

[131] A. KAHN, J. MARZAT, H. PIET-LAHANIER, and M. KIEFFER - *Cooperative Estimation and Fleet Reconfiguration for Multi-Agent Systems*. 3rd IFAC Workshop on Multi-Vehicle Systems, IFAC Proceedings Volumes, 48(5):11 - 16, 2015.

Acronyms

CML	(Concurrent Mapping and Localization)
CoSLAM	(Collaborative SLAM)
DMPC	(Distributed MPC)
FTC	(Fault-Tolerant Control)
GPS	(Global Positioning System)
GPU	(Graphics Processing Unit)
HQP	(Hierarchical Quadratic Programming)
HTN	(Hierarchical Task Networks)
INS	(Inertial Navigation System)
IQC	(Integral Quadratic Constraints)
LFT	(Linear Fractional Transformation)
LMI	(Linear Matrix Inequality)
LTI	(Linear Time Invariant)
LTV	(Linear Time Varying)
MILP	(Mixed Integer Linear Programming)
MIQP	(Mixed Integer Quadratic Programming)
MLD	(Mixed Logical Dynamical)
MPC	(Model Predictive Control)
mp-MILP	(multi-parametric Mixed Integer Quadratic Programming)
MRAW	(Model Recovery Anti Windup)
OIST	(Output to Input Saturation Transformation)
OISTeR	(Output to Input Saturation Transformation extended for Robustness)
PD	(Proportional-Derivative)
PWA	(Piecewise Affine)
PWL	(Piecewise Linear)
QP	(Quadratic Programming)
SIMD	(Single Instruction, Multiple Data)
SLAM	(Simultaneous Localization And Mapping)
TEC	(Theory of Event Coding)
UAV	(Unmanned Aerial Vehicle)

AUTHORS

Philippe Bidaud received his PhD in Robotics in 1984 and he was recruited as a CNRS researcher in 1985. He began his career at the Laboratory of Mechanics and Robotics at the Pierre and Marie Curie University in Paris. He subsequently directed several research groups at Robotics Laboratory of Paris. Meanwhile. In 1998, he was hired as a professor at the University Pierre and Marie Curie. He became deputy director of the Robotics Laboratory of Paris in 1999 and then director in 2002 before creating the Institute of Intelligent Systems and Robotics which he provided leadership until 2013. Today Philippe Bidaud is Scientific Director for Information Technologies and Sciences at ONERA. For several years now, his research activities focuses design and control complex robotics systems specifically all-terrain robot, dextrous manipulation systems, humanoid robotics, interactive robotics systems and more recently space robotics. All of his work is the subject of more than 230 publications in journals and international conferences and a dozen of patents.

Laurent Burlion graduated from the French engineering school Ensieta (now Ensta Bretagne) in 2003 and received the same year the M.Sc. degree in Control at the IRRCyN in Nantes. Dr. Burlion received his Ph.D degree from the University of Paris Sud in 2007. He was a control Engineer with the Laboratoire de Recherches Balistiques et Aerodynamiques from 2003 to 2010. Since 2010, he has been working as a research scientist at the French Aerospace Lab (ONERA). His current research interests are focused on the problems of input saturations and output constraints with application to UAVs and satellites.

Henry de Plinval graduated from Ecole Polytechnique, France, in 2006 and received his M.Sc. in Aeronautics and Astronautics in 2006 from MIT (USA) and his Ph.D in automatic control from the University of Toulouse in 2014. Since 2008, he is working with ONERA. His areas of interest include guidance, navigation and control, mostly for UAVs,with a particular focus on visual servoing for VTOL UAVs. Since 2015, he is the UAVs Program Director at ONERA, covering the related activities in the areas of automatic control, artificial intelligence, fluid mechanics, optronics, radar, telecommunications, ...

Thomas Loquen is researcher with ONERA, Information Processing and Systems Department (DTIS). In 2006, he completed his Engineering degree and M.Sc in Electrical engineering and Automatics from INPT-ENSEEIHT. In 2010, he completed his Ph.D. degree in Control Systems within the INP-Toulouse, by working on stability analysis of hybrid systems. His main interests concern hybrid dynamical systems (modelling, robust analysis) and embedded systems (joint works with computer science researchers of ONERA).

Julien Marzat graduated as an engineer from ENSEM (INPL Nancy) in 2008 and completed his PhD Thesis from University Paris Sud XI in 2011. Since then, he has been working as a Researcher at ONERA. His research interests include guidance, control, estimation and fault diagnosis with application to autonomous vehicles and aerial robotics.

Cédric Pralet graduated from SUPAERO (French engineering school in aeronautics and space) in 2003. He passed his PhD in Computer Science in Toulouse in 2006. He is now working as a research engineer at ONERA. His research interests concern constraint-based combinatorial optimization and automated planning and scheduling. His research activities are applied to space and aeronautics.

J.-M. Biannic, G. Hardier,
C. Roos, C. Seren,
(ONERA)

L. Verdier
(AIRBUS Operations SAS, EGYA)

E-mail: Georges.Hardier@onera.fr

DOI: 10.12762/2016.AL12-14

Surrogate Models for Aircraft Flight Control: Some Off-Line and Embedded Applications

The use of surrogate models is now very widespread in engineering activities to replace high-fidelity models, computation codes or simulators that are too complex or time consuming. This is especially the case in the aerospace field: although physical models are available, derived from aerodynamics, structural dynamics or flight mechanics, quite often they cannot be used just as they are. Hence, simplified representations need to be developed in order to achieve some tasks involving for instance optimization, parameter identification, embedded implementations, and so on. This paper illustrates some of these aspects in the field of aircraft flight control systems. After outlining the characteristics of the preferred surrogate models, the techniques that have been developed for constructing these models efficiently are presented. Then, three relevant off-line and on-line applications are described, with the surrogate models being respectively used: to create parsimonious representations useful for building Linear Fractional Representations for analysis and design of control laws, as intermediate models in the modeling of rigid aircraft to facilitate the identification process of aerodynamic nonlinearities from flight tests, and to obtain embeddable models used for the virtual sensing of some flight parameters required to schedule the control/protection laws.

Introduction

A number of activities in aeronautical engineering rely on the availability of models to represent the real behavior of the aircraft. For example, let us quote the analysis and design of flight control laws, the study of the handling qualities, the fault monitoring process, the prediction of hazardous behaviors, or the implementation of simulators used to train the pilots and to validate hardware and software systems. One feature of the aeronautical field is that many physical models are available, derived from aerodynamics, structural dynamics or flight mechanics. Accordingly, the development of Flight Control Systems (FCS) often makes use of model-based techniques (Figure 1), but these models cannot always be implemented just as they are because of high orders, strongly nonlinear behaviors or, more generally, because they have become increasingly complex in recent years. On the other hand, surrogate models are very useful to replace

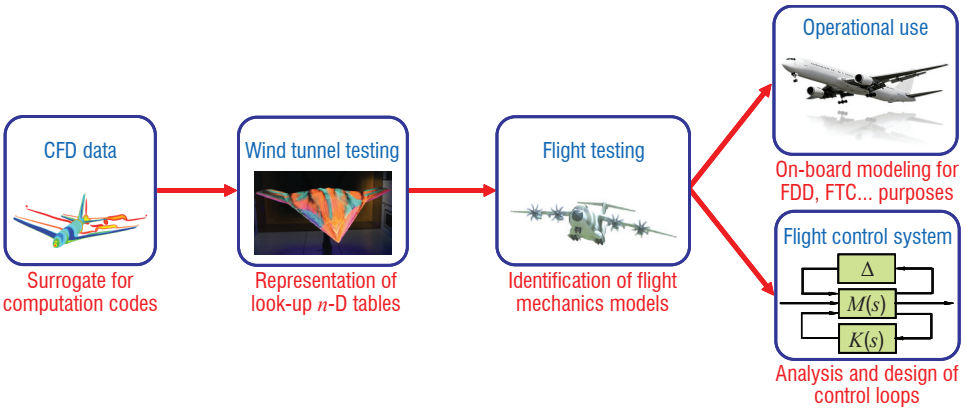


Figure 1 – Some development stages likely to benefit from surrogate models

the system or the reference model when the latter is too complex or time consuming for achieving some tasks (optimization, parameter identification, embedded implementation, etc.).

Consequently, a wide range of models and associated methods has been developed for building surrogate models efficiently, both accurately and parsimoniously [1][7][8][34][54]. Within this range, Neural Networks (NN) are recognized nowadays as an efficient alternative for representing complex nonlinear systems, including dynamical ones [14][21][49]. As a result, one can also benefit from efficient tools permitting the static nonlinearities to be modeled, such as those involved in the representation of aerodynamic coefficients resulting from CFD computations, or from wind tunnel or flight testing [26][40][47][52][55]. Usually, these data are only available in the form of look-up tables and hence are not very convenient for on-board implementation [4][22]; that is why analytical and differentiable approximations are contemplated, with lower memory requirements also. For instance, special types of NN can be advantageously used to design grey-box neural models arranged to accurately represent the aircraft aerodynamic coefficients appearing in the flight nonlinear equations of motion [5]. Such NN permit the physical readiness and the structure of aerodynamics to be preserved in the final surrogate model (as opposed to black-box approaches often promoted in the field), as illustrated below for the pitching moment coefficient Cm in clean configuration:

$$Cm = SlPd\left[\underbrace{Cm_0^{NN}}_{f(M)} + \left(x_{CG} - \underbrace{x_{AC}^{NN}}_{f(M)}\right)\underbrace{Cz_\alpha^{NN}}_{f(M)}\alpha + \underbrace{\kappa_{NL}^{NN}}_{f(Pd,M)}\underbrace{\Delta Cm_{NL}^{NN}}_{f(\alpha,M)} + \ldots\right] \quad (1)$$

In (1), the parameters α, M, S, Pd and l refer respectively to the aerodynamic Angle of Attack (AoA), Mach number, reference area, dynamic pressure and mean aerodynamic chord. x_{CG} and x_{AC} correspond to the longitudinal abscissa of the Center of Gravity and Aerodynamic Center. $\left[Cm_0\ x_{AC}\ Cz_\alpha\ \Delta Cm_{NL}\right]$ and $\left[\kappa_{NL}\right]$ are respectively rigid-body aerodynamic and static aeroelastic neural approximations of nonlinear effects (as indicated by the NN exponent) which contribute to the global model of Cm. From this simple example, it can be stated that if such surrogate models can be created from the initial data, several model-based techniques will be facilitated, both for off-line and on-line applications.

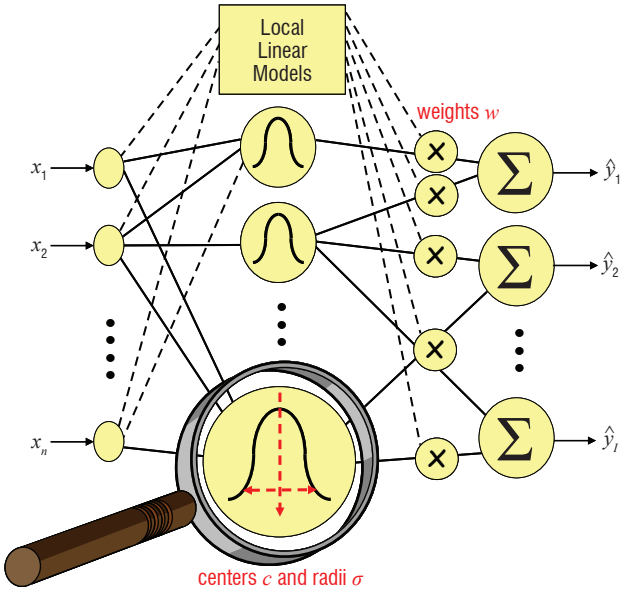


Figure 2 – RBF/LLM-type surrogate models

To cover these aspects, the paper is organized as follows. The first section outlines the characteristics of the surrogate models preferred for meeting the previous needs, and gives an overview of the techniques that have been developed for constructing these models efficiently. The three following sections are devoted to a short description of some relevant applications in the field of flight control systems:

- to derive simplified parsimonious representations useful for building Linear Fractional Representations (LFR) of lower complexity for analysis and design of control laws,
- as intermediate models to facilitate the identification process of aerodynamic nonlinearities from flight tests,
- to obtain embeddable models used for the virtual sensing of some flight parameters required to schedule the control/protection laws (tolerance to sensor faults). Concluding remarks and perspectives end the paper.

Surrogate models for aircraft modeling

The equations of motion of a rigid body A/C are derived from the fundamental principles of mechanics. Thus, the external forces involved in these equations arise from propulsion, aerodynamics and gravitational attraction. The major source of complexity in this model mostly comes from the highly nonlinear aerodynamic effects, and to a lesser extent from the propulsion components. Consequently, the modeling of the aerodynamic part within the whole flight envelope is the main challenge of many model-based applications. The following subsections outline the characteristics of the nonlinear surrogate models that were selected, and which play a pivotal role in the practical applications described afterwards, and focus on different aspects of flight control systems.

Nonlinear regression for Linear-in-their-Parameters models

At first, it is noteworthy that a nonlinear model can be either linear or nonlinear with regard to its internal parameters. Within the framework of NN, the latter case corresponds for example to Multi-Layered Perceptrons (MLP) [8][14] as well as to Radial Basis Function (RBF) networks [7][17][37][53] when the nonlinear parameters of the radial units (centers c and radii σ in Figure 2) are also optimized in addition to the linear ones (weights w in Figure 2). Clearly, this is the most general formulation since Linear-in-their-Parameters (LP) models are nothing but a special case, and it is the origin of the NN theoretical properties as parsimonious approximators. However, the joint optimization of the whole set of model parameters (linear plus nonlinear) practically results in ill-posed problems, which are likely to converge only by enforcing strong regularization constraints in the optimization process. This issue can be overcome by splitting the parameters into linear and nonlinear subsets [16], but LP models are always quite common practice anyway, because more simple and robust algorithms can be adopted, derived from the classical methods in use for estimating the parameters of linear regressions. Some advantages for preferring RBF versus MLP are:

- a possible integration of physical constraints,
- they make it easier to perform a local adjustment from heterogeneous data distribution or from sparse partial data relative only to portions of the flight domain,
- their grey-box architecture is better suited to integrate some initial knowledge (keeping the aerodynamic model readable) and to manage the optimization process (w.r.t. pure black-box networks),

- they develop local approximations instead of global ones making constructive algorithms possible (for determining the number of units),
- the various parameters (linear and nonlinear) can be sequentially optimized during the learning process.

On the other hand, most of the existing tools related to those local models are based on some initial user-defined architecture. For RBF, this amounts to setting *a priori* the number of kernels as well as their distribution in the input space. Although this distribution can be more or less automatically defined from the data repartition, and although some pruning is possible after learning, this is not a real structural optimization. Most often, only input data are used for choosing the kernel distribution without considering the output values, and hence the match between the model and the problem complexity cannot be guaranteed. Actually, by taking advantage of the features of LP models, both parametric and structural identification can be jointly proceeded to for extracting a suitable set of regressors from the available data. The procedure should determine the whole set of parameters involved in the nonlinear regression, and neither the number of kernels nor their distribution would have to be known in advance. The algorithm described in the sequel of this section makes use of the best techniques in the fields of Orthogonal Least-Squares (OLS) [9], local regularization [39], Separable Nonlinear Least-Squares (SNLS) [16], as well as a global evolutionary optimization that hybridizes local and global search [12].

The approach relies on a regularized constructive process that adds new regressors progressively, in terms of both a performance cost and the generalization errors. The LP models that we consider hereafter comply with the generic form:

$$\hat{y}_k = f(x_k) = \sum_{j=1}^{m} w_j r_j(x_k) \qquad (2)$$

where x_k is a vector gathering the n explanatory variables (α, M, Pd in the example of Eq. (1)), and y_k is assumed to be scalar to alleviate the developments. The functions r_j are the nonlinear regressors to be defined, while w_j are the regression parameters also to be determined during the optimization process, as well as the number m of regressors (unknown *a priori*). At this time, the regressors r_j can represent either the monomials of a polynomial expansion, the radial functions of RBF or Local Linear Models (LLM), etc. [34], but we will focus in the sequel on the use of local models (Figure 2). From N available data samples (x_k, y_k) representing the tabulated coefficient to be modeled, the LS minimization of the approximation errors yields to consider the criterion:

$$C = \sum_{k=1}^{N} \left(y_k - f(x_k)\right)^2 \quad \left[+\sum_{j=1}^{m} \lambda_j w_j^2\right] \qquad (3)$$

where adding an [optional] regularization term enables large values of the w_j parameters to be penalized, and thus the conditioning of the problem to be improved in case of model overfitting (ridge regression). By grouping the regularization parameters into the diagonal matrix $\Lambda = diag\{\lambda_j, j = 1,\ldots,m\}$ (also to be adapted), the well-known solution to this LS problem is given by:

$$\hat{w} = (R^T R + \Lambda)^{-1} R^T y = H^{-1} R^T y \qquad (4)$$

by denoting R_j as the j^{th} column of the regression matrix R, written as $R_j^T = \left[r_j(x_1) \quad r_j(x_2) \quad \cdots \quad r_j(x_N)\right]$, and H^{-1} as the covariance matrix of the estimation errors. It is noteworthy that, if $\Lambda = \lambda I_m$, this results in global regularization, whereas local ridge regression is expected when all of the λ_j are different from each other, every parameter w_j being weighted with its own penalty. When radial functions are used as regressors for instance, this form of regularization permits the resulting smoothness to be controlled in a local fashion [37]. If the individual λ_j are optimized, this smoothness can then be adapted in the different regions of the input space where the regressors are located. From (4), the output prediction error can be simply expressed as:

$$\varepsilon = y - \hat{y} = y - R\hat{w} = \left(I_N - RH^{-1}R^T\right) y = P y \qquad (5)$$

A matrix P appears in (5), which is the projection matrix in the simplest case (no penalty). It projects the vector y defined in a N-dimensional space in the m-dimensional subspace spanned by the model. This matrix plays a crucial role in the regression properties, and is involved in most of the relationships, *e.g.* the cost function $C = y^T P y$. To choose the regressors r_j, we will focus on methods based on forward selection, as opposed to another class of methods that consists in first selecting a full set of candidates, and then removing the less relevant ones one by one (backward elimination). Forward selection starts with an empty subset, and the regressors are added one at a time in order to gradually improve the results. Therefore, the final number of regressors is not known in advance, and the computational cost is reduced since the regression size will become large only if it is required to reduce the modeling error. We will examine in the sequel how the candidates can be chosen or generated from scratch. Let us first consider the question of how and when to stop this forward selection process. The available data being limited in number and sometimes noisy, the model that we seek should correspond to the best compromise between the performances evaluated on the training samples and the extrapolated performances evaluated when the model is used with other inputs absent from the initial data base. This is referred to as validation or generalization error. With a few not very restrictive assumptions, it can also be proved that the expected cost can be expanded as $E(C) \approx \mathrm{var}(\xi) + \mathrm{biais}(m) + \mathrm{var}(\xi)\tilde{m}/N$ (over all possible data sets), where $\mathrm{var}(\xi)$ is the variance of the noise ξ, and the effective number of parameters $\tilde{m}$ (usually less than m) depends on the regularization parameters.

To tackle this well-known bias/variance dilemma, the most common technique consists in splitting the available data set into two groups, the 1st one being used for estimating the parameters w and the 2nd one for the validation stage. Although irrelevant in practice, this trick leads to a more interesting concept, the virtual "Leave-One-Out" method (LOO) and its variants [29]. This technique generalizes the previous split by suggesting the virtual use of N subsets to be made, each of them including only N–1 samples, prior to computing the model performance by averaging the N estimations. The idea is to achieve a theoretical prediction (*i.e.*, without really performing the corresponding optimizations) of the generalization errors resulting from the withdrawal of every sample included in the data base. All of the available data are thus effectively used for the optimization process, but the validation stage takes also into account the consequences of discarding any of them. Moreover, the interest of this technique is reinforced in the case of LP models, since a very simple analytical expression of the generalization error can then be derived. By denoting as $f_k(x_k)$ the predicted output obtained for the k^{th} sample when the model is optimized from the N–1 remaining

samples, the LOO variance of the generalization error (also known as PRESS for Predicted REsidual Sum of Squares) can be effectively written as [37]:

$$\sigma_{LOO}^2 = \frac{1}{N}\sum_{k=1}^{N}\left(y_k - f_k\left(x_k\right)\right)^2 = \frac{1}{N}\sum_{k=1}^{N}\left[\varepsilon^{N-1}(k)\right]^2 = \frac{1}{N}\, y^T P diag(P)^{-2} P y \quad (6)$$

Forward selection methods

Orthogonal Least Squares

Thanks to this estimation of the generalization error, it is possible to rule on the benefits from adding a new regressor to the model, at every step of the forward selection process. On the other part, when using LP models, it is straightforward to determine by advance the result of elementary operations, like adding or subtracting any regressor. Again, quite simple analytical formulae can be established to cover these situations [37], which avoid retraining the model from scratch. At every step of a constructive approach, the chosen new regressor will thus be the one reducing the most the value of the cost function. The process will be pursued that way until: either a user-defined minimum value of the criterion C_m is reached (knowing that this criterion will continue to decrease as long as m increases), or the generalization error (*e.g.* the LOO one estimated by σ_{LOO}^2) stops decreasing which reveals model overfitting. Otherwise, determining the pool of regressors within which the selection will operate is of course a key point for these approaches. There are several alternatives, which can modify the final results and the model parsimony. This aspect is discussed afterwards.

Forward selection is computationally efficient, but constructive algorithms can be sped up even further thanks to a preliminary orthogonalization process, making use of the famous Gram-Schmidt technique. Moreover, this procedure permits the successive regressors to be decoupled from each other, and hence their individual contribution to be evaluated regardless of those already recruited for the modeling. The principle of the method is based on factoring the regression matrix as $R_m = \tilde{R}_m U_m$, where U_m is upper triangular, and where the columns of the orthogonalized regression matrix $\tilde{R}_m = [\tilde{r}_1 \;\; \tilde{r}_2 \;\; \dots \;\; \tilde{r}_m]$ are such that $\tilde{r}_i^T \tilde{r}_j = 0$ for $i \neq j$. Consequently, when adding a new regressor r_{m+1}, corresponding to the $(m+1)^{th}$ column of the regression matrix R, only its projection perpendicular to the space already spanned by the m first regressors needs to be considered, and can contribute to a further reduction of the criterion [37]. This projection yields the recurrence:

$$\tilde{r}_{m+1} = r_{m+1} - \sum_{j=1}^{m} \frac{\tilde{r}_j^T r_{m+1}}{\tilde{r}_j^T \tilde{r}_j} \tilde{r}_j \quad (7)$$

From P, assuming no weight penalty ($\Lambda = 0$), an expression can be derived for the expected reduction of the cost, depending only on the output vector y and the orthogonal regressor $\tilde{r}_{m+1}$ (decoupling). The regression parameters (orthogonal $\tilde{w}_m$ and ordinary ones $\hat{w}_m = U_m^{-1}\tilde{w}_m$) are then also easily derived [37]:

$$\begin{cases} P_m = I_N - \sum_{j=1}^{m} \frac{\tilde{r}_j \tilde{r}_j^T}{\tilde{r}_j^T \tilde{r}_j} \Rightarrow \; C_m - C_{m+1} = y^T (P_m - P_{m+1}) y = \frac{(\tilde{r}_{m+1}^T y)^2}{\tilde{r}_{m+1}^T \tilde{r}_{m+1}} \\ \tilde{w}_m^T = [\tilde{r}_1^T y \;\; \cdots \;\; \tilde{r}_m^T y] \; diag\{(\tilde{r}_j^T \tilde{r}_j)^{-1}, j = 1,\dots,m\} \end{cases} \quad (8)$$

Local regularization and Separable Nonlinear Least Squares

Fortunately, the advantages of forward selection and regularization techniques (whether local or global) can be combined without too much complexity in the resulting algorithms. Furthermore, when radial functions are used as regressors (*e.g.* in the case of RBF or LLM), an analytical solution can usually be found [37], allowing the optimal value of each λ_j (for a local regularization) to be determined. However, this value is related to those of the $(m-1)$ other parameters λ_j, and an iterative process is required *a priori*. Another approach is explained below, resulting in a joint or alternate procedure to optimize both regressors and regularization parameters. It can also help to reduce the modeling size by removing a few regressors that have become irrelevant. Effectively, if a parameter value λ_j converges towards $+\infty$, the j^{th} regressor can be eliminated without any damage to the performances. Thus, introducing a local regularization into the forward selection incidentally enables useless regressors to be pruned without having to implement backward elimination. By combining OLS and local regularization, the LOO/PRESS validation error becomes $\varepsilon^{N-1}(k) = \varepsilon_m^N(k) / \eta_m^N(k)$, where the modeling error $\varepsilon_m^N(k)$ appears weighted by a coefficient $\eta_m^N(k)$ which contributes to the computation of the variance of the validation error [9]. A major advantage of this expression is to facilitate the implementation of a recurrent formulation, resulting in a very effective algorithm [9]:

$$\begin{cases} \varepsilon_m^N(k) = y_k - \sum_{j=1}^{m} \tilde{r}_j(k)\tilde{w}_j = \varepsilon_{m-1}^N(k) - \tilde{r}_m(k)\tilde{w}_m \\ \eta_m^N(k) = \eta_{m-1}^N(k) - \frac{\tilde{r}_m^2(k)}{\tilde{r}_m^T \tilde{r}_m + \lambda_m} \end{cases} \quad (9)$$

initialized only from the data outputs (no regressor to start the process): $C_0 = (y^T y)/N$, $\varepsilon_0(k) = y_k$ and $\eta_0(k) = 1$ (for $1 \le k \le N$). Regarding the joint search for optimal values of the λ_j (in addition to those of the model parameters w), it amounts to minimizing the cost $C(w,\Lambda) = \varepsilon^T\varepsilon + w^T\Lambda w$ and can be solved within the framework of Bayesian learning [9]. Finally, this iterative process involves the evaluation of the $\tilde{w}_j$ parameters in terms of the current λ_j, followed by an adaptation of the λ_j and so on. On the other hand, despite the unquestionable pros of the previous forward selection, the choice of each regressor is optimal iff it is considered by itself, *i.e.* for a given pre-selection of the pool of its predecessors. Disregarding this matter results in sub-optimality, which is the cost to pay for avoiding having to deal with a nonlinear complex optimization, involving both the regression parameters and the parameters of the regressor kernels. Above all, it has to do with the major issue that the number of regressors is not known in advance, and has to be inferred at the same time. However, it is possible to win on both counts, and thus to improve the forward selection, by adding some extra optimization stages to the previous constructive algorithm. The idea is to periodically call the global positioning of the selected regressors into question, in order to minimize the criterion value for a given model size, before continuing with the addition of new terms.

To improve the course of these optimization stages (time and conditioning), it is wise to benefit from particular LS techniques, referred to as SNLS, and developed for LP models [16]. Denoting by u the vector gathering all of the internal parameters defining a regressor (*e.g.*, the centers and radii of a RBF node), the sum-squared error (3) is expressed without loss of generality as: $C(u,w) = \varepsilon^T\varepsilon$ with $\varepsilon = \hat{y} - y = R(u)w - y$. It is worth noting that a regularization

term is no longer needed since the number of regressors is fixed for the time being. From this expression, the trick of the SNLS methods comes from the fact that the optimal value of the parameters w results directly from the value of the parameters u. In other words, for a given u, w is available by solving an ordinary LS problem: $w(u) = \left[R^T(u)R(u)\right]^{-1} R^T(u)y = H^{-1}R^T y$. Therefore, it would be useless and even detrimental to process both sets of parameters at the same time. By using the previous expression of $w(u)$ every time the cost function needs to be evaluated, it appears that we can consider C as a function of u only. The optimization stages can thus be restricted to this vector, reducing the dimension of the parameter space. It is also noteworthy that they will converge nicely and quickly even if 2nd order standard algorithms (Gauss-Newton type) are used, which are known for their sensitivity to the initial conditions. There are two reasons for this: the conditioning is greatly improved by the SNLS formulation, and the nonlinear optimization starts under good conditions thanks to the initial parameters u available from the constructive algorithm. Finally, after every optimization stage, a new orthogonalization of the regressors obtained from the resulting u_{opt} is also necessary before starting the forward selection again.

Determining the set of regressors

To implement the previous forward selection, two options are available: ① to first define an initial pool of candidate regressors from which the most relevant ones will be selected, ② to determine each regressor individually as the process goes on, which generally amounts to optimizing the kernel functions in the input space. Within class ① is the whole range of classical and direct methods that locate the regressor kernels quite arbitrarily: in a subset of the data samples, on the knots of a lattice derived from a gridding of the input space, by using data clustering or self-organization techniques. These approaches have the advantage of being very simple to implement and of providing a pool of candidate regressors almost instantly. On the other hand, they do not provide any guarantee regarding the balance between the distribution of the regressors and the real complexity of the modeling problem, since they do not use all available information (only input data are used). More elaborate methods attempt to fill this gap by building regression trees that do use the values of data outputs [38]. They recursively partition the input space, approximating the output values by their average value computed from the samples included in each partition. The main weakness of these methods is to be very greedy, the bias/variance dilemma resulting in the following interrogation: when should the branching out of the tree stop, and how should it then be pruned?

Class ② is related to optimization techniques, but to avoid the problems inherent to classical methods (convergence, sensitivity to initial values) global optimization is preferred, among which evolutionary algorithms have done particularly well for some years. New techniques have thus appeared such as Genetic Algorithms (GA), boosting search, or simulated annealing [10][53]. In return for their conceptual simplicity, these classical evolutionary algorithms can be computationally very expensive by requiring a number of cost function evaluations, and they often require the tuning of several internal parameters. More recently in the 90's, a new metaheuristics also arising from biological inspiration (bird flocking or fish schooling) was imagined, known as Particle Swarm Optimization (PSO). The collective behavior of the particles looks like a swarm of living beings, and the most relevant metaphor certainly concerns the bees because an individual having discovered a good spot passes on the information to the others, and is used to direct their next moves. Therefore, the swarm represents a set of autonomous and interacting agents, cooperating to solve a problem. The members of a group benefit from the accidental discoveries as well as the experience acquired by other individuals. Similarly to the evolutionary case, the method is based on an iterative and stochastic process [12]. At iteration i, the position of the particles $\vec{u}_i$ (which includes all of the parameters to be optimized) and their velocity $\vec{v}_i$ are updated as follows:

$$\begin{cases} \vec{v}_{i+1} = c_1\vec{v}_i + c_2U(0,1)\otimes\left[\vec{p}_i - \vec{u}_i\right] + c_3U(0,1)\otimes\left[\vec{g}_i - \vec{u}_i\right] \\ \vec{u}_{i+1} = \vec{u}_i + \vec{v}_{i+1} \end{cases} \quad (10)$$

where $\vec{p}_i$ corresponds to the best position ever reached by particle i, and $\vec{g}_i$ to the best position ever reached by the pool of its informants. $U(0,1)$ is a random number between 0 and 1 chosen following a uniform distribution, and $\otimes$ is a symbolic operator for the element-by-element product of two vectors. Thus, it appears that a particle updates its velocity through a weighted linear combination of three behaviors [12]: an *adventurous* behavior for preserving the acquired velocity (inertial component of the motion), a *conservative* behavior consisting of getting closer to its best position (*cognitive* component of the motion), and a *sheeplike* behavior for getting closer to its best informant (*social* component). It is also worth noting that the required number of particles remains quite low as opposed to usual GA populations, and that PSO algorithms are very robust regarding their tuning: coefficients c_i and swarm size T [13].

From this historical and basic version, a number of variants were studied during the last decade to improve the performances. They are based on various strategies for determining the informants (fixed or random topology), population splits (memory vs explorer swarms), the proximity distributions used to guide the bounded motion of the particles, etc. Adaptive versions were also developed to save the user from having to set these elements *a priori*, by self-adjusting them during the run. Hence, a recent extension is aimed at having several tribes of particles cooperating, with the idea that many swarms should be more efficient to explore the different regions of the search space than only one [12]. The most promising techniques have been selected and implemented in the PSO code developed by ONERA to optimize the regressor positioning (which is part of the *koala* tool described below). A detailed description of this software is beyond the scope of this paper. Thus, only a brief survey of the main functionalities offered is given hereafter, with some references for readers interested in obtaining more details:

- *fixed and adaptive topologies* → from static (star, ring, Von Neumann) to dynamic ones (*e.g.* Delaunay neighboring) [24],
- *particle displacement* → standard, with constriction factor, FIPS and weighted FIPS versions of the velocity update laws [28],
- *hybrid local/global method* → to speed up the convergence with direct search (improved Nelder-Mead, Delaunay tessellation for the initial simplex),
- *multiswarm strategies* → for competing swarms or for partitioning the search domain into several subregions [51],
- *diversity analysis* → to provide information about the swarm dispersion and to refine the convergence tests [35],
- *swarm initialization* → from random to low discrepancy sequences (Hammersley, centroidal Voronoï diagram) [13],
- *competitive multirun* → to benefit from several topologies, algorithm variants and tuning,
- *charged vs neutral particles* → cooperation of particles with different physical properties [3].

A Kernel Optimization Algorithm for Local Approximation (KOALA)

The coupling of this PSO algorithm with the constructive approach based on forward selection allows structural and parametric optimizations to be proceeded to jointly for various types of regressors with local basis (as opposed to [10] where a basic version of PSO is used). In the *koala* software developed by ONERA for that purpose, this approach is applied to various kernel-based NN such as RBF and LLM, provided some adjustments of the orthogonalization method. LLM generalize RBF [34], by replacing the linear weights (scalar w for RBF networks) by an affine expression depending on the model inputs (see the dotted connections in Figure 2). They gave rise to the famous *LOLIMOT* algorithm (LOcal LInear MOdel Tree). It is thus expected that fewer radial units will be required to achieve the same accuracy in most applications. The generic formulation (2) used to represent LP models remains thus suitable, but needs only to be adapted to this case by using an extended set of regressors $r_l^{\#}$, because a new kernel now breeds a subset of regressors and not only one as it was for RBF:

$$\hat{y}_k = f(x_k) = \sum_{j=1}^{m}\left(\sum_{i=0}^{n} w_{ji} x_k^i\right) r_j(x_k) = \sum_{l=1}^{m(n+1)} w_l\, r_l^{\#}(x_k) \quad (11)$$

denoting by x_k^i (for $i = 1$ to n) the value of the i^{th} input variable for the k^{th} data sample, and setting $x_k^0 = 1$ to include the constant terms of the local affine modeling in the 2nd sum. Hence, it will remain possible to use the constructive algorithms developed for any type of regression, but some slight adaptations are required to take the peculiarities of the kernel functions $r_l^{\#}$ into account. When adding or subtracting terms, the group of regressors sharing the same kernel r_j needs to be considered as a whole, and no longer separately as was the case for RBF, polynomials, etc. In any case, the algorithm is aimed at gradually selecting a series of regressors by optimizing their kernel parameters, *i.e.* the ellipsoid center c and radius σ related to each radial unit in Figure 2 (both vectors). Considering the n explanatory variables x^i, the components of the vector u (particle position) can be ordered arbitrarily for instance as $u_i = c_i$ if $1 \le i \le n$, and $u_i = \sigma_{i-n}$ if $n+1 \le i \le 2n$. To sum up, the improved performance of the *koala* tool results from two complementary aspects: on the one hand, applying efficient OLS-based forward selection and SNLS optimization to a more powerful modeling (LLM) and, on the other hand, implementing a new PSO algorithm that outperforms the standard one.

Application #1: building efficient LFR for control analysis and design

For this first application topic, the motivation is to use surrogate models to improve the creation process of Linear Fractional Representations (LFR) of reduced complexity. It happens that the final LFR object relies on rational functions which hence can be considered as simple surrogate models, but actually we will see in this section that more complex surrogate models will be used during the building process to obtain these rational functions *in fine*. Let us remember that a LFR is a model where all of the known and fixed dynamics of a given system are placed together in a linear time-invariant plant M, while the uncertain and varying parameters are stored in a perturbation matrix Δ (Figure 3). LFR modeling is now a widely spread and a very efficient tool in the fields of system analysis and control design. It notably allows the robustness properties of uncertain closed-loop plants to be evaluated (*e.g.* using μ-analysis or Lyapunov-based methods), and robust control laws (especially using H_∞ approaches) or gain-scheduled controllers to be designed [56]. However, the efficiency of the aforementioned analysis and synthesis techniques strongly depends on the complexity of the considered LFR, which is measured in terms of both the size of the matrix Δ and the order of the plant M. An increase in complexity is usually a source of conservatism, and can even lead to numerical intractability.

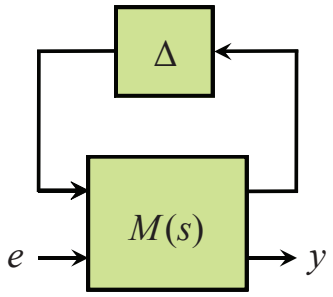


Figure 3 – LFR object

In most industrial applications, physical systems are described using a mix of nonlinear analytical expressions and tabulated data. Therefore, a two-step procedure has to be implemented to obtain a suitable LFR: a linear model with a rational dependence on the system parameters is first generated, and then converted into a linear fractional form. Several techniques exist, such as object-oriented realization, to perform the latter transformation. Although the minimality of the resulting LFR cannot be guaranteed, symbolic preprocessing techniques, as well as numerical reduction, usually enable complexity to be overcome. Efficient software such as the LFR Toolbox for Matlab© is also available (see [27] and references therein for an overview of LFR modeling).

On the other hand, the preliminary issue of converting the tabulated or irrational data into simple yet accurate rational expressions has been paid much less attention, although it is of significant practical importance. In the aeronautic field for example, most aircraft models include tabulated aerodynamic coefficients determined by CFD, wind tunnel experiments or flight tests, and several controller gains depend on the flight parameters in a tabulated fashion. The motivations for addressing the issue of tabulated data approximation are twofold. The first one is of a physical nature: computing parsimonious rational expressions, for which the number of terms in the numerator and denominator is as low as possible, is a natural way to prevent data overfitting and to ensure a smooth behavior of the model between the points used for approximation. On the other hand, building a LFR from a polynomial or a rational expression $f(x^1,\ldots,x^n)$ results in a block diagonal matrix $\Delta = \mathrm{diag}[x^1 I_{p_1},\ldots,x^n I_{p_n}]$. The number p_j of repetitions of each parameter x^j in Δ is strongly linked to the number of occurrences of x^j in f. Although this is not an exact rule, the trend is as follows: the fewer the occurrences of x^j in $f(x^1,\ldots,x^n)$ are, the smaller the size of Δ will be. In other words, no matter how efficient the LFR generation tools can be, they are of little help if the rational expressions to be converted are unnecessarily complex. Hence, the need to obtain tractable LFR for control analysis or design purposes is another strong motivation for generating sparse rational expressions.

For a given accuracy, an intuitive idea is to determine a rational function for which the numerator P and denominator Q are two polynomials of the lowest possible degrees. This fairly simple strategy is followed by most existing methods. A classical linear least-squares (LS) technique is notably implemented in the LFR Toolbox [27] in case the rational function is restricted to be polynomial. In the general case, a nonlinear LS technique, implemented for example in the

Curve Fitting Toolbox of Matlab©, strives to minimize the approximation error, whereas a Quadratic Programming problem solution [6] ensures that the resulting rational function intersects a set of intervals containing the data. However, all of these techniques suffer from the same drawback: all admissible monomials of P and Q are usually nonzero, regardless of their real ability to model the data. More generally, the question of which terms should be included in the model is often addressed by trial-and-error, or even ignored in practice. A way to deal with this question is to use orthogonal LS (OLS), which allows the ability of each monomial to efficiently model the data to be evaluated, and therefore only the most relevant ones to be selected leading to sparse expressions. This approach was applied in [32][33][41] to model aeronautical data with polynomials, but practical methods leading to rational expressions are still lacking. Yet, the additional degrees of freedom offered by such expressions are likely to lead to simpler expressions and thus to smaller LFR [43][44].

In this context, a new method has been developed to compute sparse rational approximants, *i.e.* as few monomials in P and Q as possible, by using an indirect approach that first builds a parsimonious model based on LLM networks, before translating the result into a fractional form. Note that a direct approach for computing a rational approximant in a single step thanks to a symbolic regression technique is proposed in [18], which uses another recent evolutionary algorithm (Genetic Programming) to select sparse monomials. A first idea for such an indirect approach capitalizing on the tool *koala* would be to convert (11) *a posteriori* into a rational form. By choosing Gaussian radial functions, this regression is expressed as the sum of m terms, the j^{th} one being for any x:

$$f_j(x) = \left(\sum_{i=0}^{n} w_{ji} x^i\right) r_j(x) = \left(\sum_{i=0}^{n} w_{ji} x^i\right) \exp\left[-\sum_{i=1}^{n} \frac{\left(x^i - c_{ji}\right)^2}{\sigma_{ji}^2}\right] \quad (12)$$

It is thus possible to use Pade approximants of the exponential function, so as to replace it by a rational function in reduced form $\exp_{[p,q]}$. The latter is expressed as the quotient of two polynomials of p^{th} and q^{th} degrees, and the corresponding approximant to $f_j(x)$ becomes a rational function of $(2p+1)^{\text{th}}$ and $2q^{\text{th}}$ degrees for every explanatory variable x^i. However, obtaining high-quality approximants (*e.g.*, decreasing rapidly to 0 as x^i increases) would require large values of q (with $q - p > 2$ or 3). Hence, the degree of the resulting rational function would be penalized, with no guarantee as regards the accuracy of the global regression $f(x)$. Consequently, a more relevant approach consists in replacing the exponential function straight away by such an approximant, and then using this new kernel form straightaway during the optimization of the regression. The simplest transform corresponds to the reduced form $\exp_{[0,1]}$, which yields a sum of m components like:

$$f_j(x) = \left(\sum_{i=0}^{n} w_{ji}\, x^i\right) \Bigg/ \left(1 + \sum_{i=1}^{n} \left(x^i - c_{ji}\right)^2 / \sigma_{ji}^2\right) \quad (13)$$

Accordingly, another class of models was added to the RBF/LLM kernels proposed by *koala*, based on the Pade approximant $\exp_{[0,1]}$. Thanks to this new form, it is also worth noting that the computational complexity of the surrogate modeling can be adjusted to limited coding requirements, if any, prior to on-board implementation. This remark does not apply to LFR models, but rather to other applications for which the exponential operator, for instance, is not available on aircraft computers (see Application #3 on embeddable models for virtual sensing of flight parameters).

Finally, it must also be mentioned that the post-processing of the resulting regression, prior to the derivation of the LFR, makes use of the Matlab© Symbolic Toolbox. Again, several options exist for gathering the m components $f_j(x)$ into a single rational function: global expansions of the numerator/denominator, factorization of the denominator, and sum of elementary rational terms. The latter appears to be the most relevant since it favors some simplifications when building the final LFR. A factor of two (in the scalar case) can usually be gained in the final LFR size. When the simultaneous modeling of several coefficients is considered, the benefit from using this approach is maximized with a LFR size usually reduced by a factor equal to the number of coefficients if the complexity of the nonlinear coefficients is like-for-like (and hence does not require a significant increase in the number of kernel functions).

This work takes place within the framework of a more general ONERA project aimed at developing a *Systems Modeling, Analysis and Control* (SMAC) toolbox [42]. This Matlab/Simulink© library is being developed to provide both researchers and control engineers with a complete set of tools for making the design, tuning and validation of control laws easier. More precisely, the purposes of the SMAC project are to control aeronautical vehicles throughout their whole flight domain in the presence of nonlinearities, uncertainties, external disturbances and imperfectly measured or estimated data, while obtaining strong guarantees w.r.t. the stability margins and the performance levels. A free version of SMAC can be downloaded from w3.onera.fr/smac, which includes three kinds of tools:

- **Modeling tools**, which allow the considered physical systems (usually represented in an industrial context by using a mix of nonlinear analytical expressions and tabulated data) to be described as a single parameterized model. Typically, they are aimed at creating accurate LFR with sizes that are as reduced as possible, in order to facilitate the subsequent use of the design and the analysis tools described in the next two items. The APRICOT library (*Approximation of Polynomial and Rational-type for Indeterminate Coefficients via Optimization Tools*) includes a set of optimization tools to convert numerical data into simple yet accurate polynomial or rational expressions [43][44], and notably implements the adaptation of the tool *koala* described in this section. A limited version is available at w3.onera.fr/smac/apricot, and can be applied to simple cases ($n \leq 2$, $N \leq 100$). The GSS library (*Generalized State Space*) then converts the resulting expressions into low-order LFR. Note that the GSS library replaces and extends the LFR toolbox [27], which is no longer maintained.

- **Control design tools.** The Convex Synthesis library is dedicated to convex synthesis of LTI and LFT feedback controllers using Youla parameterization. The SAW library is a collection of Matlab/Simulink© tools for *Saturated systems analysis and Anti-Windup design*, and the OISTeR library (*Output to Input Saturation Transformation extensions for Robustness*) allows output saturations to be transformed into input saturations, which can then be handled with the SAW library. An additional library will be available soon. It will combine robustified nonlinear dynamic inversion techniques, structured H_∞ synthesis and anti-windup compensation, so as to produce simple yet powerful controllers, which can be easily implemented but do not require any interpolation as is the case with classical gain-scheduling techniques.

- **Analysis and validation tools**, which permit the robustness properties of the resulting closed-loop systems to be evaluated. Let us mention the SMART library (*Skew-Mu Analysis based Robustness Tools*) [42], which implements most of the μ-analysis based algorithms developed by ONERA over the last 15 years, and allows systems with parametric uncertainties and unmodeled dynamics to be handled. A frequency-domain and a time-domain IQC-based techniques are also available in two dedicated libraries. They make it possible to consider time-varying parameters and hard nonlinearities such as saturations, deadzones and sector nonlinearities, in addition to model uncertainties.

To illustrate the use of some of the tools included into the APRICOT library, and relying on surrogate modeling to create LFR, a realistic example is processed (Figure 4) involving a set of 3 aerodynamic coefficients depending

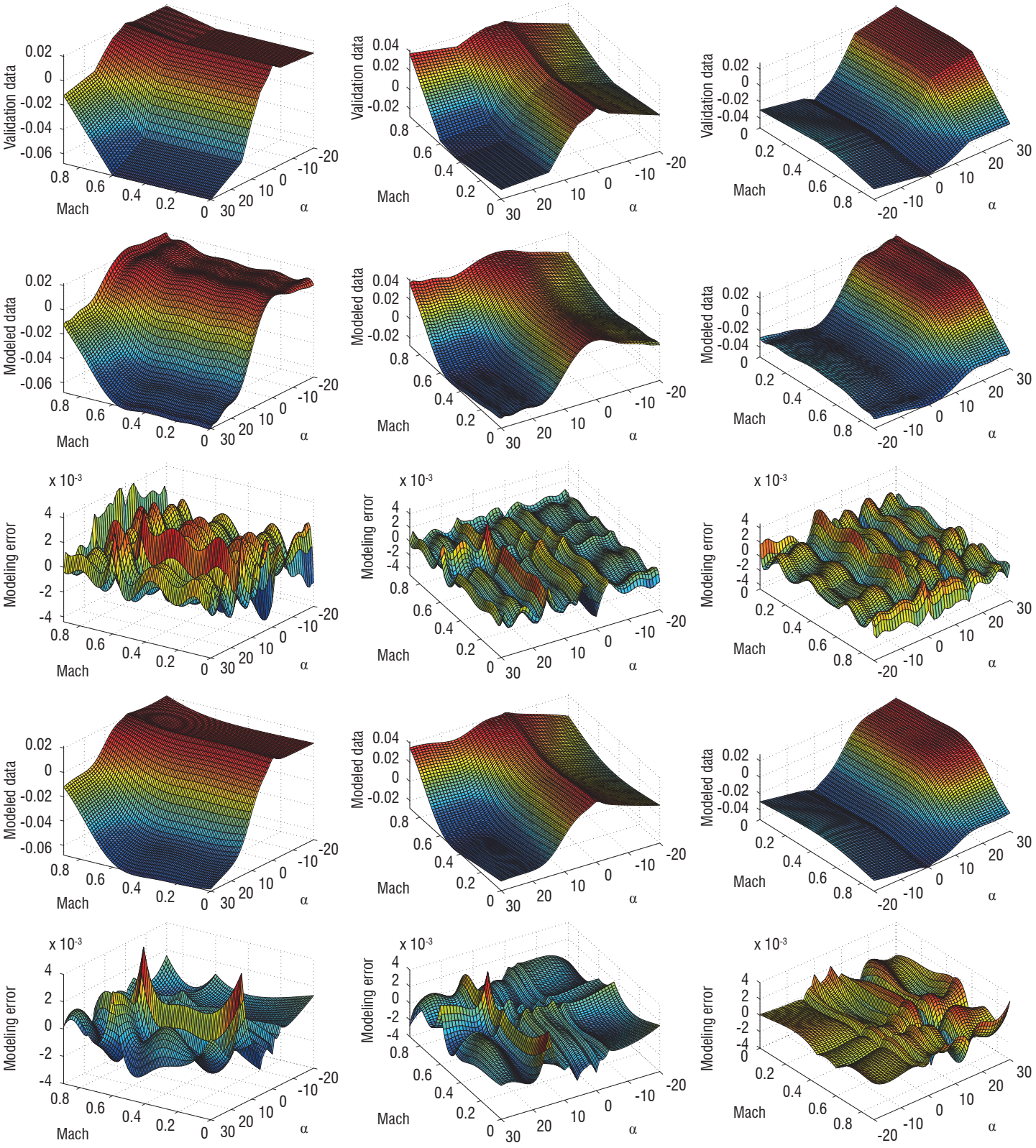


Figure 4 – Comparison of LFR results achieved by polynomial and rational models (1st row=reference data, 2nd and 3rd rows=polynomial approximations and modeling errors achieved by *olsapprox* 4th and 5th rows= rational approximations and modeling errors achieved by *koala*)

on 2 flight parameters (Angle of Attack α and Mach number). More details on this benchmark are available at w3.onera.fr/smac/?q=apricot_example1. These coefficients are depicted on a fine 50x90 mesh in the top row of Figure 4, a gridding which is used as a validation set to evaluate the approximation results achieved with a rougher 40x60 mesh of learning data. The middle rows of Figure 4 display the results achieved by using a polynomial surrogate model to create the LFR (routine *olsapprox* of the library which implements the same OLS-based forward selection to add relevant monomials one by one). A maximum degree of 12 is set for each explanatory variable and for the monomials, resulting in a size of 40 for the final LFR. On the other hand, the bottom rows of Figure 4 display the results achieved by using a rational surrogate model to create the LFR (routine *koala* described in the previous sections). A maximum degree of 12 is also set for a fair comparison, resulting in the creation of 6 kernels and in a size of 24 for the final LFR. In addition to providing a much smaller LFR size, the modeling accuracy is also better (improvements in the RMS errors of 33%, 15% and 25%, respectively). For this type of application, *i.e.* seeking a common model for a set of rather complex coefficients, using a powerful approach to construct a rational approximant presents a distinct advantage. Although the degrees of the polynomials would be globally the same if expanded, the conversion process of the rational expression into a LFR fully benefits from the factorized form of the rational function which is shared by all of the coefficients to be modeled.

Unfortunately, things are not that black-and-white and counter-examples prove that there is no definite answer for choosing either one or the other surrogate model. For example, when seeking LFR models in matrix form involving a set of low-complexity coefficients, the pros of rational expressions can be wasted during the final conversion stage into LFR form. In this case, accurate polynomial approximants can be achieved with low degrees, and can provide very satisfactory LFR with much smaller sizes. This is illustrated by the following aircraft benchmark, also available with the SMAC toolbox, corresponding to a civil transport A330-like aircraft. The goal is to obtain low-order but accurate open-loop LFT models covering a significant part of the operating domain (during the approach and landing phases), in order to perform control law validation and eventually to find worst cases for improving the design. To achieve this, a realistic nonlinear model of the aircraft is linearized at various values of the airspeed, altitude, mass, Center of Gravity location, and temperature. Accordingly, a family of 3000 linear models is obtained from which short-term longitudinal and lateral models are extracted. In this benchmark, the main challenge is not the complexity of the coefficients to be interpolated, but rather their high number (for longitudinal models, one has 26 varying coefficients distributed in a sparse 6x8 matrix), as well as the significant number of explanatory variables.

These two properties tend to generate high-order LFT models especially when using rational interpolation techniques. Consequently, the LFT-based representations are computed here by using polynomial interpolation only. A first and very accurate (global relative error = 0.5%) model is obtained by a standard LS technique (routine *lsapprox* of APRICOT) using polynomials of the third degree (resulting in full expressions with 55 monomials per coefficient). The global size of this first model is 109. Next, an enhanced OLS optimization technique is used (routine *olsapprox*)

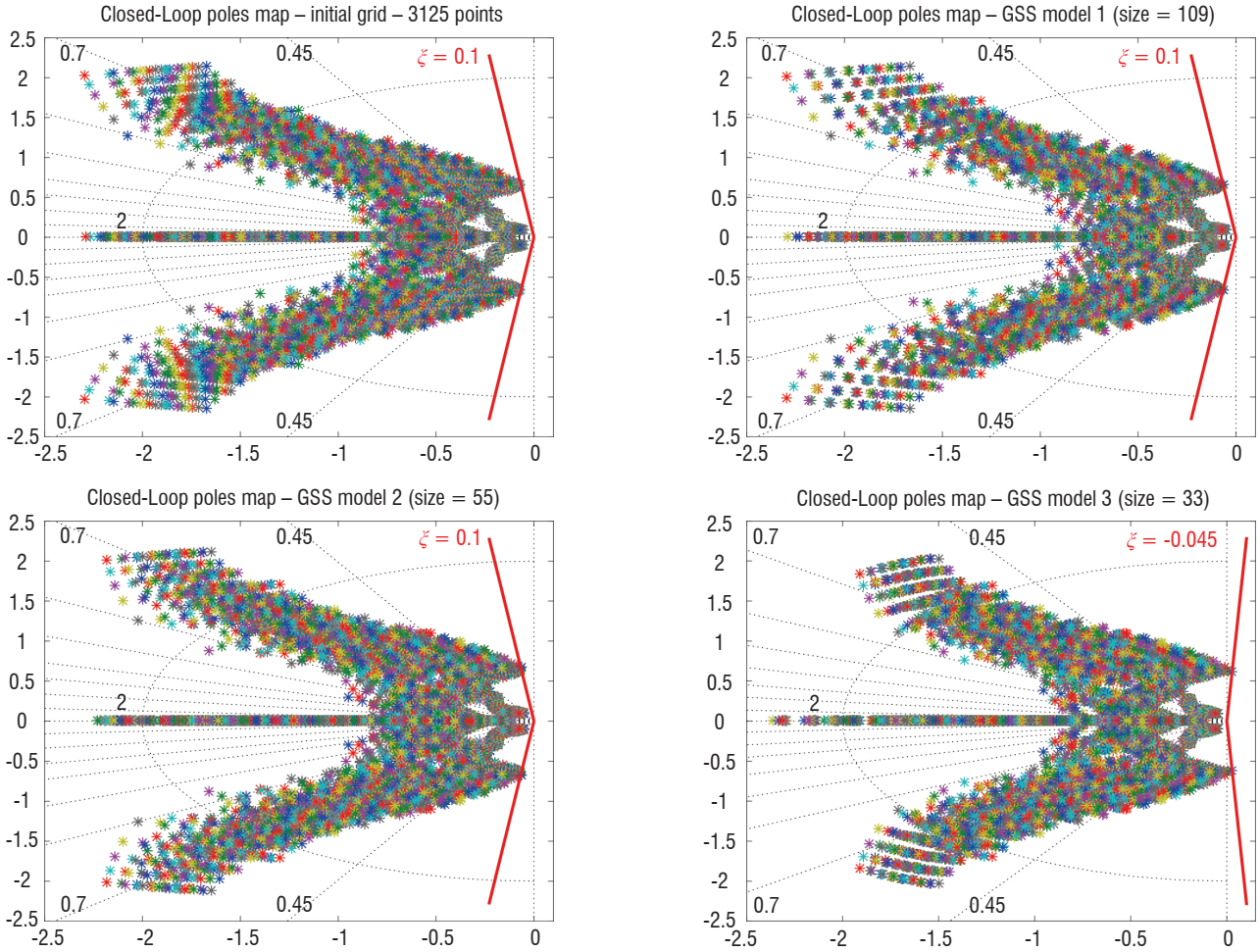


Figure 5 – Comparison of the longitudinal closed-loop poles of a civil aircraft over a large operating domain before and after LFT modeling (upper-left = reference data, upper-right = high-fidelity surrogate model, lower-left = medium-fidelity surrogate model, lower-right = low-fidelity surrogate model)

which enables the sparsity to be better controlled. The required accuracy is successively set to 1% and 5%, resulting in very sparse polynomials for the 26 coefficients with a number of monomials from 3 to 26 (10 on average) in the 1st case, and from 2 to 12 in the 2nd case (4 on average). Two reduced-order LFT models (referred to as GSS Models 2 and 3 respectively) are obtained. Their global sizes are significantly reduced to 55 and 33, which makes these models quite attractive for robustness analysis provided that the degraded accuracy does not compromise the analysis.

To evaluate the impact of the quality of the approximation, a preliminary validation of the models is easily performed by comparing closed-loop pole maps in the complex plane for a given basic control system designed to stabilize the nominal plant. The result of this validation is visible in Figure 5. The upper-left subplot shows the closed-loop poles directly obtained from the initial family of 3000 linear models. Next, the upper-right, lower-left and lower-right subplots respectively display the closed-loop poles corresponding to high, medium or low-order GSS models. Quite interestingly, one observes that the medium-order model (size 55) is still perfectly representative of the initial set. On the other hand, the lower-order model is not accurate enough in the whole operating domain; however, further investigations have shown that this model is still quite satisfactory in a reduced domain, and hence can be used for a more local validation. This benchmark shows that even simpler (polynomial) surrogate models can be obtained using the constructive OLS-based optimization techniques described in the previous sections, which are quite efficient to obtain low-order reliable LFR. The order reduction achieved by these advanced algorithms is essential for the robustness analysis tools. These will later provide stability and perfor-mance certificates more efficiently, or alternatively some worst cases useful to further improve the controller design.

Application #2: identification of aerodynamic non linearities

The concept of model identification refers to a set of tasks required to determine, and then to tune, a suitable modeling likely to explain the experimental behavior of the system. This involves choosing the type of mathematical relationships linking the i/o observed variables (often denoted as structural identification), as well as adjusting the unknown parameters of these equations (the so-called parametric identification). The early developments in system identification date back to the 70's, but this topic remains the subject of new developments nowadays, especially for aeronautics. They apply to the modeling of both rigid aircraft, described by the flight mechanics equations, and flexible aircraft where structural deformations are considered (see [5] and references therein). Owing to its mission, ONERA plays an intermediate role between academy and its main industrial partners. Within the framework of identification as in others domains, this role presupposes that new promising methods are investigated, adapted and transposed whenever necessary, in order to be evaluated through aeronautical applications. Accordingly, ONERA has been working with French aircraft manufacturers (Airbus and Dassault Aviation) on a large spectrum of themes and methods [5], applying these through a succession of industrial programs from the A320 to the A380, the Rafale, UAVs, etc. A close cooperation with the Flight Mechanics and Simulation Department of Airbus has been running for about 30 years, and has led to the implementation of several software codes in their identification toolbox for an operational use [25]. This continues nowadays through research programs aimed at improving the industrial process that allows the aerodynamic model to be fitted more rapidly into a larger part of the flight envelope.

Identification is aimed at producing accurate models to represent the behavior of the real aircraft. Given that pre-flight modeling obtained from CFD, wind tunnel, or ground tests is seldom faithful enough, this accuracy is achieved thanks to a final updating stage from a set of particular flight tests devoted to the ultimate model refinement. However, the aerodynamic models used for example for rigid aircraft, are also becoming drastically complex since from now on they integrate several effects that were disregarded before, or simply because the airplanes themselves have become much more complicated. Let us mention the A400M for the propeller blasts, or the A380 with an unprecedented proliferation of control surfaces. Hence, the identification becomes a much trickier task, but is also more crucial than it was in the past. Although physical models are available, derived from aerodynamics, structural dynamics or flight mechanics, quite often they cannot be implemented into the identification algorithms just as they are, because of high orders or strongly nonlinear behaviors. Hence, they require simplified grey-box representations to be developed, of linear or nonlinear types. These surrogate models facilitate the processing of the multivariate aerodynamic nonlinearities, usually complex and poorly structured. In addition, some constraints should be respected:

- the aircraft simulation requires continuous-time differential equations to be integrated,
- *a priori* knowledge about the predicted A/C behavior should be considered,
- a physical understanding and interpretation of the results is mandatory and introduces additional constraints into the optimization process.

The need to reduce the duration and the cost of the identification tests taking place during the first flights of a new airplane also requires specific techniques to be developed for designing, and then for processing this type of tests, without degrading the quality of the resulting models.

The variety of problems and models under consideration entails having a wide range of identification techniques available. Obviously, these include basic methods, such as LS or Maximum Likelihood and their variants, estimators based on Kalman filtering, etc. Practically, they result in 3 major categories of methods [5][23] for minimizing the Equation Errors (EE), Output Errors (OE) or Filter Errors (FE). Most of these methods are not directly usable as they are and need to be adapted to the peculiarities of aeronautical problems. For example, a very restrictive point comes from the requirement of coupling the identification tools with industrial simulators, in order to facilitate the implementation of the results and incremental updating of the aerodynamic data. This is all the more restrictive because these simulators also suffer from increasing complexity, involving a drift of the computational costs which is hardly compatible with the number of simulations required by the identification procedure. Another major concern arises from the will to automate and to systematize the test processing, to reduce its duration and to facilitate its progress. Actually, the whole set of available flight tests represents a huge amount of data, and hence some semi-manual steps in the process are especially tedious for the engineers responsible for sifting through the data. Automation also enables a global and joint processing of many tests, as well as the gradual introduction of new tests as they become available during the preliminary test campaigns.

Techniques based on linear or weakly nonlinear models are efficiently used as a first step of the identification process by processing only tests flown under close flight conditions [25]. However, when a global model is sought, valid over an extended area of the flight domain and including all aircraft specific nonlinearities, appropriate approaches and techniques must be applied. The task is all the more complex because the aerodynamic non-linearities are only available in the form of multivariate look-up tables (depending on configuration, Mach, AoA, sideslip, deflection angles, and dynamic pressure) which are not suited for identification algorithms. In industry, this global modeling is typically obtained after a long iterative process mainly based on EE approaches, and the result can be dependent on the skill of the performing engineers. For that matter, ONERA has been developing a so-called *hybrid identification* approach for several years, which might also achieve a more automated processing of the flight data. The ultimate goal would be to tune all of the model parameters in a single step (linear and nonlinear ones) using all of the available test data. The concept of *hybrid identification* refers to the hybridization between classical algorithms and specific techniques based on surrogate models intended to handle the aerodynamic nonlinearities. The RBF/LLM networks described in the previous sections are implemented for that purpose. They are particularly well suited for modeling complex and unstructured nonlinear systems, whether static or dynamic ones [11][47][52][55]. In the hybrid approach, they are typically used to replace the look-up tables describing the various coefficients, *e.g.* those appearing in equations similar to (1) [26]. This allows an algorithmically efficient identification to be performed and, additionally, this does not bias the results against *a priori* knowledge (*e.g.*, a predefined look-up index). This kind of implementation is grey-box type since it preserves the physical meaning as well as the structure of the aerodynamic developments used by industry [4][22].

The methods classically used (EE/OE/FE algorithms) have been implemented in a tool developed in cooperation between Airbus and ONERA (named *IdPy*, see Box 1), and this has been achieved through extensive algorithmic adaptations (Figure 7). Two complementary options are available for the identification: either time dependent coefficients are directly compared with their counterpart extracted from flight data throughout the sequence of tests available (EE), or they are integrated into the flight dynamic equations in order to minimize the errors between measured and simulated state variables (OE/FE). Both approaches benefit from the analytical and differentiable formulations of the surrogate models, which make it possible to perform quasi-exact parameter optimization (unlike purely numerical approaches using finite differences). Much CPU time is also saved for computing the derivative estimates required by the sensitivity equations, which is quite valuable especially for the greediest algorithms (OE/FE).

Whatever the approach, it is crucial for the aerodynamic corrections to remain physically acceptable. Practically, the purpose of an

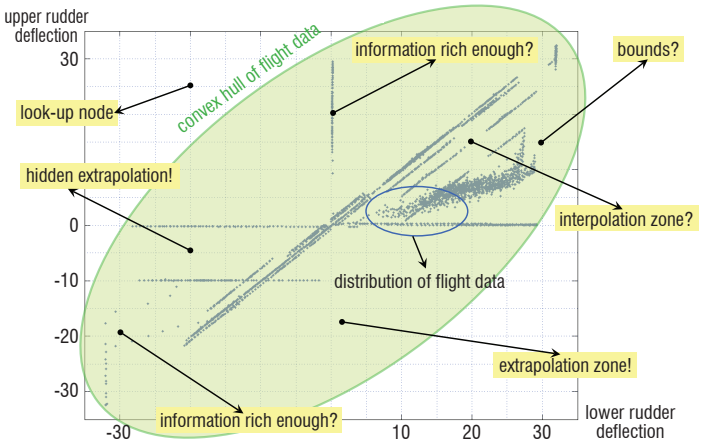


Figure 6 – Some traps and issues related to heterogeneous data distributions

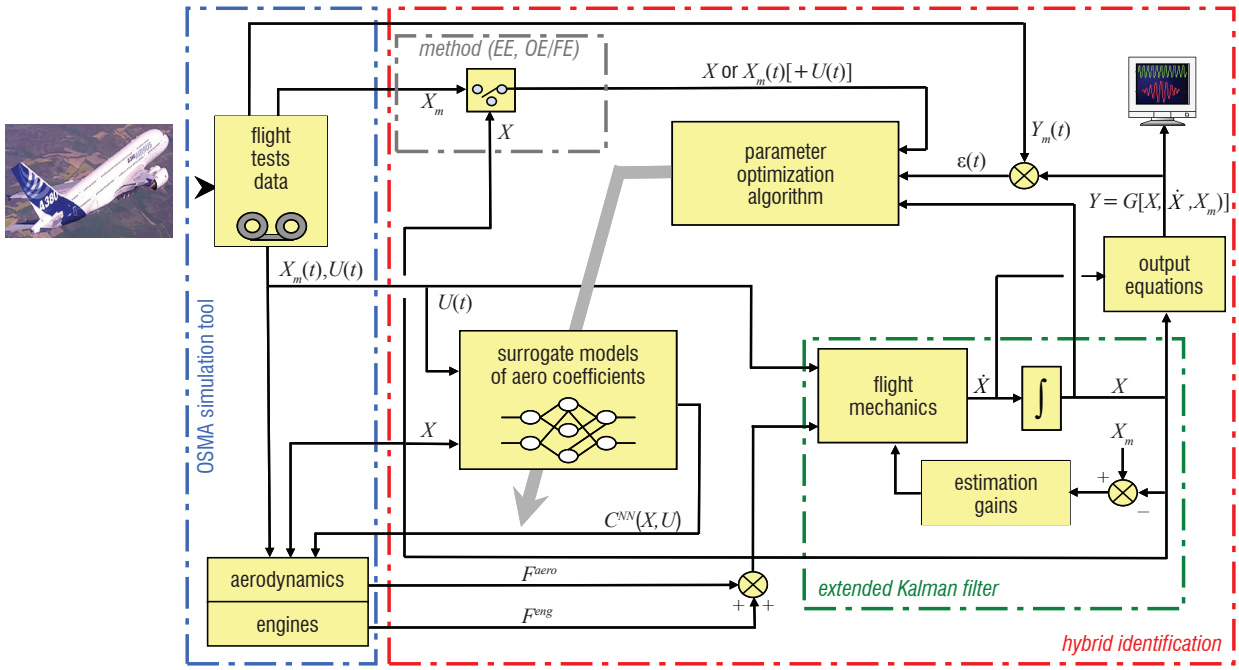


Figure 7 – Principles of the hybrid identification (X = simulated states, X_m = frozen states, Y = simulated outputs, Y_m = measured outputs, U = control deflections, F^{aero} = aerodynamic forces/moments, F^{eng} = engine forces/moments, G = measurement equations, C^{NN} = surrogate model of aerodynamic coefficients)

Box 1 - Identification of the A350 pitching moment in clean configuration

The following results have been achieved by using the tool *IdPy*, which stands for **Id**entification with **Py**thon, which is the most recent tool jointly developed by Airbus and ONERA. Compared to the previous tools, the novelty provided by *IdPy* comes from directly using OSMA, the Airbus simulation tool [4][22], thanks to a python interface that provides all of the functionalities required to match the simulation loop and the algorithm needs. In that way, the computations done by the identification tool are in line with the end user requirements. This common interface is a key point for spreading the results effectively to the operational context and for using other Airbus post-processing tools. *IdPy* follows an agile process: any new Airbus requirement to deal with a new issue can be implemented, tested by ONERA and validated by Airbus within a few weeks. *IdPy* is mostly generic since all Airbus models are operated through OSMA; hence, *IdPy* can be deployed very quickly for any aircraft, from the oldest A320 aircraft to the most recent A350 aerodynamic model. *IdPy* also offers new capabilities to operational end users with respect to the legacy Airbus tools: to run a complete identification process chaining EE, OE and FE minimizations within the same framework, and to process large flight data sets since the amount of flight test maneuvers has tripled during the last decade.

This user case corresponds to the pitching moment Cm, as expressed by (1). The 3 rigid coefficients to be estimated are the moment at zero lift $Cm_0 = f(M)$, the aerodynamic center position $x_{AC} = f(M)$ and the non-linear term $\Delta Cm_{NL} = f(\alpha, M)$. In the aerodynamic model, these coefficients are in the form of look-up tables; 15 parameters describe the linear coefficients and 868 parameters are required for the non-linear coefficient. For each coefficient, a surrogate model can be used as a grey-box to estimate an additional term w.r.t. the pre-flight model, following an EE approach to start with the identification process. Two look-up tables are set for the linear coefficients Cm_0 and x_{AC} with 13 parameters, and a LLM network is introduced for the non-linear coefficient ΔCm_{NL}. In the present case, the initialization of the network does not make use of a constructive approach, which would result in a more parsimonious model. This is a simple mapping of the explanatory variables done by hand, just considering the input space, the output, the flight domain achieved and the extrapolation objective. Figure B1-1 superimposes the initial kernel distribution with the flight test data in the input space of the non-linearity (AoA, Mach). The density of kernels is raised at high Mach number (green ellipsoids) to ensure enough flexibility during the updating process, whereas a few kernels are set up beyond the domain explored during the tests (blue ellipsoids) to facilitate a recovery of the pre-flight values (orange ellipsoids). Even with this straightforward parameterization, the number of parameters to be processed is 129, much lower than the 868 ones required by the original look-up table.

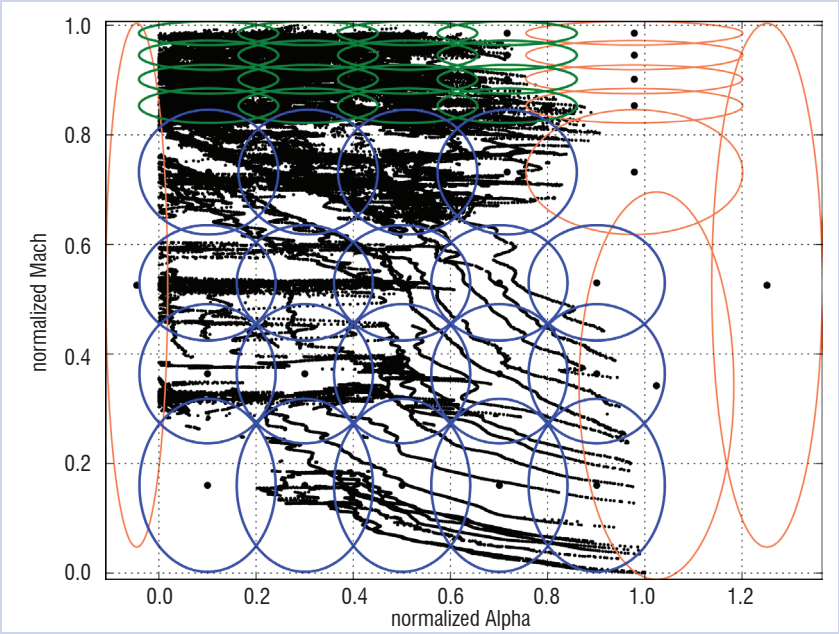


Figure B1-1 – Kernel and data distribution in the (α, M) plane

With regard to the flight test data, 374 flight test maneuvers have been selected for this identification step corresponding to over 2×10^5 data samples (about 6 flying hours). The identification is done in a single step for both linear and non-linear coefficients. Additive biases on the pitching moment are also estimated for each maneuver to account for shifts in the measurements. Several constraints are also added into the process, to get suitable updates and to regularize the optimization:

- to constrain the model output to zero in the linear part, assuming that the pre-flight linear domain was properly defined from CFD computations,
- to smooth the behavior of the surrogate models (look-up tables and LLM network),
- to extrapolate the estimation results in the boundary parts of the flight domain covered by the available data,
- to smoothly connect the updated and pre-flight models in the far-off regions.

The identification results obtained within about one hour of computation time, and involving 3 optimization iterations, are shown in Figure B1-2. The left part of the figure corresponds to the pre-flight "predicted" nonlinear coefficient ΔCm_{NL}, and the right part to the updated one.

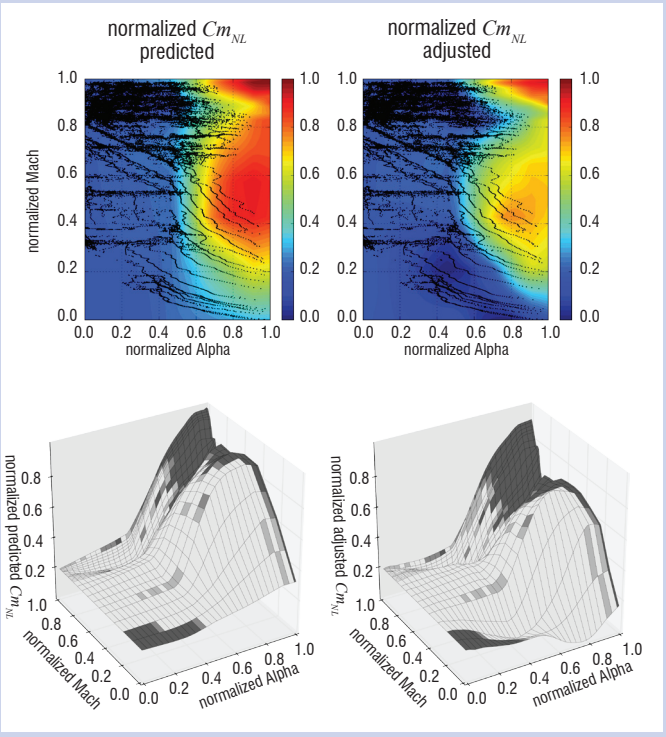


Figure B1-2 – Comparison of pre-flight and identified coefficients ΔCm_{NL}

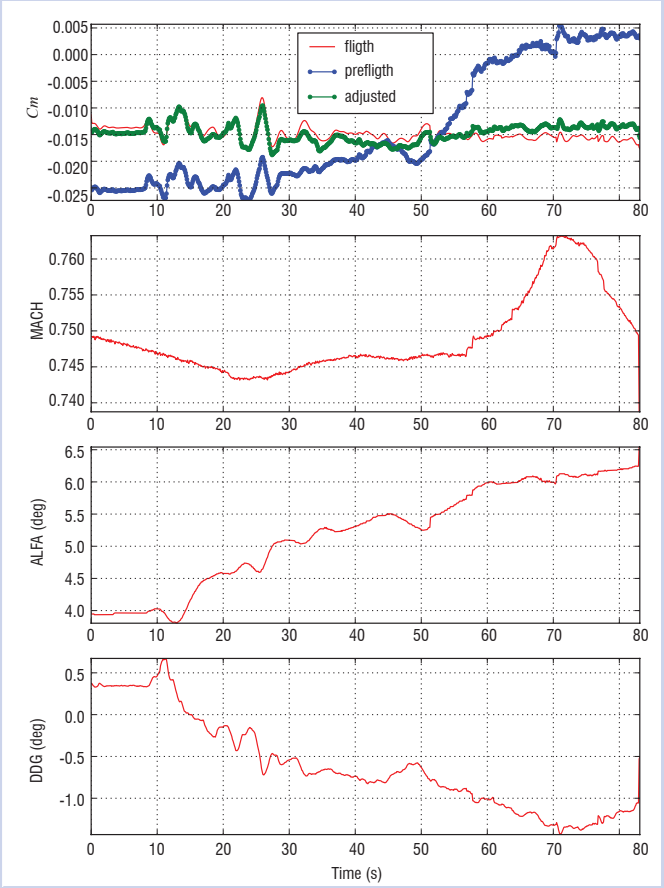


Figure B1-3 – Pitching moment during a turn maneuver with pitch-up at high Mach (from top to bottom: pitching moments, Mach, AoA, elevator deflection)

The heterogeneous data distribution is superimposed to the contour lines of the coefficient in the upper part of Figure B1-2, and is also displayed via a color code in the 3D plots (from white to black mesh depending on the proximity of the data samples). Finally, the plots of Figure B1-3 compare the pitching moment reconstructed from the flight data with the pre-flight and updated values, for one maneuver among the 374 flight tests. This user case also shows that the industrial requirement of coupling the identification tool with the Airbus simulator OSMA has been satisfied, and is not so detrimental to the industrial process since a large set of data can be managed within a reasonable time, which was not possible with the previous *standard* Airbus tools.

automated identification over large areas of the flight domain has raised a new need: specifying constraints to be followed by the non-linearities (*i.e.*, the outputs of the surrogate models). Constraints are a way to compensate for insufficient or sparse test data (Figure 6) or to introduce some kind of expertise into the problem. For instance, freezing the output levels may be required in some zones (*e.g.*, ΔCm_{NL} in (1) should remain zero at low AoA and low Mach, so that it does not interfere with the other terms of the development). It may also be desired to smooth the nonlinearities, or to connect identified and pre-flight models in areas where no flight data is available, and hence *a priori* knowledge from CFD or wind tunnel tests should be preserved [31]. These constraints can be enforced by robust regression mechanisms relying on various forms of penalties w.r.t. constrained values, smoothing, regularization, etc. This in turn raises the question of choosing and tuning these hyper-parameters, which should also be as automated as possible.

All of these issues, which are drawn from an identification stage of the A380 rudder efficiency, are illustrated by Figures 6-8. The coefficient represented in Figure 8 corresponds to a strongly nonlinear effect, contributing to the yawing moment gradient due to the rudders. For the A380 aircraft, this is a 3D coefficient depending on the sideslip angle and on the rudder deflections (lower and upper surfaces). Figure 6 corresponds to a cross-section view of the distribution of the available flight test data for a given value of the sideslip angle (*i.e.*, a thin slice of the 3D volume). The green ellipsoid is a 2D section of the convex hull of the flight data, and the background grid represents the look-up nodes of the pre-flight look-up table. To facilitate the industrial process, incremental updates of the aerodynamic data must be provided to be used by the OSMA simulator, and this requires the estimated corrections to be projected onto these look-up nodes. The issue highlighted by Figure 6 is that this projection should only be validated in some regions, due to the heterogeneous distribution of the input data in the domain.

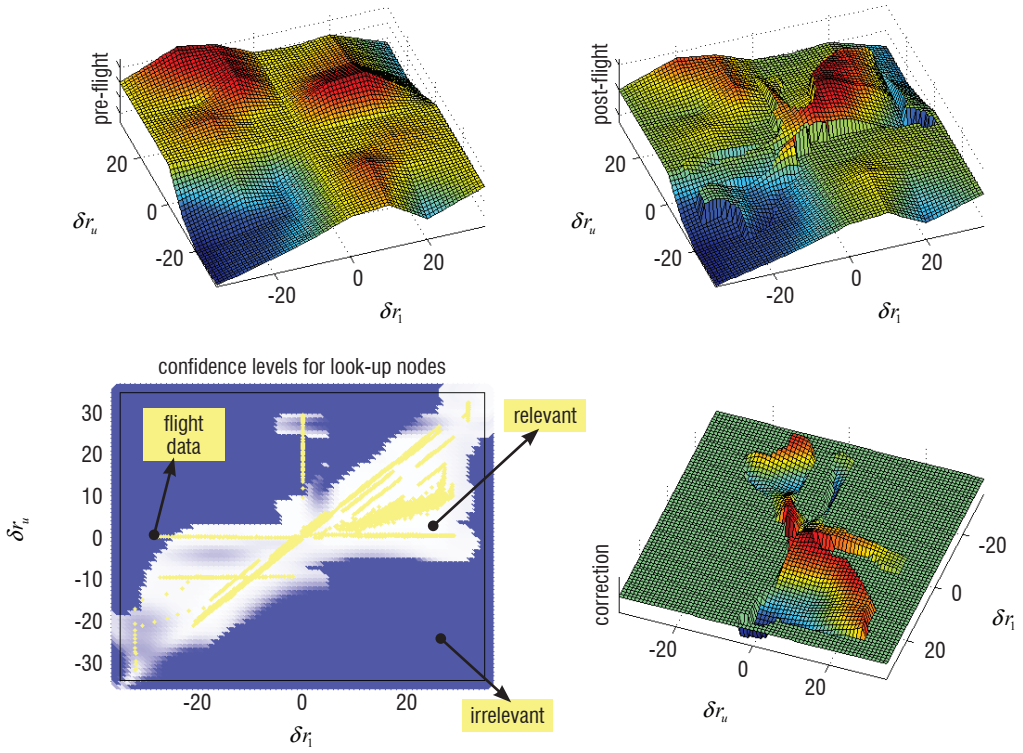


Figure 8 – Identification of the rudder efficiency on the yaw axis: sectional view of a 3D coefficient $C = f(\beta, \delta r_l, \delta r_u)$

To make this decision and to detect hidden extrapolation zones for instance, confidence levels can be associated with the identification results by computing useful indicators like the leverages (diagonal terms of the hat matrix [2][30]). Such confidence levels are displayed in the left bottom part of Figure 8, with a color code going from white to blue according to the local relevance of the results. Since this plot is only a sectional view of a more general 3D computation, the flight data displayed in yellow in this plane do not account on their own for these confidence levels. Finally, to smoothly connect the identified and pre-flight models, a robust regression procedure is used during a post-processing stage, involving a weighted LS optimization for which the penalties are inferred from the confidence levels, and an extended criterion is defined by introducing regularization constraints on the curvature of the outputs. Figure 8 illustrates the results achieved by this process with a surrogate LLM-type model (18 kernels) optimized using the *koala* tool described in the previous sections. From the right bottom plot of Figure 8, it can be seen that the pre-flight surface is reshaped by incremental updates only in the informative zones.

Application #3: embeddable models for virtual sensing

The development of electrical FCS, known as Fly-By-Wire, and the increasing level of automation have resulted in advanced capabilities for detecting, protecting and optimizing A/C flight guidance and control [50]. However, this higher level of automation requires the availability of some key flight parameters to be extended, particularly both aircraft AoA and calibrated airspeed, to keep the nominal flight control laws activated, thus avoiding any switch to degraded control modes (*e.g.*, *alternate* or *direct* laws as opposed to *normal* flight control laws for most of civil transport A/C). Hence, the monitoring and consolidation of those signals are significant issues, usually achieved via many functionally redundant sensors, allowing the way in which those flight parameters are measured to be directly enlarged (*hardware redundancy*). Unfortunately, this solution is detrimental to aircraft weight, power consumption and space requirements, and requires extra maintenance needs. Another alternative consists in benefiting from specific physical and dynamical relationships which link some A/C states. Advanced signal processing techniques can achieve real-time estimates of the critical flight parameters and yield dissimilar signals, by merging all or part of the available measurements through a model-based simulation of the aircraft flight mechanics (*analytical redundancy*). The resulting filtered and consolidated information is delivered under unfaulty conditions by estimating an extended state vector including wind components, measurements biases and modeling errors, and can replace failed signals under degraded conditions (*virtual probes*). To this aim, adaptive estimation schemes, which make use of integrated Fault Detection and Diagnosis (FDD) methods, can be developed in order to provide to the FCS with the ability to accommodate itself to potential sensor failures automatically.

Together with its industrial partners, ONERA has been working on this topic for a decade now and has exploited the potential and developed adaptive versions of several nonlinear state estimation techniques. This is the case of the well-proven Extended Kalman Filtering (EKF) [19][20][48], which permits basically all available sensor information to be merged while enhancing the measurements through a modeling that can describe the expected A/C behavior with more or less accuracy. The resulting estimates can be processed in parallel with the measured signals, thus contributing to extend the way the flight parameters are both monitored and consolidated. Several schemes allow that extra information to be introduced into the overall FDD architecture (see Figure 9): processing the virtual sensor just like the real ones, using it only in duplex or simplex modes (after one or several faults have already occurred), or using it to form an $(n+1)$-plex

sensing channel (*e.g.*, quadruplex in the case of 3 available sensors). The use of analytical redundancy for FCS is not new, as opposed to virtual sensors which have been studied more recently [36]. Practically, a major obstacle to an operational use of this kind of approach springs from their implementation in real time on embedded computers and especially from the computational burden of the associated algorithms. This issue is of primary importance when dealing with civil transport aircraft due to the certification purposes which impose stringent requirements to be fulfilled. That is why, to limit the complexity, surrogate models have been widely harnessed to derive simplified representations of the A/C aerodynamics (and possibly of the thrust forces). In addition, the selected analytical models, of an LLM type, offer other substantial benefits for encoding the EKF-based estimation algorithm, while remaining consistent with the implementation constraints related to the FCS for aviation industry.

Subsequently, this compliant adaptive estimation algorithm embeds a dedicated nonlinear state space representation, denoted as M_{NL} in the sequel, which describes the flight dynamics. This is derived from both kinematic relationships and longitudinal flight mechanics equations (see [19] for more details). In order to obtain an accurate and suitable model, the most relevant rigid-body aerodynamics, as well as static aeroelastic and possibly propulsion effects, are modeled using LLM networks. Such a dynamic modeling can be mathematically formulated as follows:

$$M_{NL}:\begin{cases}\dot{x}(t)=f\left(x(t),y_m(t)\right)+v(t)\\ y(t)=g_{LLM}\left(x(t),u(t),y_m(t)\right)+w(t)\end{cases} \quad (14)$$

In (14), $x(t)$ and $y(t)$ designate the state and output vectors of the nonlinear state space representation respectively; $u(t)$ refers to the input vector and gathers the deflection of all aircraft control surfaces (ailerons, spoilers, rudder, elevators and horizontal stabilizer); $y_m(t)$ (subscript m will refer to the measurements in what follows) is a vector of measurements, which includes also the time histories of several measured flight parameters, especially those of the longitudinal/lateral variables whose dynamics are not modeled in $\dot{x}(t)$. Subscript LLM indicates that the analytical expressions associated with the nonlinear output equations contained in function g_{LLM} exhibit the LLM approximations used to model at least the aerodynamic lift force, plus the total static mean gross thrust force and pitching moment, depending on the filter formulation. The process equations appearing in f correspond to a set of common kinematic relationships, augmented by slowly time-varying parameters (accounting for the accelerometer biases, the wind speed components and the modeling error in the lift force coefficient) whose dynamics are represented by random walks. Vectors $v(t)$ and $w(t)$ are process and observation Gaussian white noises. They are introduced to model errors in both the modeling M_{NL} and the measurements $y_m(t)$. It is assumed that they are characterized by zero-mean, uncorrelated and mutually independent processes with estimated covariance matrices denoted by Q and R respectively.

Based on previous nonlinear flight dynamics modeling introduced in (14), the proposed Adaptive EKF (AEKF) algorithm resumes, but also adapts, the main processing steps of the nonlinear Kalman filtering theory. Considering the standard filtering equation given in continuous-time (see [23]) with $p \in N^*$ measurements available every dt, and assuming that the correction term remains constant over $[t_k;t_{k+1}=t_k+dt]$, one can derive:

$$\dot{\hat{x}}\left(t_k|t_k\right)=f\left(\hat{x}\left(t_k|t_k\right),y_m\left(t_k\right)\right)+\left[\frac{K^i\left(t_{k+1}\right)}{dt}\times\begin{Bmatrix}z^i\left(t_{k+1}\right)\\ -g^i_{LLM}\begin{pmatrix}\hat{x}\left(t_{k+1}|t_k\right),\\ u\left(t_{k+1}\right),y_m\left(t_{k+1}\right)\end{pmatrix}\end{Bmatrix}\right]\begin{matrix}i\leftarrow 1\\ i\leftarrow 2\\ \vdots\\ i\leftarrow p\end{matrix} \quad (15)$$

with: $\hat{x}\left(t_{k+1}|t_k\right)=\hat{x}\left(t_k|t_k\right)+\int_{t_k}^{t_{k+1}}f\left(x(t),y_m(t)\right)dt$

and then: $\hat{x}\left(t_{k+1}|t_{k+1}\right)=\hat{x}\left(t_k|t_k\right)+\int_{t_k}^{t_{k+1}}\dot{\hat{x}}\left(t_k|t_k\right)dt$

In the differential equation (15), the gain matrix K is computed as in the discrete-time case. (15) also shows that the correction step is processed sequentially, measurement by measurement, which permits the computational complexity to be alleviated significantly. Indeed, this procedure avoids the matrix inversion traditionally required for the calculation of the correction gain matrix K, and replaces it by p scalar division(s). This sequential update is strictly equivalent to the standard computation of the EKF gain (which involves a matrix inversion each time new measurements become available) as both process and output noises are assumed to be uncorrelated (diagonal estimated covariance matrices Q,R). Then, given that sensor reliability can be periodically assessed by means of specific fault detection techniques (FDD is a component of the AEKF), the calculation of the K matrix in (15) can be adapted continuously over time, in case of detected and isolated faulty measurements, by simply omitting them

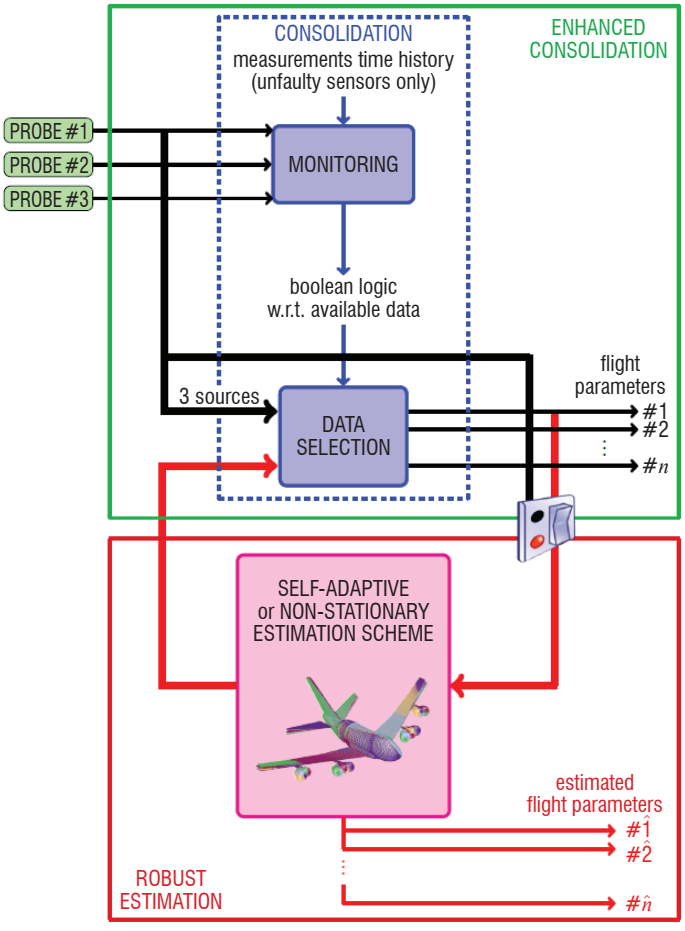


Figure 9 – Consolidation scheme using a virtual sensor

within the sequential update process. Usually, another difficulty for implementing such an estimation scheme on embedded computers arises from the double linearization that must be performed in real-time for both the system process (around $\hat{x}(t_k|t_k)$) and the output equations (around $\hat{x}(t_{k+1}|t_k)$). Resorting to numerical finite differences does not appear to be a viable solution w.r.t. the capabilities of current on-board computers since this implies multiple calls to the nonlinear flight dynamics modeling M_{NL}, as well as numerous multidimensional save/restore points. Fortunately, regarding the details of the equations which compose both nonlinear f and g functions, the linear tangent matrices $A = \partial f / \partial x$ and $C = \partial g_{LLM} / \partial x$ can be obtained analytically by means of:

- resorting to kinematic relationships to describe A/C dynamics (*i.e.*, in f). The resulting set of differential equations is known to be fully differentiable analytically w.r.t. the state vector $x(t)$ considered. Consequently, the analytical expressions of all A matrix coefficients can be established *a priori*, and their values can be calculated over time as the estimated state vector $\hat{x}(t_k|t_k)$ is continuously updated,

- using LLM networks which have the property of also being analytically differentiable w.r.t. their inputs (explanatory variables). Thus, given that the g_{LLM} function corresponds to a set of very simple static relations (based on standard trigonometric functions or vector norm calculations), except for the vertical load factor equation which relies on prior knowledge involving a LLM of the lift force coefficient, the nonlinear observation equations can again be differentiated once, in an analytical way, in order to derive the C matrix. The latter will be refreshed as the predicted state vector $\hat{x}(t_{k+1}|t_k)$ is periodically calculated.

These remarks make possible an off-line derivation of all of the analytical expressions associated with the terms composing both A and C matrices. Once all of these mathematical expressions have been determined, their valuation is carried out naturally over time through a single call to the internal model of the estimator. Besides, in most aircraft applications, these matrices will also appear to be sparse, thereby reducing the numerical complexity of the proposed approach slightly more. Deriving the Jacobian matrix $C = \partial g_{LLM} / \partial x$ is not as straightforward as for A. Indeed, the nonlinear output equations used for the prediction step of the AEKF require prior knowledge of the A/C aerodynamics. However, the latter can be limited to the aircraft lift force coefficient, which is generally represented by a complex nonlinear modeling in order to be accurate enough. Although surrogate models offer an efficient and powerful solution for representing such complex physics, there is also another benefit of using such analytical modeling, related to differential calculation. Let us consider the general LLM representation with $N_i \in N^*$ explanatory variables, $N_o \in N^*$ outputs and $N_c \in N^*$ kernels:

$$\forall k \in [1, N_o] \qquad y_k(e) = \sum_{i=1}^{N_c} \left(\sum_{j=0}^{N_i} w_{ij} e_j \Big/ \left(1 + \sum_{k=1}^{N_i} \frac{(e_k - c_{ik})^2}{\sigma_{ik}^2} \right) \right) \qquad (16)$$

and: $y = [y_1 \;\; y_2 \;\; \dots \;\; y_{N_o}]^T$

In (16), the set of parameters $(w_{ij}, c_{ik}, \sigma_{ik})$ must be optimized so that the output vector y match some reference data for given input vector values $e = [e_1 \;\; e_2 \;\; \dots \;\; e_{N_i}]^T$. Given that the exponential operator is not available in most dedicated versions of the SCADE© formalism used in the aviation industry, the LLM kernels adopted in this application are based on the Pade approximant $\exp_{[0,1]}$ (leading to a set of rational function(s)). This illustrates that the computational complexity associated with any modeling involved in new potential solutions can be adjusted to limited encoding capabilities, despite their algorithmic sophistication. Subsequently, it is noteworthy that the partial derivative of any given LLM output $y_k, k \in [1, N_o]$ with respect to the vector e of explanatory variables, the so-called sensitivity, can be easily computed from (16) s.t.:

$$\forall k \in [1, N_o] \qquad \frac{\partial y_k}{\partial e} = \sum_{i=1}^{N_c} \frac{\partial}{\partial e} \left(\frac{\sum_{j=0}^{N_i} w_{ij} e_j}{\left(1 + \sum_{k=1}^{N_i} \frac{(e_k - c_{ik})^2}{\sigma_{ik}^2} \right)} \right) \in M_{1 \times N_i}(R) \qquad (17)$$

It is of primary interest to note that the differentiation of the Pade approximant in (17) leads to an analytical expression which again depends on parameterized rational functions. Therefore, the calculation of $C = \partial g_{\mathrm{LLM}} / \partial x$, which combines standard differentiation developments (analytically known thanks to flight mechanics equations) and LLM sensitivities as in (17), can also be carried out analytically as for $A = \partial f / \partial x$. Consequently, this adaptive EKF-based estimation algorithm can be implemented on-board with a low complexity, and used to reconstruct any faulty information so as to maintain nominal flight control laws.

As was aforementioned, the adaptation of the nonlinear state estimation is managed by a dedicated FDD technique since the EKF solely becomes unsuitable on its own in faulty cases, given that it assumes noisy, but healthy, measured signals. This FDD technique can take on several forms and can result from any signal- or model-based approach provided that it ensures both fault detection and isolation capabilities. To say a word about the detection and adaptive part of the method, a candidate EKF reconfiguration method has been experienced in [20] to monitor the measured signals from sensors and reduce the estimation errors induced by those potential incorrect measurements to recover acceptable performances. This method accomplishes both detection and isolation of specific single or multiple abrupt faults (constant bias, stuck value, strong drift, etc.) before adjusting the sequential measurement update process of the EKF.

The FDD relies on Fisher-Snedecor statistical hypothesis tests, recursively processed in order to periodically assess the reliability of the sensors. In the case of invalid test results, this pure signal-based technique will instantaneously detect and isolate unexpected sensor faults characterized by High-Frequency (HF) signatures beyond the A/C dynamics bandwidth. All declared and identified faulty measurements will then be denied for data fusion, thus achieving fault-free estimates of the key flight parameters, namely the AoA and the Calibrated AirSpeed (CAS). In a few words (see [46] for more details), these stages are accomplished using several signal processing operations which consisting, firstly, in extracting the HF components from the considered time-varying measurements and, secondly, in applying an appropriate inverse Auto-Regressive whitening filter. Then, a statistical analysis of the resulting HF residues compares a given reference variance (characteristic of a healthy signal) with a time-varying one (calculated over a sliding window), and permits the AoA and CAS states to be determined (healthy or faulty measurements).

The AEKF methodology was evaluated with data drawn from real flight tests performed on a civil transport aircraft (A340-600). These data correspond to an operational flight path profile, comprising

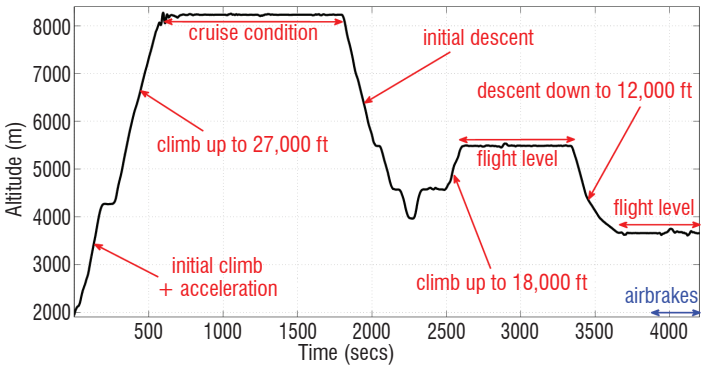


Figure 10 – Flight profile of the real test data

climb, cruise and descent flight conditions with heading changes (see Figure 10) which span more than 1 hour. The results gathered after processing these real flight data are presented in Figure 11. The simulated additive AoA and CAS faults (constant biases) are both applied at time t = 20 sec, and are quickly detected and isolated by the FDD mechanism used for sensor monitoring (detection time less than 1 sec). The results show that the mixed predicted/estimated flight parameters reconstructed afterwards remain valid over a long time horizon (more than 1 hour) under realistic changing flight conditions after the sensor reliability is affected by a double failure. With respect to these real data, the residual errors associated with the parameters to be reconstructed reach maximum peak values equal to 1° for AoA and 15 knots for CAS in case of a double fault. It is worth noting that these results also prove the robustness of the developed approach w.r.t. the modeling errors. Actually, the internal model used by the estimator corresponds to a simplified one resulting from a LLM representation of only a subset of the aerodynamic coefficients. Moreover, several effects (*e.g.*, the lateral ones) were ignored (ailerons) or simply approximated (spoilers). Besides, the thrust dynamics appear also roughly modeled, resulting in errors on forces/moments varying from 50% to 100% w.r.t. the theoretical ones. Hence, the internal model approximates the best reference model, which itself only approximates the real (unknown) aircraft.

Conclusion and prospects

This paper highlights the potential value of introducing surrogate models in some stages of the development of FCS for aircraft. The special types of neural networks, which have been selected for their attractive properties, were firstly described, as well as the constructive procedure implemented to optimize their internal parameters. Then, three specific applications related to modeling, identification and control aspects were considered and illustrated by processing real aircraft data and real aerodynamic models. These applications extend from the exploitation of the first flight tests of a new airplane to in-service monitoring of the flight parameters, through the intermediate steps required for designing the control laws. To complete this picture, it is worth noting that other stages of aircraft development are likely to benefit from surrogate-based approaches. For instance, at both ends of the process, the Multidisciplinary Design Optimization (MDO) during the preliminary design of a new airplane, and the clearance of the flight control laws involving worst-case analysis prior to aircraft certification, are two candidates for which surrogate models are commonly investigated. These applications, although slightly different, involve a costly multiobjective optimization process, and hence share the need to limit the number of high-fidelity computations by means of low-fidelity models, which can replace a significant proportion of the computationally demanding simulations.

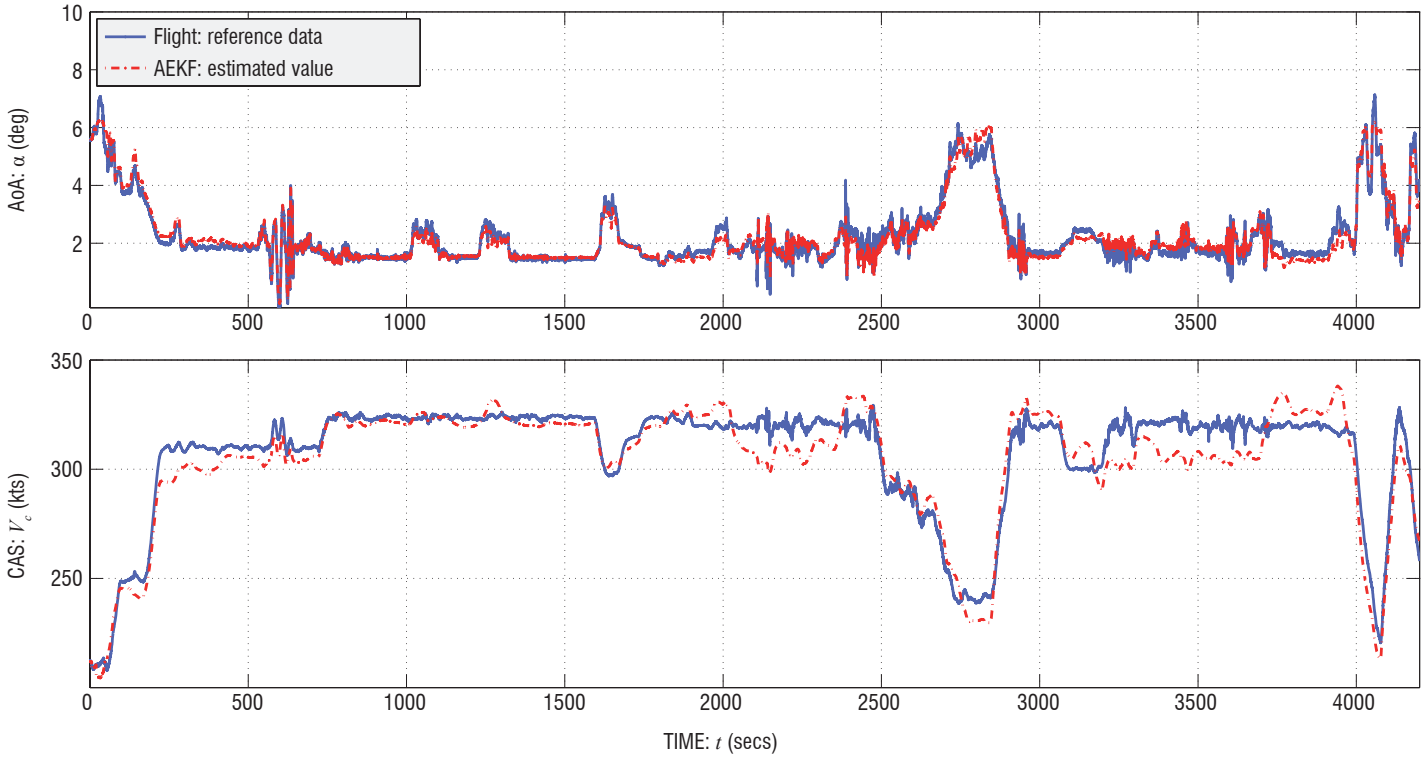


Figure 11 – AEKF results obtained by processing the flight data and mimicking a complete loss of both AoA and airspeed information at t = 20 sec

Concerning the first application topic (building efficient LFR for control analysis and design), room for improvement still exists and would be based on a better sparsity of the surrogate-based indirect approach by pruning useless explanatory variables in the final rational expressions. This would contribute to enlarging the practical user cases for which rational approximants are superior to polynomial ones and can provide smaller sized LFR. Though basically suited to off-line modeling applications, it should be noted that the *koala* tool can also be used for on-line implementation of LFT controllers. Actually, the usual way of computing scheduled LFT gains, for instance, is not relevant for embedded controllers because its computational burden is generally far too high. Rather than pre-computing tabulated values of the control gains, which can be a tricky option for high dimensional functions, very accurate and parsimonious models can be achieved thanks to the APRICOT library. The surrogate model will fully benefit the practical implementation since the (factorized) rational expressions are consistent with the stringent coding requirements for aircraft computers (see [15] for example).

The second application topic (identification of aerodynamic non linearities) illustrates that the transfer of new tools to industry is quite a long term process. The end users need to gain confidence in the results provided by new tools and also need time to appropriate these tools and their functionalities. In the long run, surrogate models might not only be intermediate models introduced into the identification process to facilitate the optimization stages, but may also represent a new way of representing aerodynamic nonlinearities in a wider part of the industrial process. This would avoid the present upstream and downstream conversions: from look-up tables to neural networks, and then backward reprojection of the updated results in the tables used *in fine* by the simulation environment. Regarding prospects, the development of multicriteria optimization is a very promising topic for the identification process. As mentioned in the section devoted to the second application, this process is still tedious today, involving a series of semi-manual steps which are time-consuming and would benefit from being more automated. Some of these steps are typically based on dealing with various criteria: tracking the time domain histories of the measured flight parameters by means of both EE and OE fitting criteria, considering also the values of modal frequencies and dampings in the frequency domain, and offering a compromise between the raw solution of the optimization problem and the discrepancies w.r.t. *a priori* knowledge encoded in the pre-flight data. Multiobjective optimization is a way of coping with these criteria of various types, and also of providing the end users not only with a unique "optimal" solution, but with a set of solutions (the so-called Pareto surface), some of which are likely to offer more satisfactory trade-offs for those conflicting goals. As the whole process should be achieved over large areas of the flight domain and should include all of the aerodynamic nonlinearities, surrogate modeling should be an integral part of this multicriteria identification, which is a challenging topic for future developments.

The third application topic (embeddable models for virtual sensing) is a good example of the aforementioned remark: the industrial process could benefit from a common way of representing aerodynamic nonlinearities by means of surrogate models. If this were the case, the updated models resulting from the identification stage could be directly used as internal models in the estimation scheme. Besides, it is worth noting that the same simplified aircraft surrogate modeling can be shared by the AEKF and by any model-based FDD techniques (as for instance the one used in [45]). In addition to being useful for civil aircraft, this virtual sensing strategy can also be beneficial for under-equipped vehicles (UAV, small airplanes). In all cases, the grey-box representation selected is essential for preserving only the most relevant aerodynamic terms in the embedded surrogate model, and for proceeding to a sensitivity analysis as regards the performances of the estimator. Another possibility offered by this modular and grey-box architecture, combined with the local properties of the RBF/LLM-type networks, would be to adjust the internal model of the estimator in a local and incremental way. Such an adjustment could be achieved in terms of the real performances under operational flight conditions, with standard flight profiles. Hence, the model would be finally fitted to the behavior of the estimator, resulting in a surrogate model that may be slightly different from a pure knowledge-based model. Finally, on-going works related to this topic are aimed at directly dealing with the anemometric measurements available on-board (pressures and temperature) instead of the CAS information, in order to improve the detection capability of the combined FDD and estimation schemes ■

References

[1] R. BABUSKA, H. VERBRUGGEN - *Neuro-Fuzzy Methods for Nonlinear System Identification*. Annual Reviews in Control, Vol. 27, p. 73-85, 2003.

[2] D. A. BELSLEY, E. KUH, R. E. WELSCH - *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & sons, New York, 1980.

[3] T. M. BLACKWELL, P. J. BENTLEY - *Dynamic Search With Charged Swarms*. GECCO'02, New York, USA, July 2002.

[4] C. BOUCHEREAU, D. LIOT - *Aerodynamic Specification for A340/A330/A321 Flight Training Simulators*. Airbus TR 456.006/90, Issue 8, April 1993.

[5] A. BUCHARLES *et al.* - *An Overview of Relevant Issues for Aircraft Model Identification*. AerospaceLab, Issue 4, http://www.aerospacelab-journal.org/al4, 2012.

[6] O. S. CELIS, A. CUYT, B. VERDONK - *Rational Approximation of Vertical Segments*. Numerical Algorithms, 45, p. 375-388, 2007.

[7] S. CHEN, S. A. BILLINGS, C. F. N. COWAN, P. M. GRANT - *Nonlinear System Identification Using Radial Basis Functions*. International Journal of Systems Science, 21, p. 2513-2539, 1990.

[8] S. CHEN, S. A. BILLINGS - *Neural Networks for Nonlinear Dynamic System Modelling and Identification*. International Journal of Control, 56, p. 319-346, 1992.

[9] S. CHEN, X. HONG, C. J. HARRIS, P. M. SHARKEY - *Sparse Modelling Using Orthogonal Forward Regression with PRESS Statistic and Regularization*. IEEE Transactions on Systems, Man and Cybernetics, Part B, 34 (2), p. 898-911, 2004.

[10] S. CHEN, X. HONG, B. L. LUK, C. J. HARRIS - *Nonlinear System Identification Using Particle Swarm Optimization Tuned Radial Basis Function Models*. International Journal of Bio-Inspired Computation, 1 (4), p. 246-258, 2009.

[11] J. CHO, J. LAN, G. K. THAMPI, J. C. PRINCIPE, M. A. MOTTER - *Identification of Aircraft Dynamics Using a SOM and Local Linear Models*. 45th Midwest Symposium on Circuits and Systems, Tulsa, USA, August 2002.

[12] M. CLERC - *Particle Swarm Optimization*, ISTE, London, 2006.

[13] M. CLERC - *Why Does it Work ?* International Journal of Computational Intelligence Research, 4 (2), p. 79-91, 2008.

[14] G. DREYFUS *et al.* - *Neural Networks*. Methodology and applications, Springer, 2005.

[15] G. FERRERES, G. HARDIER, C. SEREN - *Adaptive Control of a Civil Aircraft Through On-Line Parameter Estimation*. SysTol'16, Barcelona, Spain, September 2016.

[16] G. GOLUB, V. PEREYRA - *Separable Nonlinear Least Squares: the Variable Projection Method and its Applications*. Inverse Problems, 19 (2), p. R1-R26, April 2003.

[17] G. HARDIER - *Recurrent RBF Networks for Suspension System Modeling and Wear Diagnosis of a Damper.* IEEE World Congress on Computational Intelligence, Vol. 3, Anchorage, USA, May 1998.

[18] G. HARDIER, C. ROOS, C. SEREN - *Creating Sparse Rational Approximations for LFR Modeling Using Genetic Programming.* 3rd IFAC Intal Conf on Intelligent Control and Automation Science, Chengdu, China, 2013.

[19] G. HARDIER, C. SEREN, P. EZERZERE - *On-Line Estimation of Longitudinal Flight Parameters*. SAE AeroTech Congress and Exhibition, Toulouse, October 2011.

[20] G. HARDIER, C. SEREN, P. EZERZERE - *Model-Based Techniques for Virtual Sensing of Longitudinal Flight Parameters*. International Journal of Applied Mathematics and Computer Science, 25(1), March 2015.

[21] S. HAYKIN - *Neural Networks: a Comprehensive Foundation*. IEEE Press, MacMillan Ed., New York, 1994.

[22] S. HUTASSE, T. DUCHAMP, B. BARRIETY - *Model Specification AER A350*. Airbus TR V00EDD081441, issue 19.0, March 2016.

[23] R. V. JATEGAONKAR - *Flight Vehicle System Identification - A Time Domain Methodology.* AIAA Progress in Astronautics and Aeronautics, Frank K. Lu Ed., Vol. 216, Reston, 2006.

[24] J. LANE, A. P. ENGELBRECHT, J. GAIN - *Particle Swarm Optimization with Spatially Meaningful Neighbours*. IEEE Swarm Intelligence Symposium, St Louis, USA, September 2008.

[25] D. LIOT, A. BUCHARLES - *Outils d'identification de la mécanique du vol latérale des Airbus*. Symposium RTO System Concepts and Integration Panel, Madrid, 1998.

[26] C. F. LO, J. L. ZHAO, R. DELOACH - *Application of Neural Networks to Wind Tunnel Data Response Surface Methods*. 21st AIAA Aerodynamic Measurement Technology and Ground Testing Conference, Denver, USA, June 2000.

[27] J. F. MAGNI - *Linear Fractional Representation Toolbox for Use with Matlab*. Available at http://w3.onera.fr/smac/lfrt with the SMAC Toolbox, 2006.

[28] R. MENDES, J. KENNEDY - *The Fully Informed Particle Swarm: Simpler, maybe Better.* IEEE Transactions on Evolutionary Computation, 8 (3), June 2004.

[29] G. MONARI, G. DREYFUS - *Withdrawing an Example from the Training Set : an Analytic Estimation of its Effect on a Nonlinear Parametrised Model*. Neurocomputing, Vol. 35, p. 195-201, 2000.

[30] G. MONARI, G. DREYFUS - *Local Overfitting Control via Leverages*. Neural Computation, Vol. 14, p. 1481-1506, 2002.

[31] E. A. MORELLI, D. G. WARD - *Automated Simulation Updates Based on Flight Data*. AIAA AFM, Hilton Head, USA, August 2007.

[32] E. A. MORELLI, R. DeLoach - *Wind Tunnel Database Development Using Modern Experiment Design and Multivariate Orthogonal Functions*. 41st AIAA Aerospace Sciences Meeting and Exhibit, Reno, USA, 2003.

[33] E. A. MORELLI - *Global Nonlinear Aerodynamic Modeling Using Multivariate Orthogonal Functions*. Journal of Aircraft, 32(2), 1995.

[34] O. NELLES, R. ISERMANN - *Basis Function Networks for Interpolation of Local Linear Models.* 35th IEEE Conference on Decision and Control, 1, Kobe, Japan, December 1996.

[35] O. OLORUNDA, A. P. ENGELBRECHT - *Measuring Exploration/Exploitation in Particle Swarms Using Swarm Diversity*. IEEE Congress on Evolutionary Computation, Hong Kong, China, June 2008.

[36] M. OOSTEROM, R. BABUSKA - *Virtual Sensor for Fault Detection and Isolation in Flight Control Systems*. Fuzzy Modeling Approach. Proc. of the 39th IEEE Conference on Decision and Control, Sydney, December 2000.

[37] M. J. L. ORR - *Introduction to Radial Basis Function Networks*. Technical Report of the Center for Cognitive Science, Edimburgh University, Scotland, April 1996.

[38] M. J. L. ORR, J. HALLAM, A. MURRAY, T. LEONARD - *Combining Regression Trees and Radial Basis Function Networks*. International Journal of Neural Systems, 10 (6), December 2000.

[39] M. J. L. ORR - *Local Smoothing of Radial Basis Function Networks*. Intal Symp. on Artificial Neural Networks, Hsinchu, Taiwan, 1995.

[40] J. R. RAOL, R. V. JATEGAONKAR - *Artificial Neural Networks for Aerodynamic Modeling*. TR IB 111-94/41, DLR Braunschweig, Germany, October 1994.

[41] C. ROOS - *Optimization Based Clearance of Flight Control Laws*. In Varga-Hansson-Puyou, Generation of LFRs for a flexible aircraft model, §4. Lecture Notes in Control and Information Sciences, Springer-Verlag, 2011.

[42] C. ROOS - *Systems Modeling, Analysis and Control (SMAC) Toolbox: an Insight into the Robustness Analysis Library*. IEEE Multiconference on Systems and Control, Hyderabad, India, 2013.

[43] C. ROOS, G. HARDIER, C. DOLL - *A Comparison of Techniques to get Sparse Rational Approximations for Linear Fractional Representations.* 29th Congress of the International Council of the Aeronautical Sciences, St Petersburg, Russia, September 2014.

[44] C. ROOS, G. HARDIER, J. M. BIANNIC - *Polynomial and Rational Approximation with the APRICOT Library of the SMAC Toolbox*. IEEE Conference on Control Applications, IEEE Multi-conference on Systems and Control, Antibes, France, October 2014.

[45] S. SAMAR, D. GORINEVSKY, S. BOYD - *Embedded Estimation of Fault Parameters in an Unmanned Aerial Vehicle*. IEEE Conference on Control Applications, Munich, Germany, October 2006.

[46] P. A. SAMARA, G. N. FOUSKITAKIS, J. S. SAKELLARIOU, S. D. FASSOIS - *A Statistical Method for the Detection of Sensor Abrupt Faults in Aircraft Control Systems*. IEEE Transactions on Control Systems Technology, 16(4), p. 789-798, 2008.

[47] S. SEHER-WEISS - *Identification of Nonlinear Aerodynamic Derivatives Using Classical and Extended Local Model Networks.* AST 15(1):33-44, February 2011.

[48] C. SEREN, G. HARDIER - *Adaptive Extended Kalman Filtering for Virtual Sensing of Longitudinal Flight Parameters*. 2nd IEEE International Conference on Control and Fault Tolerant Systems, Nice, France, October 2013.

[49] J. SJÖBERG, H. HJALMARSSON, L. LJUNG - *Neural Networks in System Identification*. IFAC SYSID, Copenhagen, July 1994.

[50] P. TRAVERSE, I. LACAZE, J. SOUYRIS - *Airbus Fly-by-Wire: a Total Approach to Dependability*. 18th IFIP World Computer Congress, p. 191-212, Toulouse, 2004.

[51] K. TROJANOWSKI - *Multi-Swarm that Learns, Intelligent Information Systems*. Vol. XVI, p. 121-130, 2008.

[52] C. C. DE VISSER - *Global Nonlinear Model Identification with Multivariate Splines – Application to Aerodynamic Model Identification of the Cessna Citation II*. Ph.D. Thesis, TU Delft, 2011.

[53] X. X. WANG, S. CHEN, C. J. HARRIS - *Using the Correlation Criterion to Position and Shape RBF Units for Incremental Modelling*. International Journal of Automation and Computing, 3 (4), p. 392-403, 2006.

[54] D. WEDGE, D. INGRAM, D. McLEAN, C. MINGHAM, Z. BANDAR - *On Global-Local Artificial Neural Networks for Function Approximation*. IEEE Transactions on Neural Networks, 17 (4), p. 942-952, July 2006.

[55] E. DE WEERDT, Q. P. CHU, J. A. MULDER - *Neural Network Aerodynamic Model Identification for Aerospace Reconfiguration.* AIAA GNC, San Francisco, August 2005.

[56] K. ZHOU, J. C. DOYLE, K. GLOVER - *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, 1996.

Acronyms

A/C	(Aircraft)
AEKF	(Adaptive Extended Kalman Filtering)
AoA	(Angle of Attack)
CAS	(Calibrated AirSpeed)
CFD	(Computational Fluid Dynamics)
EE	(Equation Error)
FCS	(Flight Control System)
FDD	(Fault Detection and Diagnosis)
FE	(Filter Error)
GA	(Genetic Algorithm)
HF	(High-Frequency)
LFR	(Linear Fractional Representation)
LLM	(Local Linear Model)
LOO	(Leave-One-Out)
LP	(Linear-in-their-Parameters)
LS	(Least Squares)
MLP	(Multi-Layered Perceptron)
NN	(Neural Network)
OE	(Output Error)
OLS	(Orthogonal Least Squares)
OSMA	(Outil de Simulation des Mouvements d'un Avion)
PRESS	(Predicted REsidual Sum of Squares)
PSO	(Particle Swarm Optimization)
RBF	(Radial Basis Function network)
SMAC	(Systems Modeling, Analysis and Control toolbox)
SNLS	(Separable Nonlinear Least Squares)

Jean-Marc Biannic graduated from SUPAERO in 1992, and received his PhD degree in Robust Control Theory. He joined ONERA as a research scientist in 1997, and has been accredited as a PhD supervisor since 2010. He is the author or co-author of 20 journal papers, more than 50 conference papers, many book chapters, teaching documents, a tutorial book on multivariable control, and Matlab toolboxes. He has participated in several European projects and Garteur Groups (on PIO and nonlinear control). From 2012 to 2016, he led a research project involving 10 research scientists for the development of the SMAC toolbox for Systems Modeling, Analysis and Control.

Georges Hardier holds a PhD in Automatic Control from SUPAERO. After joining ONERA, he was in charge of the development of autopilots and control laws for a series of French warships for two decades. He now has over 30 years of experience in parametric estimation, modeling and identification techniques, applied to the aeronautics industry. Recently, he implemented on-line estimation methods for FDD/FTC during the European project RECONFIGURE. He is also interested in surrogate modeling and evolutionary techniques for optimization, for which he has developed a series of algorithms and tools (*e.g.*, the APRICOT library of the SMAC toolbox).

Clément Roos graduated from SUPAERO in 2004 and holds a PhD in Automatic Control. He joined ONERA as a research scientist in 2007. He often takes part in industrial projects with Airbus and Dassault, and was notably involved in the European projects GARTEUR-AG17 and COFCLUO. His research interests focus on aircraft modeling, robustness analysis and control laws validation, as well as nonlinear design based on robustified dynamic inversion schemes and anti-windup synthesis. He is the author or co-author of several papers, book chapters, teaching documents and Matlab toolboxes.

Cédric Seren is a research scientist at ONERA. He graduated in 2003 from SUPAERO and received his Ph.D. degree in 2007. During his Ph.D., he worked on A/C flight tests protocol optimization for flight dynamics identification, using new evolutionary algorithms. Since 2007, his activity has been essentially focused on both nonlinear modeling and estimation for aircraft fault detection, isolation and recovery, as well as on mathematical optimization. Recently, he has implemented adaptive estimation techniques robust to sensors faults during the EU project RECONFIGURE.

Laurent Verdier, graduated from ENSICA in 1999 before joining Airbus in 2000 in the Flight Dynamics Simulation department. He acts as a technical competence leader for free-air aerodynamic identification and is the focal point for R&T activities. For 16 years, he has been involved in the flight dynamics identification of several Airbus programs (A340-600, A318, A380, Single Aisle Sharklet, A350) from the design of the flight test maneuvers to the validation of the final simulation tool. His research interests focus on flight test protocol optimization, identification methods, and flight test measurement processing, and he has also co-supervised several PhD works.

B. Hérissé, G. Hervieux, K. Dahia, J.-M. Allard, J.-C. Sarrazin
(ONERA)

E-mail: bruno.herisse@onera.fr

DOI: 10.12762/2016.AL12-15

Component-Based Simulation for Real-Time Experiments of Advanced Aerospace Systems

This paper presents recent advances at ONERA in simulation engineering. Simulation is used for many purposes, from performance evaluation of algorithms to realistic emulation of the real world, including real components such as humans. To design general and scalable simulation software, a flexible architecture needs to be designed. A component-based architecture approach that allows multilevel capitalization and high collaborative sharing between systems experts and computing experts has been developed at ONERA. This architecture allow hybrid simulation involving real devices in the simulation loop to increase the Technology Readiness Level of novel algorithms by emulating real embedded environment. Human-In-the-Loop simulation is also an important tool to study the human interactions with a system. Some recent research projects are presented to illustrate our simulation platforms.

Introduction

Simulation is a more and more demanding tool

The main role of computer simulation is to simulate physical or non-physical phenomena that intervene in a system. This is achieved using analytical and numerical models but also black box functions stemming from an unknown subsystem, for example. Using these models, simulation platforms are often dedicated to facilitating design of systems and evaluation of their performance, their safety and their reliability (for controller design, mechanical design, etc.). For technical or cost reasons, using simulation for testing new systems involving complex algorithms, new technological components or testing human behavior in specific contexts can be the best way to achieve the project requirements according to its Technology Readiness Level (TRL). The scientific projects in which ONERA is involved have different TRL. Therefore, simulation platforms have to be sufficiently flexible to meet all research needs.

Simulation has evolved along with computers. From the 80s, object programming languages allowed the ability to work with advanced software components to be exploited. The first simulations were often monolithic, but tended to complex architectures that address systems and system of systems [1] [2] very well. Software design took a new direction, with business issues considered in the architecture design to meet new objectives: the capitalization and the maintainability of code. The component-based approach that is being developed at ONERA and detailed below illustrates this need to produce a reference code useful for business experts as well as for software experts.

With the increasing power of computers, simulation tools can provide results with more and more precision and always faster. However, scientist and engineer demand is also increasing with this respect. Indeed, the improvement of output precision implies more complex models that need more computation power on specific architectures. Therefore, in order to respond efficiently to the requirements, hardware and software architectures for simulations have to be heterogeneous and multiple. Thus, a key challenge for software engineers is to design simulations that manage multithreading, distributed architectures and hardware acceleration (using GPU or FPGA chips for instance). Another challenging problem consists in ensuring repeatability, reliability and portability of the simulation tool.

Simulation interacts with the real world

Increasingly, the use of simulation is extended to interact with the real world, including humans, using physical devices as interfaces (network communication devices, haptic interfaces, flight pilot simulator with emulated cockpit, etc.). Such simulation platforms are made to ensure a complete immersion of the target component being assessed in its environment. For example, the analysis of human behavior (stress, panic, etc.) in a cockpit can be evaluated correctly only if the simulation environment is sufficiently accurate and coherent in terms of motion and vibration especially. Moreover, training and education can use high fidelity simulation for a better experience (3D simulation with haptic devices in medical education for example).

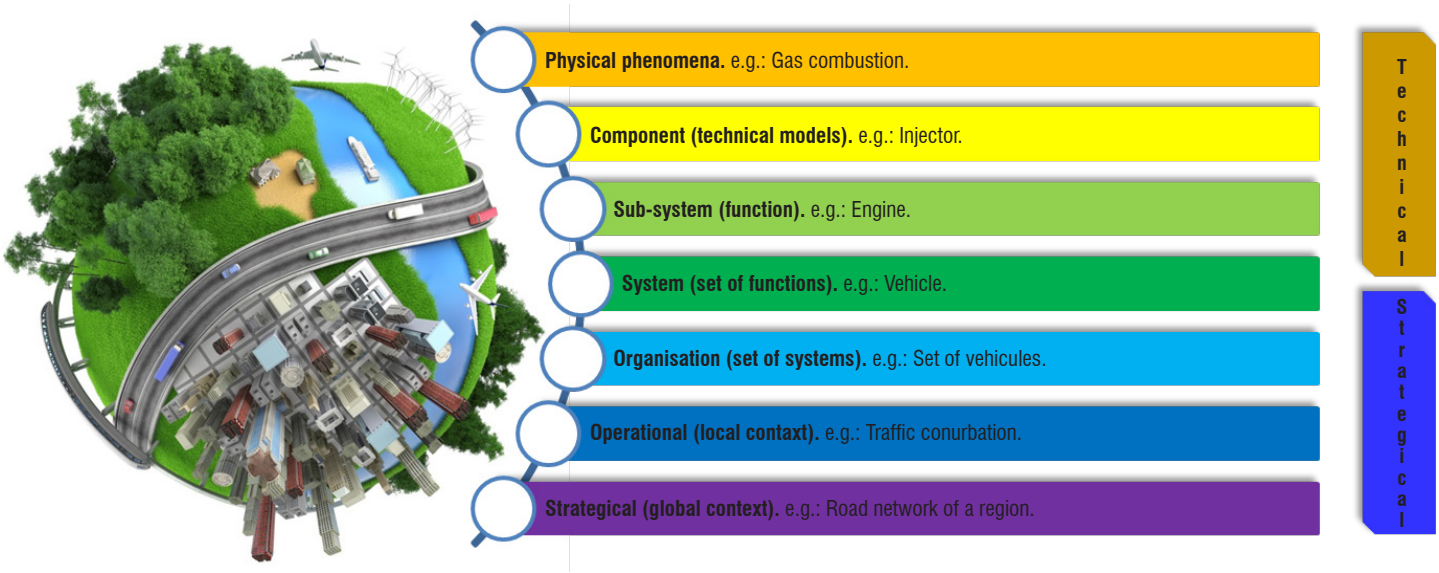


Figure 1 – Simulation levels of Systems

Thus, simulation of complex systems can include several specific simulators for each actor of the considered system, as well as real data stemming from a running system, such as a radar detection system or a decision platform. The simulation can also interact with humans via a flight simulator or a supervision platform. Moreover, hardware devices, such as embedded computers and sensors, can be incorporated into the simulation to assess their performance under realistic conditions before considering their use onboard a real system. In this paper, some of ONERA's simulation means are presented through some past and current research developments that involve real components in the simulation loop.

Designing relevant simulation

For a simulation platform to be fully reliable, the right behavior is fundamental to correctly emulate the real world. Therefore, evaluation tools are necessary. Taking advantage of the fact that all input and output signals are observable in a simulation test, such simulation tools can be designed. For example, in a Simulink simulation [3] or a LabVIEW simulation [4], all wires connecting component blocks can be checked in real-time with available viewers. Most real-time operating systems can be simulated to observe the full system behavior for every time step, using the associated tool such as WindView for VxWorks [5]. Profiling tools are also available for the recent Robotic-Operating-System [6] used for robotic applications. However, work remains to be done to interpret this flow of data in real-time and autonomously, so that supervision of the simulation and reconfigurations are possible. For example, communication based on an Ethernet network must be robust to packet loss using dead reckoning techniques. Such tools make it possible to verify that the simulation provides the right output related to the given input only empirically. To be fully reliable, formal proof should be provided in order to ensure a deterministic behavior [7].

Another important challenge raised by simulation is to make both the input data easy to enter in the simulation and the output result useful for the user. Ideally, the user should not have to understand how the tool is designed. This necessitates suitable autonomous interpreters that bridge the gap between the simulation and the user. Moreover, the simulator should be able to adapt its models to the required output precision by means of automatic model reduction, for example.

Simulation platforms at ONERA

This paper is aimed at presenting some of the simulation platforms available at ONERA through some recent or ongoing projects. It is organized as follows. In the next section, the architecture of the simulation Platform developed by engineers at ONERA is presented. Then, we focus on three topics of interest for our current and future advanced testing developments: interconnection of simulators in a distributed system, Hardware-In-the-Loop simulation and Human-In-the-Loop simulation.

Architecture of the simulation Platform

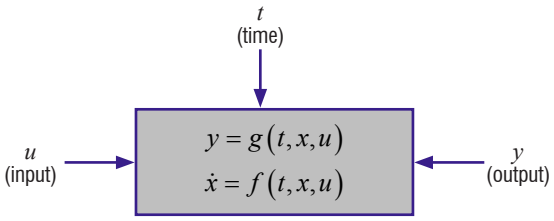
To efficiently design a software architecture understandable by both business experts and software engineers, it is necessary to develop a methodology based on a multilayer approach (Figure 1). Indeed, system studies sometimes require variable granularity models: very coarse to very fine simulation considering complex physical phenomena. The architecture implementation must be robust to this scaling problem. This implies the definition of the adequate scope for each component, as well as for the corresponding interfaces.

In the case of systems and systems of systems, this subdivision can be performed by functions or by physical architectures. Functional separation will be preferred for macroscopic levels and technical cutting up for low levels. To facilitate the work, it is helpful to use a pattern for arranging components in relation to each other. In our case, we used the SCA paradigm (Sensor / Controller / Actuator) well-known in the field of automatic control and especially suitable for addressing engineering systems.

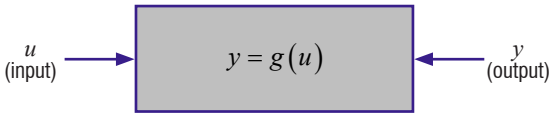
Component-based subdivision of the system

An actor in the simulation, also called agent in some cases, consists of components. A component is a block consisting of a model.

The purpose of this component consists in processing input data (the control u) to provide output data y. For technical components modeling the dynamics of a physical system, a simplified abstract representation is given below:



where g is the output function, x is the state and f is a continuous state function. For functional components, the representation does not involve time or internal state:



Definition of interfaces

For the communication between the components to be possible, it is necessary for their interfaces to be compatible. In our approach, an event mechanism is used. Once a component has finished a job, a new event is produced as a message and broadcasted directly to components that have subscribed to it. Each component is responsible for its subscription to the messages of interest. An initialization step is performed before the simulation is launched, to ensure that all of the communication channels are set up. During the simulation, a callback is run by a manager responsible for calling the broadcasting function when it is necessary.

Multi-level compatibility is ensured by a scale effect on each component. Indeed, it is possible to divide a component A into sub-components {A.1, A.2, ..., A.n}. The interface between a component A and a component B implies the interfacing between the components {A.1, A.2, ..., A.n} and B. This means that the inter-component interfaces should remain the same. However, the intra-component interfaces can be defined regardless of the outside and may be more specific or less generic.

Implementation of the simulated system

The architecture presented in Figure 2 is organized around a master (the *SimulationManager*), it is able to access all of the players (the *Entity*) in the simulation, as well as process them over time. These players consist of components that can be functional if the time is not a necessary input data for updating, or technical if a temporal integration mechanism is needed. Communication between components is ensured by a specific object (the *CommunicationManager*) attached to the master. It is responsible for transmitting messages between players. In order to manage time in the simulation, a specific component (the *Sequencer*) is dedicated to managing the execution sequence.

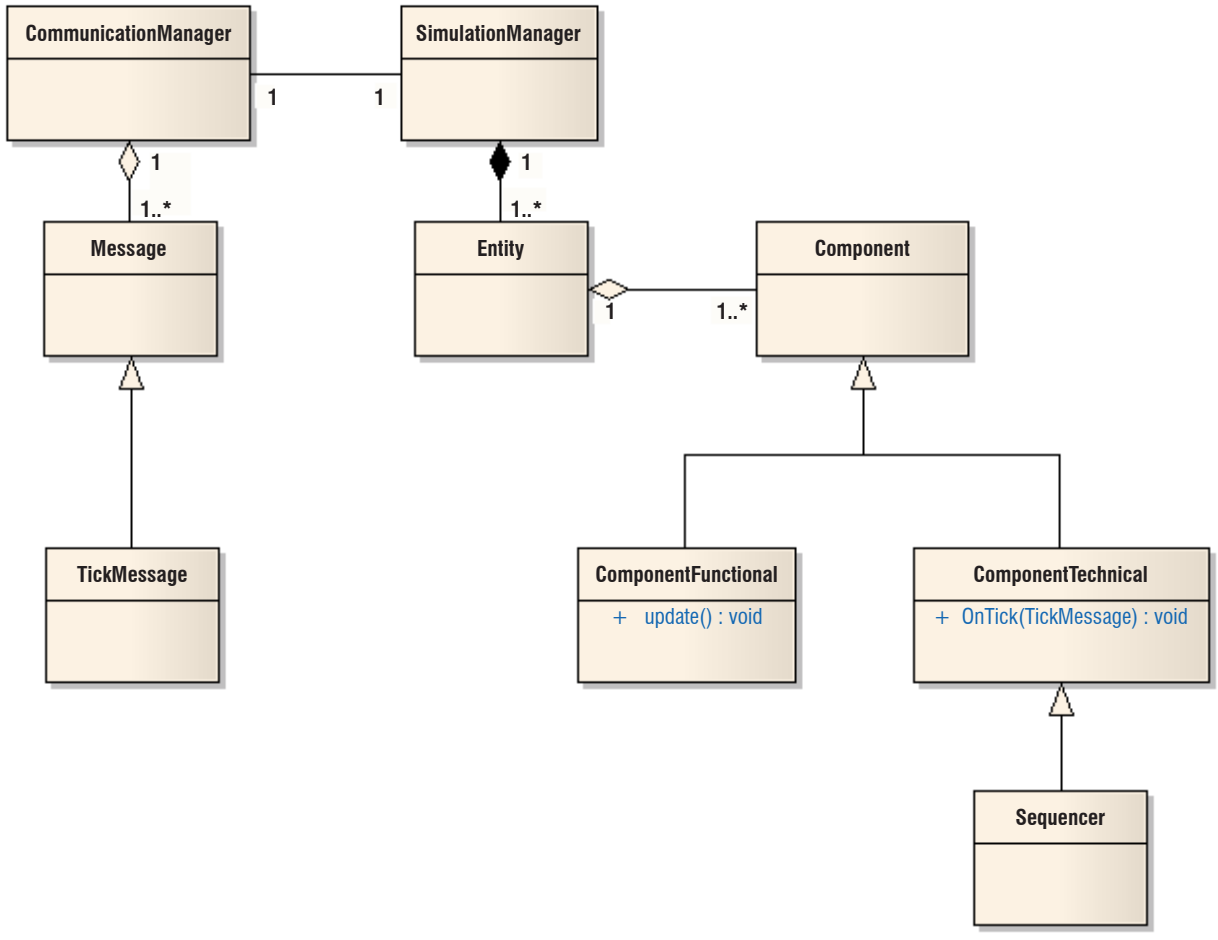


Figure 2 – Architecture of the simulation (class diagram)

Managing discrete and continuous events

The management of discrete and continuous events is done by sending messages. In the discrete case, a message can be any event sent punctually that will be processed upon receipt of the message modifying the internal state of the component. In the continuous case, a time message (*TickMessage* in Figure 2) is sent at each step of the simulation. This time message contains the current simulation time and the time step thereby achieving, for example, an integration of the component state. Messages are communicated via a callback when it is necessary. Event-messages are sent instantly. Time messages are sent at each step (*tick*) of the simulation. For the purposes of specific inter-component communication, messages are sent instantly, thereby making available up-to-date data to meet the needs of each component.

As previously said, the management of simulation is performed by a master managing the events and the progress of each player and component. For specific system needs, a new feature was implemented to allow each component to work at a different time step. Indeed, in most systems, sensors, controllers and actuators do not operate at the same speed. To be more representative of the physical reality, it is necessary to consider this time constraint in order to ensure that the system remains stable. For example, guidance and control modules of an aerospace vehicle do not operate with the same speed. Therefore, sub-tick management has been implemented in the *Sequencer*. Moreover, components need to be activated in a right order so that the loop progresses correctly over time. In our simulator, actuators, then sensors, then controllers are activated successively in the simulation loop.

Interconnected simulation

The interconnection of a simulator with an external tool is quite common. Often, some processing and even some simulators are third party tools that cannot be integrated either with the used technologies or due to the intellectual property. The interconnection is very often based on a standard of communication so that it can be facilitated. The standard must be implemented on both sides and can sometimes be expensive, especially when a specific implementation meets a specific need. Two cases are considered here: a first case addressing interoperability between simulations through the use of standard High-Level-Architecture and a second case dealing with a specific interoperability with an external tool (which is not a simulator) processing data during the simulation.

Distributed simulation

Distributed simulation offers the possibility of playing a simulation on several remote machines. In order to implement this type of simulation, the High-Level-Architecture (HLA) [8] [9] [10] standard has been used at ONERA for twenty years [11]. The previously described simulator provides a HLA gateway to map messages to the HLA standard. This standard describes the objects and shareable messages in a reference RPR-FOM [12]. The master of the simulation is a Run-Time-Infrastructure (RTI) that manages time. Messages are timestamped and a dead-reckoning mechanism allows the Federated actors to have the right information at time t knowing the information at time $(t - dt)$.

This kind of distributed simulation has been implemented many times at ONERA. In particular, tests in collaboration with the DGA (the French department of defense) in the simulation network SimDEx Defense (see Figure 3). In this experiment, some actors (Federated) were played at ONERA and the others at the DGA, thereby demonstrating the feasibility of large-scale simulations involving many actors (industry, state, etc.).

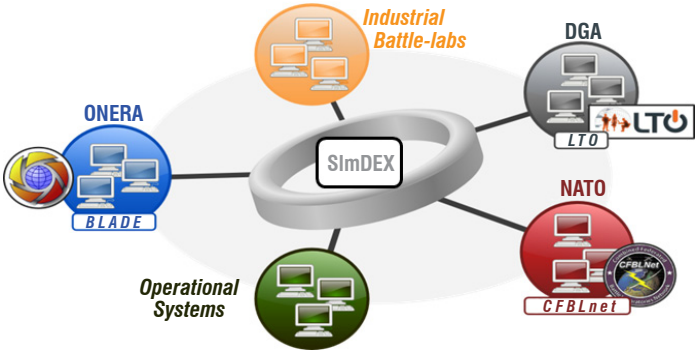


Figure 3 – HLA distributed simulation in the SimDEx network

In this example, interconnection consisted in providing the simulator with an external tool that is responsible for carrying out the fusion of data from the simulation. This tool is also involved in the simulation chain, since it may send requests for additional information via an operator. Therefore, this simulation also deals with the Human-In-the-Loop problem with this contribution of a human during the simulation. From the point of view of computer implementation, a specific gateway has been developed for converting the simulation events into SOAP messages. These XML SOAP messages are specific to our implementation. The data fusion tool only reacts to data from the simulation by processing them according to the current time of receipt. These messages are timestamped and sent by the simulator. A buffer mechanism was provided at the gateway for more flexibility in the transmission of data over the network. A synthetic scheme is proposed below in Figure 4. Note the specified processing loops representing a process. It was necessary to create a thread for receiving the simulation data so that the simulation loop is not blocked. The exchanges are ensured by two SOAP connections: sending simulator data (uplink) is done by the server to the client from the external tool; receiving data from the external tool simulator is done by the client simulator (downlink).

This work has contributed to highlighting the interconnection capacity of the simulator with an external tool. The effort involved is more important when using a standard communication protocol is not possible. Preference is given to the use of gateways that can be capitalized without creating a strong connection with the simulator. Indeed, in our case, the gateway converts the simulation messages to SOAP messages without any useful gains (only a translation is carried out).

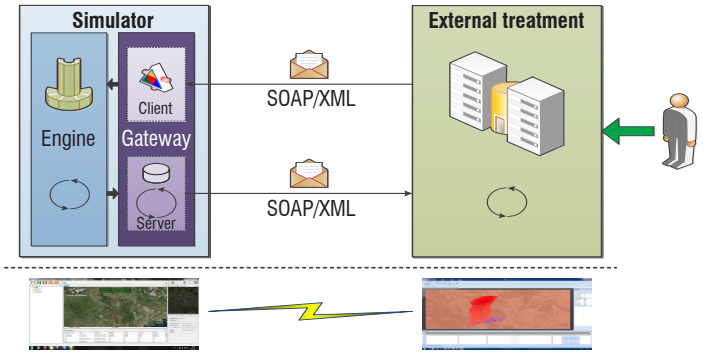


Figure 4 – Example of a specific interconnection

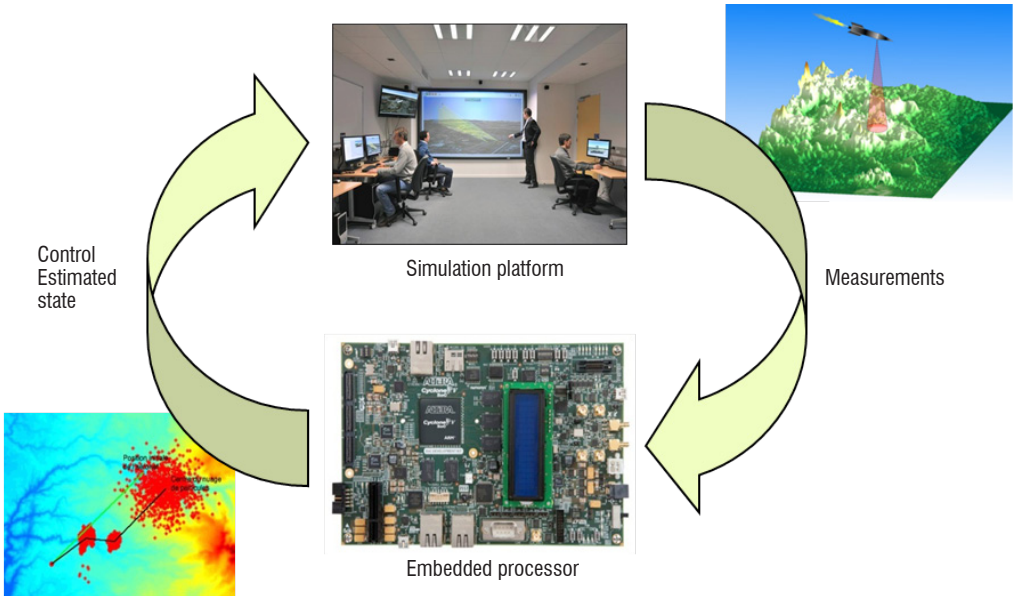


Figure 5 – Presentation of a HIL Simulation

Hardware components in the simulation loop

The purpose of the Hardware-In-the-Loop (HIL) simulation is to evaluate a real sub-component of an embedded system inside a simulation of the other components of the system. This method is widely used for system validation and verification. This is particularly interesting when the system is very complex and includes many hard components (sensors, actuators, controllers, interfaces, etc.). Testing each component with real experiments could be very expensive, unacceptable or even impossible. For example, the control of a vehicle used under critical conditions requires many tests to ensure its robustness to every possible scenario, the HIL simulation then offers a cheap, safe and repeatable method for this purpose.

A HIL simulation is a real-time loop including three main components. The embedded hardware is the component that needs to be evaluated. This often consists of a controller alone. However, real sensors and actuators can also be plugged into this controller. In that last case, another hardware component is necessary to emulate the plant model. For example, if a vehicle motor and its controller are tested in a HIL simulation, the motor load needs to be emulated by another physical component. If a vision-based algorithm is under test including the camera, images of the modeled environment need to be generated [13]. The HIL simulation also includes a real-time computer implementing the plant model. Eventually, an I/O device is used for the communication. Indeed, for a complete evaluation of the embedded component, the I/O interface is needed to communicate with the simulation using the real physical signals that are either feeding an actuator or generated by a sensor. Such a testing approach is often used in the design chain of a system controller. The HIL simulation constitutes the last test before validation with the real physical system after the other validation steps have been performed, that is, Model-In-the-Loop (MIL), Software-In-the-Loop (SIL) and Processor-In-the-Loop (PIL) simulations. Only the HIL and the PIL simulation can be used to verify that the execution time in the embedded processor fits within the required time. For PIL simulation, the interface between the embedded controller and the plant model is ensured by a standard communication such as Ethernet, while for the HIL simulation, the plant model is implemented on a hard real time computer, so that full real time simulation can be performed.

When a full complex system, such as the aerospace systems encountered at ONERA (missile systems for example), cannot be physically built and tested, PIL and HIL simulations offer a way for experimenting new algorithms in real time and under real conditions. Using this approach, the next section presents the experimentation of an advanced navigation and guidance algorithm currently being done in the DCPS department at ONERA.

Experiments of embedded advanced GNC algorithms in a HIL simulation

Experiments on advanced navigation and guidance algorithms are currently conducted in the DCPS department using the presented approach to demonstrate their performance in a realistic complex scenario.

The problem considered consists of both an interceptor missile and a cruise missile. The goal of the interceptor is to intercept the possibly maneuvering target (Figure 6). A ground station involving a radar detects the target and sends a predicted intercept point [14] to the

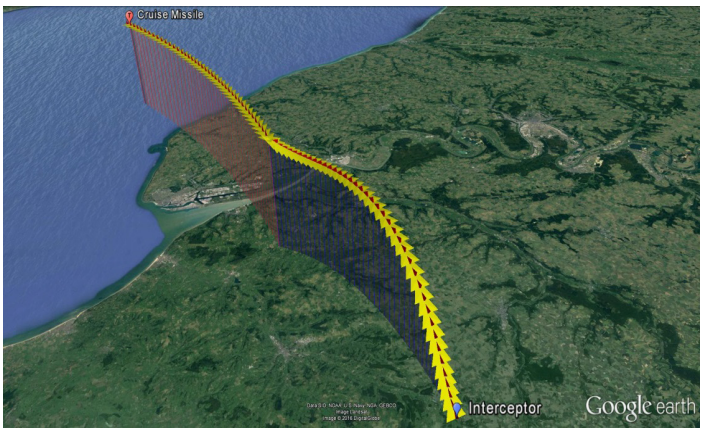
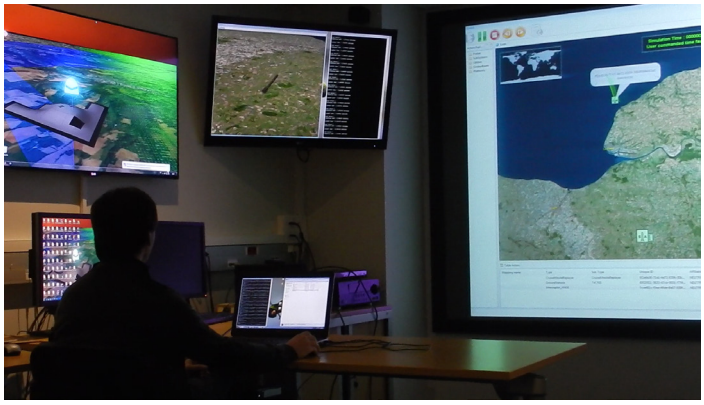


Figure 6 – Presentation of the intercept scenario

interceptor so that it can ensure the rendezvous. For this scenario, a controller is embedded in the interceptor for the midcourse guidance [15] and a second controller is embedded in the cruise missile for navigation [16]. A short video presenting this project is available at the following URL: http://www.aerospacelab-journal.org/sites/aerospacelab.onecert.fr/files/playlists/al12-15-video1.flv



Video 1 – Hardware-In-the-Loop simulation for missile interception

The computing consumption of classical algorithms used in most aerospace systems is low. Indeed, these algorithms often rely on analytical equations that can be computed very fast. Thus, cheap and low-power embedded processors are sufficient to ensure real time performance. Such classical algorithms are not satisfactory for complex missions involving navigation without GPS or optimal guidance, as considered here. Thus, advanced algorithms to address such problems have been developed. For the navigation of the cruise missile, specific particle filters were designed [16] [17] and for midcourse guidance of the interceptor, an indirect shooting method was designed [15]. These two algorithms are computationally demanding, since the first relies on the Monte-Carlo method and the second relies on iterative Newton methods [18]. Good performance is demonstrated in simulation results. However, it remains to be shown that these techniques can be embedded with the same level of performance, that is, real time capabilities need to be verified on an embedded processor.

To this end, a PIL simulation has been carried out: Figure 7 briefly presents the architecture of the simulation. A computer ensures the simulation of the system consisting of three main components: the interceptor component, the cruise missile component and the ground station component. The interceptor includes a subcomponent that ensures the computation of the control and the target component includes a subcomponent for navigation. Two modes are available: a full MIL simulated mode (control and navigation are modeled) and a PIL mode (control and navigation are ensured by embedded processors). For the PIL mode, the subcomponents directly transmit input data to the embedded processor through an Ethernet interface and the output data are acquired using the same interface. The processors that are used here are two Cyclone V SoC Development Kits from Altera that allow computing acceleration using the integrated FPGA [19]. Since the interface between the simulation and the embedded processors does not emulate real signals yet, this simulation is not fully HIL. However, as a first step, the PIL simulation is sufficient to verify that the processors can compute output data within the required time.

The proposed simulation loop offers a tool that allows the team to experiment new algorithms in realistic situations, so that the practical feasibility of our algorithms can be demonstrated. However, critical aerospace applications demand a safe and deterministic behavior. Therefore, to increase the TRL of our novel algorithms, a formal verification of the developed software may be required [20]. For instance, some advanced algorithms developed at ONERA rely on optimization methods, such as the Newton method or the interior point method [21]. A formal analysis of such methods needs to verify that a given precision can be achieved for a given number of iterations. These deterministic properties are necessary to ensure that the implemented algorithms can be qualified and transferred to industry. Future work will concentrate on these important issues.

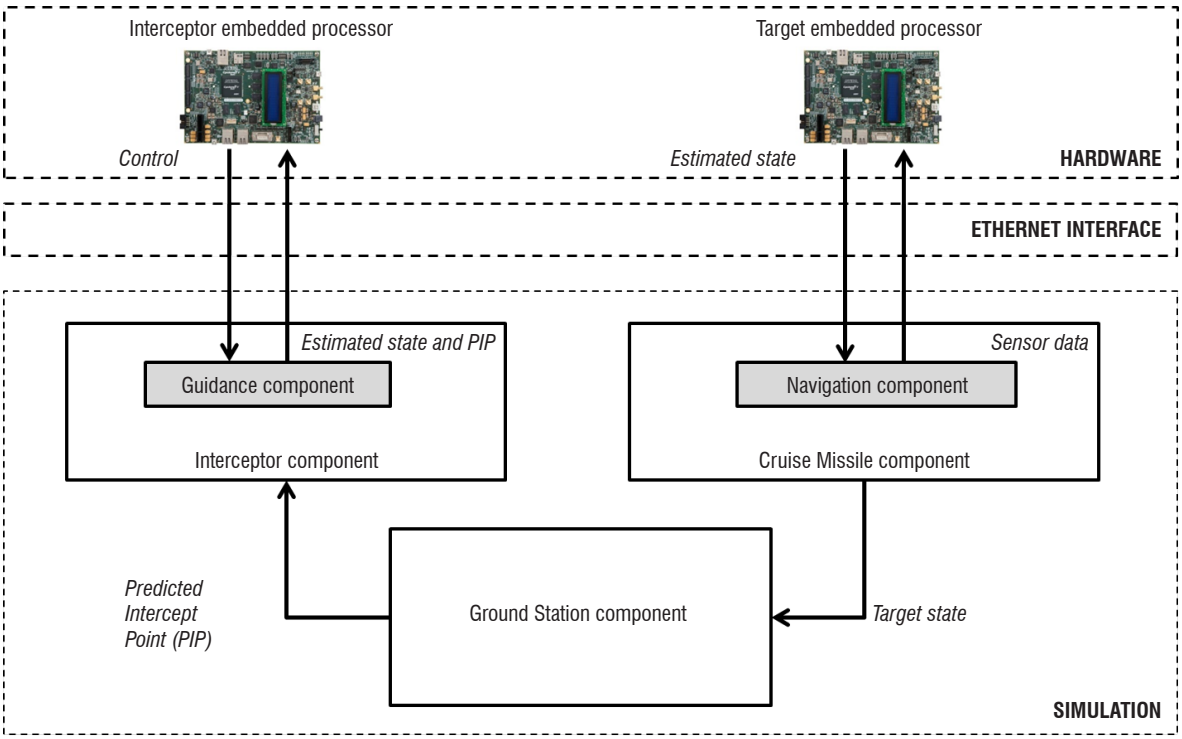


Figure 7 – Architecture of the PIL simulation

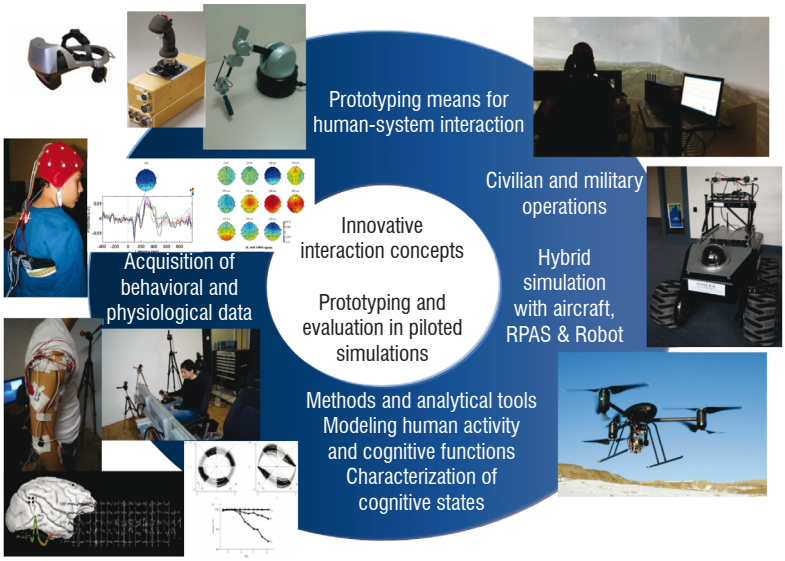


Figure 8 – Some experimental means available at ONERA

Human-in-the-loop simulation

When a Human interacts with the simulation, the simulation is called Human-In-the-Loop. Such simulations can be used for training, for example (flight simulators for pilot training), or to assist in the design of virtual objects by creating a simulated environment. Human-In-the-Loop simulation has been a focus of many research projects [22] [23] [24]. These projects generally consider how automation and humans can work together or in a joint environment. However, testing these algorithms and setting up experiments is often difficult, mainly due to lack of realistic simulations. Quantifying simulation fidelity, using an engineering metrics approach, underpins the confidence in the successful completion of the conception-design-build-test/qualification-production-operation cycle of aircraft, yet has been neglected in the aeronautical world. For fixed wing aircraft, the concept of zero flight time training using flight simulation is accepted and deemed necessary from a safety and cost standpoint. This must become the modus operandi for rotorcraft training. Simulators are commonly used to assess handling qualities and to develop crew-station technologies. Attempts to quantify overall simulation fidelity within the framework of handling quality engineering have been presented in a number of forms in recent years. In [25] [26] [27], an approach based on pilot-aircraft modeling has been developed and the handling quality sensitivity function was introduced as the basis of a quality metric. In [28] and later in [29] the use of the handling quality standard (ADS-33E PRF 22) was proposed, for deriving metrics, the rationale here being that if the simulator is to be used to optimize handling qualities, then what better parameters to judge fidelity than those defining the predicted handling. In [30] and [31], an approach using comparative measures of performance and control activity, correlated with handling quality ratings given for the same tasks when flown in simulation and in flight, was presented. In all of these approaches, the philosophy has been to try to develop a rational and systematic approach to the identification of the differences between simulation and flight, hence directing attention towards areas of deficiency. The partial success of these methods is encouraging, but only serves to highlight the need for fidelity criteria for use in design, development and product qualification. In these areas, flight simulation can be a primary source of data from which knowledge is derived, decisions are made and significant resources are committed; similar arguments can be made for the development of flight training.

A Human-In-the-Loop project at ONERA

An assessment methodology for human-in-the-loop simulation is under development at ONERA, using available experimental platforms (Figure 8) that addresses the Spatial Disorientation (SD) phenomenon, where pilots experience erroneous sensations about their orientation. One situation in which SD can occur is the go-around procedure. In some circumstances, go-around can lead to a somatogravic (i.e., vestibular) illusion (also termed false-climb) prevalent during high accelerations (or decelerations) when a pilot has no clear visual reference [32]. The "illusion" is a strong pitching sensation up (or down) when the body is exposed to high accelerations (or decelerations) [33]. This illusion is due to the limitation of the vestibular system, which detects changes in orientation without differentiating between head tilt backward and forward acceleration.

How to optimally simulate self-motion using motion simulators is still an unsolved problem, despite the fact that self-motion simulation is an essential part of all commercial flight and driving simulators. Flight simulators used for pilot training and also most driving simulators strive to simulate motion trajectories that are considerably larger than the actual range of the physical simulator device. To do this, motion cueing algorithms attempt to mimic the accelerations that act on the body during self-motion. While a larger range of movement allows for more accurate motion cueing, increasing the number of degrees of freedom and enlarging the movement range of the simulator raises the costs of the device considerably, and there are also technical limits to what kind of trajectories can be performed in a simulator due to the limited motion envelope and actuator power. It is therefore important to find techniques to believably simulate large trajectories using smaller movements that are within the limited movement range of the simulator. Presenting believable physical accelerations is an issue for psychophysical experiments that investigate the perception of self-motion in motion simulators.

Conclusion and perspectives

In this paper, the state of the ONERA's simulation developments was presented. A component-based approach has been adopted to enable researchers and engineers to integrate their system developments as easily as possible into a common and stable software. For

researchers, this offers the possibility of increasing the maturity of their innovative solutions using a very flexible simulation tool. Moreover, thanks to the modularity of the simulator, the accomplishment of multidisciplinary projects is facilitated for system engineers. In particular, the simulation platform includes hybrid simulation involving humans and real components. This allows the gap between pure simulation and real experiments to be bridged.

Future work will consist in feeding innovative technologies into the hybrid simulation platforms at Onera, so that complex future systems can be designed and tested efficiently. Thus, the simulation platform will constitute a wonderful tool to demonstrate proof of concepts for future systems and to facilitate transfer of technology. For example, using all Onera technologies integrated into a simulation platform, future complex systems such as Reusable Launch Vehicles could be fully simulated and validated, including aerodynamic phenomena, navigation, guidance and control, specific sensors and actuators, etc.

Moreover, scientifically challenging problems still need to be explored, especially with regard to the management of simulation uncertainties: how can errors and inconsistencies be detected, and how can they be managed autonomously? Furthermore, the question of choosing the right granularity of models with respect to the assessed output remains open: both the real-time property and the output consistency need to be ensured at the same time. From a technical point of view, the simulation engine has to be improved to deal with studies that require very high precisions. The current engine is compliant with technical-operational studies, but for other studies improvements must be made in the numerical analysis domain. The Discrete Event System Specification (DEVS) is a formalism especially designed for the modeling and analysis of discrete event systems, as well as continuous state systems [34]. The idea is to manage the precision required by the experts in a transparent way. The engine will perform a fully integrated control of the time and of the events, in order to achieve the goal defined by the experts at the beginning of the experiment. ■

Acronyms

TRL	(Technology Readiness Level)
HLA	(High Level Architecture)
SD	(Spatial Disorientation)
HIL	(Hardware-In-the-Loop)
PIL	(Processor-In-the-Loop)
SIL	(Software-In-the-Loop)
MIL	(Model-In-the-Loop)
DCPS	(System Design and Performance Evaluation Department)

References

[1] P. CARLE *et al.* - *Simulation of Systems of Systems*. AersopaceLab Journal, 2012.

[2] R. CUISINIER, M. BRUNEL, and S. PRUDHOMME - *Using Open Source to Build Comprehensive Battlespace Simulations*. Proc. of SimTecT, 2010.

[3] https://fr.mathworks.com/products/simulink/.

[4] http://www.ni.com/labview/.

[5] http://www.windriver.com/products/vxworks/.

[6] http://www.ros.org/.

[7] E. CLARKE, D. KROENING, and F. LERDA - *A Tool for Checking ANSI-C Programs*. International Conference on Tools and Algorithms for the Construction and Analysis of Systems, March 2004, 168-176.

[8] J.S. DAHMANN - *The High Level Architecture and Beyond: Technology Challenges*. Proceedings of the Thirteenth Workshop on Parallel and Distributed Simulation, 1999, 64-70.

[9] IEEE 1516-2010. Standard for Modeling and Simulation High Level Architecture - Framework and Rules.

[10] IEEE 1516.1-2010. Standard for Modeling and Simulation High Level Architecture - Federate Interface Specification.

[11] J. BOURRELY, P. CARLE, M. BARAT, and F. LÉVY - *Genesis: an Integrated Platform for Designing and Developing HLA applications.* Proc. of Simulation Interoperability Workshop, 2005.

[12] IEEE 1516.2-2010. Standard for Modeling and Simulation High Level Architecture - Object Model Template (OMT) Specification.

[13] N.R. GANS, W.E. DIXON, R. LIND, and A. KURDILA - *A Hardware in the Loop Simulation Platform for Vision-Based Control of Unmanned Air Vehicles*, Mechatronics. vol. 19, 2009, 1043–1056.

[14] P. PHARPATARA, R. PEPY, B. HÉRISSÉ, and Y. BESTAOUI - *Missile Trajectory Shaping Using Sampling-based Path Planning*. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2013.

[15] R. BONALLI, B. HÉRISSÉ, and E. TRÉLAT - *Analytical Initialization of a Continuation-Based Indirect Method for Optimal Control of Endo-Atmospheric Launch Vehicle Systems*. IFAC World Congress, vol. to appear, 2017.

[16] A. MURANGIRA, C. MUSSO, K. DAHIA, and J.-M. ALLARD - R*obust Regularized Particle Filter for Terrain Navigation*. IEEE International Conference on Information Fusion (FUSION), 2011.

[17] N. MERLINGE, K. DAHIA, and H. PIET-LAHANIER - *A Box Regularized Particle Filter for Terrain Navigation with Highly Non-Linear Measurements*. IFAC Symposium on Automatic Control in Aerospace, 2016.

[18] E. TRÉLAT - *Optimal Control and Applications to Aerospace: Some Results and Challenges*. Journal of Optimization Theory and Applications, vol. 154, 2012, 713–758.

[19] https://www.altera.com/products/boards_and_kits/dev-kits/altera/kit-cyclone-v-soc.html.
[20] V. WIELS *et al.* - *Formal Verification of Critical Aerospace Software.* (4), p-1., AerospaceLab, 2012.
[21] J.F. BONNANS, J.C. GILBERT, C. LEMARÉCHAL, and C.A. SAGASTIZÁBAL - *Numerical Optimization: Theoretical and Practical Aspects.* Springer Science & Business Media, Ed., 2013.
[22] G.D. PADFIELD *et al.* - *Simulation Fidelity of Real-Time Helicopter Simulation Models*. 61st Annual Forum of the American Helicopter Society, 2005.
[23] S.J. HODGE, J.S. FORREST, G.D. PADFIELD, and M.D. WHITE - *Determining Fidelity Standards for Maritime Rotorcraft Simulation.* Maritime Operations of Rotorcraft. London: The Royal Aeronautical Society, 2008.
[24] V. SHIA *et al.* - *Semiautonomous Vehicular Control Using Driver Modeling*. IEEE Transactions on Intelligent Transportation Systems, 2014, 1-14.
[25] R.A. HESS - *Identification of Pilot-Vehicle Dynamics from Simulation and Flight Test*. Advances in Aerospace Systems Dynamics, vol. 31, 1990, 151-175.
[26] R.A. HESS and T. MALSBURY - *A Methodology for the Assessment of Manned Flight Simulator Fidelity*. Journal of Guidance, Control and Dynamics, vol. 14, 1991, 191-197.
[27] R.A. HESS and W. SIWAKOSIT - *Assessment of Flight Simulator Fidelity in Multiaxis Tasks Including Visual Cue Quality*. Journal of Aircraft, vol. 38, no. 4, 2001, 607-614.
[28] G.D. PADFIELD, M.T. CHARLTON, and A.T. MCCALLUM - *The Fidelity of Hi-Fi Lynx on the DERA Advanced Flight Simulator Using ADS-33 Handling Qualities Metrics*. DRA/AS/FDS/TR96103/1, 1996, 1-152.
[29] A.T. MCCALLUM and M.T. CHARLTON - *Structured Approach to Helicopter Simulator Acceptance,The Challenge of Realistic Rotorcraft Simulation*. RAeS conference, 2001.
[30] S.K. ADVANI and C.H. WILKINSON - *Dynamic Interface Modelling and Simulation-a Unique Challenge*. Society Conference on Helicopter Flight Simulation, 2001.
[31] M.F. ROSCOE and J.H. THOMPSON - *JSHIP's Dynamic Interface Modeling and Simulationsystem: a Simulation of the UH-60A Helicopter/LHA shipboard Environment Task*. 59th Annual Forum of the American Helicopter Society, 2003.
[32] T. WILSON - *Aircraft Human Performance & Limitations*. Civil Aviation Safety Authority, 1995.
[33] A.T KERN - *Flight Discipline*. McGraw Hill Professional, 1998.
[34] B.P. ZEIGLER, H. PRAEHOFER, and T.G. KIM - *Theory of Modelling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems*. Academic Press, Ed. London, 2000.

AUTHORS

Bruno Hérissé received the Engineering degree and the Master degree in 2007 from the *École Supérieure d'Électricité* (Supélec). He obtained the Ph.D. degree from the University of Nice Sophia Antipolis in 2010. Since 2011, he has been a research engineer at ONERA, Palaiseau, France. His research interests include optimal control and vision-based control of aerial vehicles.

Gyslain Hervieux has been a research engineer in computer science at ONERA since 2006. He received his Engineering degree from "*Ecole Nationale Supérieure d'Arts et Métiers*" (ENSAM 2002) and his Advanced Master in Simulation and Virtual Reality from "Institut Image of Chalon sur Saône" (2006). His research interest is focused on computer simulation and on computer graphics.

Karim Dahia is a senior navigation engineer with the ONERA French Aerospace Lab in Palaiseau, France. His Ph.D. from the *Université Joseph Fourier* – Grenoble (France) focused on the application of particle filtering to aircraft motion estimation. Dr. Dahia's research interests include robust and optimal navigation as well as filtering for aerospace systems.

Jean-Michel Allard received his Engineering degree from the "*Ecole Nationale Supérieure d'Arts et Métiers*" (ENSAM 1996) and the "*Ecole Supérieure des Techniques Aérospatiales*" (ESTA 1997). He has been working at ONERA since 2001 as a research engineer for the System Design and Performance Evaluation Department (DCPS). His research interest is focused on inertial measurements hybridization and simulation for navigation of aircraft. He is carrying out expertises to the benefit of military programs managed by the Direction Générale de l'Armement (DGA) of the Ministry of Defence.

Jean-Christophe Sarrazin is currently Senior Research Scientist at ONERA. He received his PhD in 2003 in Health and Life Sciences from the University of the Mediterranean (Marseille, France). After two postdocs in computational neuroscience (postdoc INRIA, Nancy, France, one year) and cognitive neuroscience (postdoc Marie Curie, Université Libre de Bruxelles, two years) respectively, he obtained a permanent position at ONERA. In the Information Processing and System Branch, in which he coordinates the Human System Integration Team, he is a neuroergonomist and works on the scientific and technical development of human system integration studies. He works on the identification and the modelling of the computational principles of motor control and its modulation by high level cognitive functions. As he considers that a key to the development of HMI technologies lies in the integration of the neurosciences by industrials, he oeuvres at ONERA for the acquisition of neuroscientific knowledge and methods of investigation, specifying simulation scenarios, and the use of this newly found knowledge in improving design methods of new concepts of interaction.