Reduced sampling and incomplete sensitivity for low-complexity robust parametric optimization

Bijan Mohammadi

REDUCED SAMPLING AND INCOMPLETE SENSITIVITY FOR LOW-COMPLEXITY ROBUST PARAMETRIC OPTIMIZATION

BIJAN MOHAMMADI
UNIVERSITÉ MONTPELLIER II, MATHÉMATIQUES
CC51, 34095 MONTPELLIER, FRANCE

INT. J. NUM. METH. FLUIDS. 2013.

Abstract. The paper considers robust parametric optimization problems using multi-point formulations and makes the link with momentum based formulations. Optimal sampling issues are discussed and a procedure is proposed to quantify the confidence level on the robustness of the design. We also discuss incomplete sensitivity evaluations to take into account the computational complexity constraint. This permits to take advantage of what was previously developed for efficient mono-point design where the cost of the optimization is comparable to one state evaluations. The proposed algorithm is fully parallel and the time-to-solution is comparable to mono-point situations. Concepts are introduced through simple examples and the paper ends with the design of the shape of an aircraft robust over a range of transverse winds.

1. INTRODUCTION

The performance of a system designed for given functioning conditions often seriously degrades when these conditions are modified. Typical situations of interest involve a few (typically one or two) parameters describing the functioning of the system. Momentum based and multi-point optimizations are two widely used approaches to address robustness issues in engineering because of their conceptual simplicity. Both require several evaluations of the functional and its gradient if a descent method is considered for minimization.

Our domain of interest is aerodynamic shape optimization and we illustrate our purpose on analytical and model problems before targeting such. The question of interest is: can we propose an aircraft shape designed to have similar performances over a given range of some functioning parameters (e.g. cruise Mach numbers, sideslip angle, etc), and can we do that modifying as less as possible an existing mono-point optimization shape design loop ? and also, is it possible for the time-to-solution cost of this parametric shape design to remain comparable to the mono-point situation ?

These questions are important as today industrial robust design mainly relies on reduced order modelling and intelligent sampling [2, 3, 4, 6] which either does not use high-fidelity simulations during design or uses lower accuracy than what would be affordable in a mono-point optimization. Our aim is to propose a plausible alternative.

Key words and phrases. low complexity, robust optimization, reliability, uncertainty, incomplete sensitivity, sampling size.

One issue in the paper is how to formulate the robustness when only a few functioning parameters are involved. We see how to introduce in this formulation what we would like for the outcome of the design through a target-based weighting in the functional. We also discuss how to take advantage of the incomplete sensitivity concept to perform robust optimizations without the linearization of the state equation [1]. This cost reduction is important despite our algorithm is naturally parallel.

Because we are interested by optimization over intervals of functioning parameters optimal sampling issues are also of importance and will be discussed together with how to provide some quantitative confidence level, through Gram-Schmidt orthonormalization, on the quality of our search direction in the context of multiple gradients evaluation.

Another important issue here is the sampling size of the functioning parameter range which from the theory [11, 9, 10] should be larger by one than the size of the control space. This is presented as a necessary and sufficient condition in [11]. But, this worst-case hypothesis is only necessary if all the associated gradients (i.e. evaluated at the sampling points) are linearly independent which is never the case in optimizations involving a state equation. This is because the governing equations introduce continuity features in the behavior of the state with respect to the functioning parameters: one expects, mainly, a small perturbation in a functioning parameter to introduce a small perturbation in the state. Again, Gram-Schmidt procedure will show that large dimensional parametric optimization problems can be treated with very small sampling of the functioning parameters range with marginal losses on the gradient informations. As a consequence, the *'plus one the size of the control space' condition* for the sampling is neither necessary nor sufficient for the search space based on this sampling to be exhaustive.

Finally, it is important to notice that there is no randomness involved in this problem in the sense that for a given choice of the functioning parameter, the simulation loop is fully deterministic.

2. Robust parametric optimization

We are interested by a class of optimization problem where the cost function involves a parameter α not considered as design parameter (denoted by x):

$$\min_{x\in\mathbf{O}_{ad}} j(x,\alpha),\ \alpha\in\mathbf{I}\subset\mathbb{R}^n, \mathbf{O}_{ad}\subset\mathbb{R}^N,\ n<<N. \tag{1}$$

$\mathbf{O}_{ad}$ is the optimization admissible domain. Usually, the functioning parameters α are just a few. Typical situations of interest are where $n=1$ or $n=2$.

2.1. Choice of the functional.

One classical approach to robust optimization is to control first and second momentum (i.e. mean and variance) of the functional:

$$\min_{x\in\mathbf{O}_{ad}} \mu=\mathbb{E}(j(x,\alpha)),\ \text{such that } \sigma^2=\mathbb{E}(j^2(x,\alpha))-\mathbb{E}^2(j(x,\alpha))\leqslant TOL. \tag{2}$$

When n is large Monte Carlo simulations permit to recover these momentum with an error decreasing as $\sigma/\sqrt{M}$ with M the number of functional evaluations. One sees that the convergence rate is independent of n. But, for $n<8$, classical numerical integration over-performs Monte Carlo simulations in term of complexity based on the number of functional evaluations to recover at a given accuracy these momentum. As we are interested by small values of n, this latter should therefore be preferred which eventually leads to the optimization of a weighted sum under constraint.

Let us assume a uniform probability density function (PDF) for each of the α component and a uniform sampling $\mathbf{I}_M$ of I and define μ and σ as:

$$\mu = \frac{1}{M} \sum_{\alpha_k \in \mathbf{I}_M} j(x, \alpha_k), \ \sigma^2 = \frac{1}{M-1} \sum_{\alpha_k \in \mathbf{I}_M} (j(x, \alpha_k) - \mu)^2. \tag{3}$$

With a different PDF, the sampling $\mathbf{I}_M$ should have been chosen accordingly. But for the sake of simplicity, let us consider uniform PDF for the sequel.

Better accuracy can be achieved at given M with a numerical integration on a non uniform grid introducing weights ω_k through:

$$\mu_* = \frac{1}{\Omega} \sum_{\alpha_k \in \mathbf{I}^*_M} \omega_k j(x, \alpha_k), \ \Omega = \sum_{k=1}^{M} \omega_k. \tag{4}$$

The non uniformity, which is linked to adapted grid $\mathbf{I}^*_M$, requires some knowledge of j: we are in the context of adaptive numerical integration and mesh adaptation. We would like to consider this idea of introducing weights even working with uniform grids (i.e. $\mathbf{I}^*_M = \mathbf{I}_M$) and with the weights ω_k accounting for the kind of performance we eventually want for the design: constant performance over the functioning parameters ranges.

Another point of interest is regularity. We would like the design to have a regular behavior with respect to α. Something which is not necessarily achieved when controlling the variance expressed through expression (3). Hence, supposing that the components of α are uncorrelated, we propose to consider the following alternative easy to achieve for small n, instead of the variance by (3):

$$\sigma^2_* = \frac{1}{2} \sum_{\alpha_k \in \mathbf{I}^*_M} < \nabla_\alpha j(x, \alpha_k), \nabla_\alpha j(x, \alpha_k) >, \tag{5}$$

where $<,>$ is the Euclidean scalar product. For small n computing $\nabla_\alpha j$ is a very easy task using finite differences, for instance.

Such an approximation is also used in First-Order Second Moment (FOSM) methods [5] to provide local estimation of the variance necessary in the definition of the local reliability index (ratio of the local mean estimation to local variance estimation). σ^2_* is therefore the sum of the FOSM variance over our sampling $\mathbf{I}^*_M$. So this choice can be seen as a variance estimator better addressing regularity issues. Numerical examples will show the impact on the design of considering functionals of the form $\mu_* + \beta\sigma_*, \ \beta > 0$.

3. Algorithmic considerations and complexity

A typical iterative descent algorithm for the minimization of $J = \mu_* + \beta\sigma_*$ involving a direct simulation chain linking the parameters (x, α) to the state U solution of a state equation and to the functional j is:

(6)
$$\begin{cases} x_0, \beta, \rho, \mathbf{I}_M, l_{max}, TOL = \text{given}, \\ \text{optimization iterations } l = 1, ..., l_{max} \\ \quad \begin{cases} M \text{ parallel state equation solutions } F(U(q(x_l), \alpha_k)) = 0, \ \alpha_k \in \mathbf{I}_M \ , \\ M \text{ parallel evaluations of } j(x_l, \alpha_k), \ \alpha_k \in \mathbf{I}_M \ , \\ M \text{ parallel solutions of the adjoint state } V \text{ equation } V^t F_U(U(q(x_l), \alpha_k)) = j_U^t, \ \alpha_k \in \mathbf{I}_M \ , \\ M \text{ parallel evaluations of } \nabla_x j(x_l, \alpha_k) = j_x + (V^t F_x)^t, \ \alpha_k \in \mathbf{I}_M \ , \\ \text{define } d \text{ the descent direction, } d = \nabla_x \mu_* + \beta \nabla_x \sigma_*, \\ x_{l+1} = x_l - \rho d, \\ \text{Stop if } \|d\| \leqslant TOL, \end{cases} \end{cases}$$

where

$$\nabla_x \mu_* = \frac{1}{\Omega} \sum_{\alpha_k \in \mathbf{I}_M} \omega_k \nabla_x j(x, \alpha_k), \tag{7}$$

$$\nabla_x \sigma_* = \sum_{\alpha_k \in \mathbf{I}_M} < \nabla_\alpha j(x, \alpha_k), \nabla_{\alpha x} j(x, \alpha_k) > . \tag{8}$$

$\nabla_{x\alpha} j(x, \alpha_k)$ is obtained from $\nabla_x j(x, \alpha_k)$ by finite differences on $\mathbf{I}_M$, component by component, following what has been done for $\nabla_\alpha j$.

Despite the natural presence of parallelism in this algorithm in the M independent evaluations of the state, functional and its gradient, computational complexity remains an issue. We will discuss the reduction of the sampling size and the use of incomplete sensitivity concept in the evaluation of the gradients. This latter permits to avoid the solution of the M adjoint equations. This is particularly suitable when using black-box state equation solvers not providing the adjoint of the state variables.

4. Reducing the sampling size

Handling fine samplings is very costly for realistic situations. We would like therefore to reduce the size of the sampling and have a quantitative confidence indicator during optimization on the pertinence of this reduction. Previous works suggest that the sampling size M should be larger (by one) than the size of the control space N [11] (which is also the number of components of $\nabla_x j$). This analysis relies on the fact that one needs N vectors to fill a vector space of dimension N and with $N + 1$ there is, therefore, no more direction left for descent. But, to make sure to fill a design search space of dimension N, the gradients $\nabla_x j(x, \alpha_k)$ generated at the sampling points $\alpha_k, k = 1, .., N$ should be linearly independent. Therefore, a criteria on the size alone is not enough. It should be completed by a condition on the linear independence of the gradients. Hence, the crucial point is rather to make sure the sampling is such that the associate gradients are linearly independent in order to fill the largest possible variation space and increase the sampling size as far as one finds new independent directions. In that sense, for large N, $N + 1$ is definitely a worst-case scenario and, for small N, there is no guarantee that the gradients at a uniform sampling of size $N + 1$ are independent.

This can be summarized as: if $N >> 1$ and $\nabla_x j(x,\alpha_k) \in \mathbb{R}^N$ (has N components), $\lim_{p\to\infty} \dim(Span(\nabla_x j(x,\alpha_k),\ \alpha_{k=1,..p})) << N$.

Indeed, experiences show that in large applications, such as the 3D shape optimization problem we will discuss in 9, generating the whole search space is actually possible with a sampling of size much less than the size of the control space as the rank of the subspace generated by the gradients of the functional at the different sampling points is small compared to the number of control parameters. This is an important reduction in complexity as in a 3D shape optimization problem with hundreds of shape control parameters, $M+1$ will be too large for a practical implementation.

4.1. **Link with Gram-Schmidt orthonormalization .** The previous discussion on the optimal sampling size can be linked with the Gram-Schmidt orthonormalization of the set of the gradient of the functional generated on a fine sampling of the functioning parameter range. The best sampling is the one with the smallest size and with the gradient vectors still generating the design search space. This latter is unknown, but it is a subset of $\mathbb{R}^N$.

Suppose one has a very fine sampling of $\mathbf{I}$ called $\mathbf{I}_\infty$. Our aim is to find the smallest M such that:

$\mathbf{S}_\infty = Span(\nabla_x j(x,\alpha_k),\ \alpha_k \in \mathbf{I}_\infty) = Span(\nabla_x j(x,\alpha_k),\ \alpha_k \in \mathbf{I}_M) = \mathbf{S}_M$.

One indication on M is by Gram-Schmidt orthonormalization (in practice modified Gram-Schmidt for numerical stability) where M will be the rank at which the rest after successive projections vanishes, or rather is small enough:

$$
\begin{cases}
\mathbf{I}_\infty,\ \mathbf{S}_\infty,\ TOL = \text{given},\\
u_1 = \nabla_x j(x,\alpha_1),\\
\text{iterations } k = 1, ...\\
u_k = \nabla_x j(x,\alpha_k) - \sum_{j=1}^{k-1} \frac{<\nabla_x j(x,\alpha_k), u_j>}{<u_j, u_j>} u_j,\\
\text{until } \|u_k\| < TOL,\\
M = k.
\end{cases}
\tag{9}
$$

This is of course out of reach exactly in practice as it requires $\mathbf{S}_\infty$. It is also subject to the choice of α_1 and the way α_k are sorted. The approach gives, however, indications on how to proceed to have some confidence level on the quality of a partial search space due to a partial sampling. By verifying, for instance, if a new direction built for a randomly chosen new α is still linear combination of previous directions. If not, one should redefine the multi-point optimization including this new sample point. Hence, one could start with a sampling $\mathbf{I}_{M'}$ and make sure during optimization that we always have $M < M'$ such that:

$\mathbf{S}_M = Span(\nabla_x j(x,\alpha_k),\ \alpha_k \in \mathbf{I}_M) \subseteq Span(\nabla_x j(x,\alpha_k),\ \alpha_k \in \mathbf{I}_{M'}) = \mathbf{S}_{M'}$.

We show in problem 9 an example of how Gram-Schmidt orthonormalization can be used to determine M the size of the free subspace, defined for a given accuracy TOL, in the space of all the gradients generated for the chosen sampling $\mathbf{I}_{M'}$. Doing so at each iteration of the optimization gives a clear indication on the quality of our sampling used to define the multi-point formulation.

5. Sensitivity analysis

In algorithm 3, at each iteration of minimization we need to provide $\nabla_x j(x_l, \alpha_k)$ for different values of α_k. This is computer intensive even if it is fully parallel. Our aim is to see how to reduce the complexity of this optimization problem using incomplete sensitivity as in the case of mono-point optimization without linearizing the state equation.

We briefly recall the Hadamard incomplete sensitivity approach [1, 7]. Consider a general simulation loop linking the independent parameter x to a functional j:

$$j(x) : x \to q(x) \to U(q(x)) \to j(x, q(x), U(q(x))), \tag{10}$$

where q and U are dependent variables. In our discussion, the difference between the two is mainly on the cost of getting them. See, for instance, q as geometrical quantities and U representing state related variables, solution of some costly governing equation. The gradient of j with respect to x is obviously $\nabla_x j = j_{,x} + (j_{,q} + j_{,U} U_{,q}) q_{,x}$. The major part of the cost of this evaluation is due to $U_{,q}$. Hadamard incomplete sensitivity addresses the following context:

- the cost function j and control x have the same domain of definition $\mathbf{D}(x)$ (e.g. a shape and an aerodynamic coefficient defined over it),
- j is a product of functions such as $j(x) = f(x, q(x))\ g(U(q(x)))$.

If these requirements hold, we can use an incomplete evaluation of this gradient, neglecting the sensitivity with respect to the state, leading to the approximation $\nabla_x j = j_{,x} + j_{,U} q_{,x} = \nabla_x f g$. This is very interesting as $\nabla_x f$ can be analytically calculated in most situations. These are also locally defined on the domain of definition $\mathbf{D}(x)$ of x and do not involve the full domain of definition of the state variable U or even q. For instance, in the case of a shape optimization problem such as in problem 9, there will be no field variable linearized as everything will be defined on the shape. This locality issue is very important in parallel computing where data locality is aimed as much as possible to reduce the communications.

6. An analytical example

Let us illustrate our purpose on the multi-point definition of the functional and on the nonuniform weighting using a simple analytical example where α is a scalar and

$$j(x, \alpha) = \frac{1}{2} \sum_{i=1}^{N} (x_i - \alpha)^2,\ -0.5 = \alpha_{min} \leqslant \alpha \leqslant \alpha_{max} = 0.5,\ -5 \leqslant x_i \leqslant 5,\ N = 40. \tag{11}$$

The optimality condition for $j(x, \alpha)$ gives obviously $x_i = \alpha,\ i = 1, .., 40$.

Let us solve our problem by a gradient method and minimize the functional $\mu_* + \beta\sigma_*$ with $\beta = 0$ or 1 where μ_* and σ_* are given by (4) and (5). As we said, the weights ω_k can be modified to account for the kind of performance we eventually want for the design: constant performance over the range of α.

Suppose one has performed a first optimization with a uniform weighting ($\omega_k = 0.1$) and the observed performance is as presented in figure 1. A choice of ω_k giving more weight to regions away from where the best performance is achieved could be:

$$\omega_k = (1 - \exp(-a(\alpha_k - \alpha^{opt})^2) + b,\ 0 < b << 1, \tag{12}$$

with $a = 1/(\alpha_{max} - \alpha_{min}) = 1$ and , $b = 0.1$ and $\alpha^{opt} = 0$.

Figure 1 shows the distribution of $j(x, \alpha_k)$ after these minimization with $M = 40$. The solution to the first minimization problem is the center of $\mathbf{O}_{ad}$ which is a rather non robust design. The second minimization produces a much flatter functional over the range of α. In this case, this is achieved only through degradation in the performance of the design. In the sequel, we will see that this is not necessarily always the case.

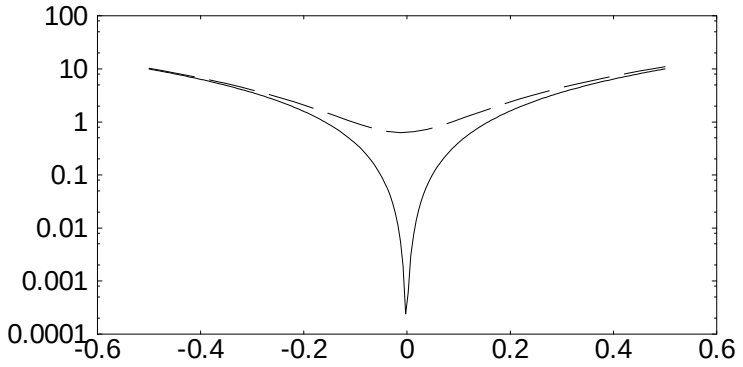

FIGURE 1. Robust minimization: $j(x_{opt}, \alpha)$ given by (11) vs. α where x_{opt} is solution after minimizing $\mu_* + \beta\sigma_*$ for $\beta = 0$ and regular weighting (continuous) and $\beta = 1$ and non uniform weighting (dashes).

7. Parametric optimality condition analysis on a model problem

Let us consider a situation where optimality conditions can be derived and the optimal solutions expressed without applying an optimization procedure. Consider a cost function given by:

$$j(x, \alpha) = x^m u_y(y = x, \alpha),\ x \in [-0.1, 0.1],\ \alpha \in [0.2, 5],\ m \in \mathbb{N}^*$$

and as state equation the following convection-diffusion equation with our α being the Peclet number:

$$u_y - \alpha^{-1}\ u_{yy} = 0,\ \text{on }]x, 1[,\ u(y = x, \alpha) = 0,\ u(y = 1, \alpha) = 1.$$

The solution of this equation is

$$u(y, \alpha) \ = \ \frac{\exp(\alpha\ x) - \exp(\alpha\ y)}{\exp(\alpha\ x) - \exp(\alpha)}. \tag{13}$$

Hence

$$\begin{aligned} u_y(y, \alpha) &= \frac{-\alpha\ \exp(\alpha\ y)}{\exp(\alpha\ x) - \exp(\alpha)}, \\ (u_y)_{,x}(y = x, \alpha) &= \frac{(\alpha\ \exp(\alpha\ x))^2}{(\exp(\alpha\ x) - \exp(\alpha))^2} = u_y^2(y = x). \\ j_{,x}(y = x, \alpha) &= x^{m-1} u_y(x, \alpha)(m + x u_y(x, \alpha)). \end{aligned}$$

The first optimality condition for $J(x) = \sum_{\alpha_k \in \mathbf{I}_M} \omega_k \; j(x, \alpha_k)$ with $\mathbf{I}_M$ a regular sampling of $[0.2, 5]$, the chosen range for α, gives critical points when:

$$\sum_{\alpha_k \in \mathbf{I}_M} \omega_k \; x^{m-1} u_y(x,\alpha)(m + x u_y(x,\alpha)) \; = \; 0. \tag{14}$$

For $m > 1$, $x = 0$ is always the optimal solution. For $m = 1$, there is no unique optimal solution for all α. But, as shown in Figure 2, the proposed formulation always reduces the sensitivity of the functional with respect to the functioning parameter α. Figure 2 shows $J'_x(x)$ and $j'_x(x, \alpha)$ with a 40×40 sampling of the (x, α) parameter space using uniform and non-uniform weightings. From figure 2, one sees that the nonuniform weighting (12) used in problem 6 can be again applied here.

The second term between parenthesis in (14) is the state linearization contribution. Dropping it leads to what we call the incomplete sensitivity. Figure 3 shows the absolute error one commits dropping this contribution. One sees that the approximation is better for large m. Also the incomplete optimality condition gives the same optimal solution for $m > 1$. As we said, this approximation only holds for special functionals made as products of control by state functions: $j(x, \alpha) = f(x)g(u(x, \alpha))$ [1].

8. Shape optimization with a reduced-order fluid model

Let us consider a situation with α being of dimension two. One considers a flow problem in two dimension in space where the pressure distribution along the shape $\Gamma(x)$ follows the so-called Newton law:

$$\begin{aligned} p(\Gamma(x), \alpha) &= \frac{1}{2}\rho_\infty \|u_\infty(\alpha)\|^2 \left(\frac{u_\infty(\alpha)}{\|u_\infty(\alpha)\|}.n(\Gamma(x)) \right)^2 \\ &= \frac{1}{2}\rho_\infty (u_\infty(\alpha).n(\Gamma(x)))^2, \end{aligned} \tag{15}$$

where subscribe ∞ denotes inflow quantity for the density and velocity and n the local outward normal to the shape. This relation gives satisfactory prediction of the pressure distribution over a simple shape (such a cylinder) in a very low-speed flow.

The control space (shape parameterization) is of dimension 3 and the functional involves pressure target over the shape $\Gamma(x)$ described by the parameter $x = (x_1, x_2, x_3)$. More precisely, the shape $\Gamma(x)$ is given by:

$$\Gamma(x) = \{(y_1, y_2), \text{such that } y_1 \in [-\pi, \pi], \; y_2 = \Pi_{i=1,3} \cos(x_i y_1)\}.$$

One would like to realize best a target pressure distribution p_{des}, taken constant here, over a given range of the inflow velocity intensity (U_{in}) and its incidence (θ) defined by $\alpha = (U_{in}, \theta)$. Using the notations above, we have:

$$j(x, \alpha) = \frac{1}{2} \int_{\Gamma(x)} (p(\Gamma(x), \alpha) - p_{des})^2 \; d\gamma. \tag{16}$$

We use a gradient method to minimize the functional $\mu_* + \beta\sigma_*$ for $\beta = 0$ and a regular weighting and then for $\beta = 1$ and a non uniform target-based weighting derived from the outcome of the first design. Figure 4 shows the performances with the uniform weighted sum and a non uniform weighting together with regularity control through (4) for a 30×30 sampling of the two dimensional functioning space (flow incidence in degree and intensity in cm/s). We see from this figure that the weighting correction should mainly account for the variations due to the angle of

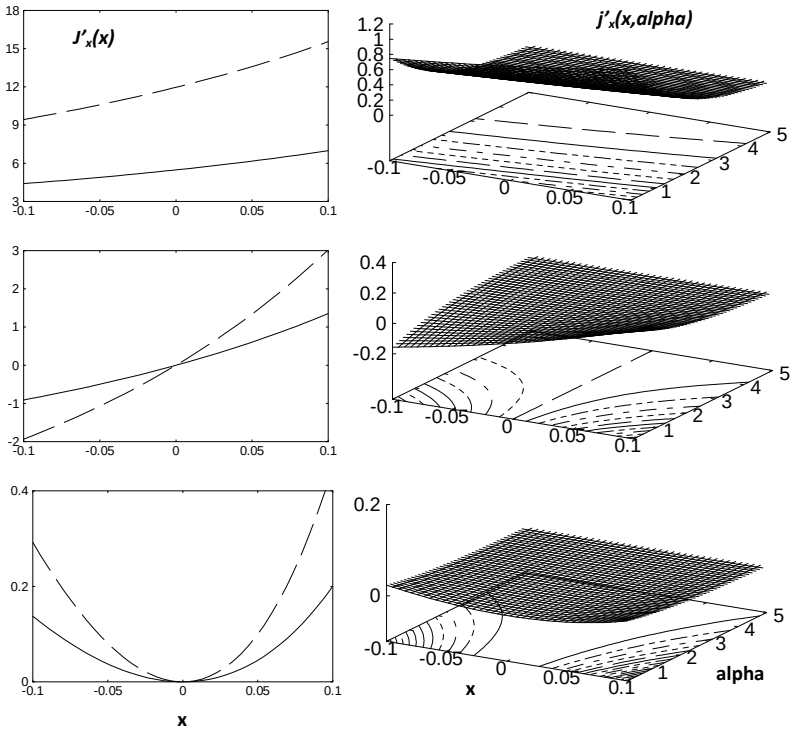

FIGURE 2. From the top $m = 1, 2$ and 3 in problem 7. $J'_x(x)$ (left column) and $j'_x(x, \alpha)$ (right column) for a 40×40 sampling of the parameter space (x, α) with a uniform (dashed) and the proposed corrections (continuous). For $m > 1$, $x_{opt} = 0$ for all α and is given by both weightings.

incidence and that the inflow intensity has less impact on the functional according to this model. We introduce therefore an anisotropic metric in the definition of ω_k:

$$\omega_k = (1 - \exp(-\|\alpha_k - \alpha^{opt}\|_*^2) + b,\ 0 < b << 1, \tag{17}$$

$$\|\alpha_k - \alpha^{opt}\|_*^2 = (\alpha_k - \alpha^{opt})^t \mathbf{A} (\alpha_k - \alpha^{opt}),$$

with $\mathbf{A}$ a 2×2 positive definite matrix which permits an anisotropic distance definition in the domain of variation of $\alpha = (U_{in}, \theta)$. In this example we consider:

$$\mathbf{A} = \text{diag}(\frac{0.01}{U_{in}^{max} - U_{in}^{min}}, \frac{1}{\theta^{max} - \theta^{min}}),\ b = 0.1,\ \text{and } \alpha^{opt} = (1.5, 0),$$

with $U_{in}^{max} = 1.6cm/s$, $U_{in}^{min} = 1.3cm/s$, $\theta^{max} = 2^o$, $\theta^{min} = -2^o$. The new optimal solution appears more stable as all the components of $\nabla_\alpha j(x_{opt}, \alpha)$ have been reduced and the functional is almost flat over the ranges of the parameters. Figure 4 also shows the shapes obtained from the two optimizations. This is non intuitive as one would have expected that the most robust shape to be also the simplest.

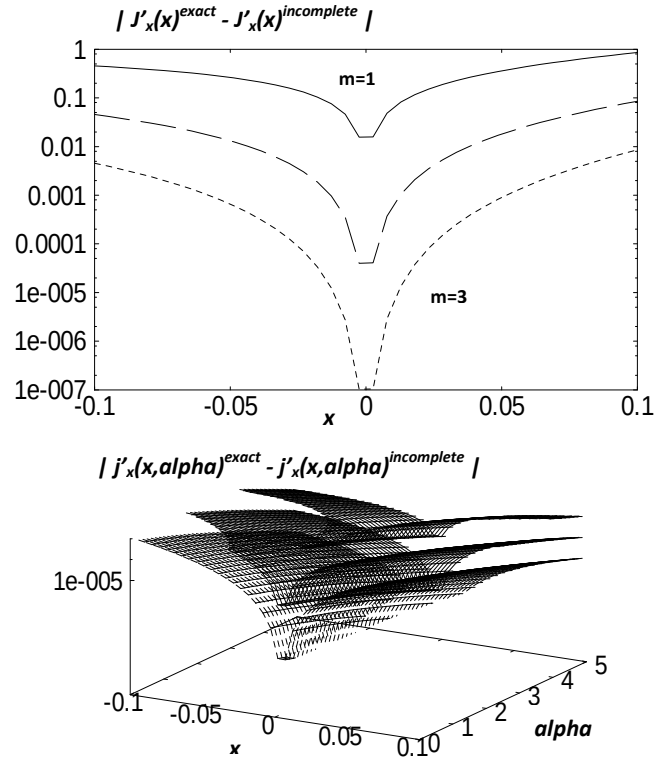

FIGURE 3. Absolute error between the sensitivity and its incomplete evaluation for $J'_x(x)$ (top) and $j'_x(x,\alpha)$ (bottom) for problem 7. The error decreases with m increasing.

8.1. **Incomplete sensitivity .** To discuss incomplete sensitivity analysis, let us for simplicity consider both x and α as scalar quantities and write in short $\Gamma(x) = x$ and write the functional (16) as a point-wise quantity:

$$j(x,\alpha) = \frac{1}{2}(p(x,\alpha) - p_{des})^2.$$

This functional is not in the validity domain of incomplete sensitivities as it is not made of a product of a function of the state variable by a function of geometrical quantities. Its gradient with respect to x gives:

$$j'_x = (p(x,\alpha) - p_{des})p'_x.$$

The incomplete sensitivity is therefore zero when dropping the state linearization contribution.

Let us consider another situation which will also be considered in 3D in problem 9. Neglecting viscous effects, designing a shape with minimum drag involves an integral on the shape of $p(x,\alpha)(u_\infty(\alpha)\cdot n(x))$. Suppose the pressure is given by formula (15).

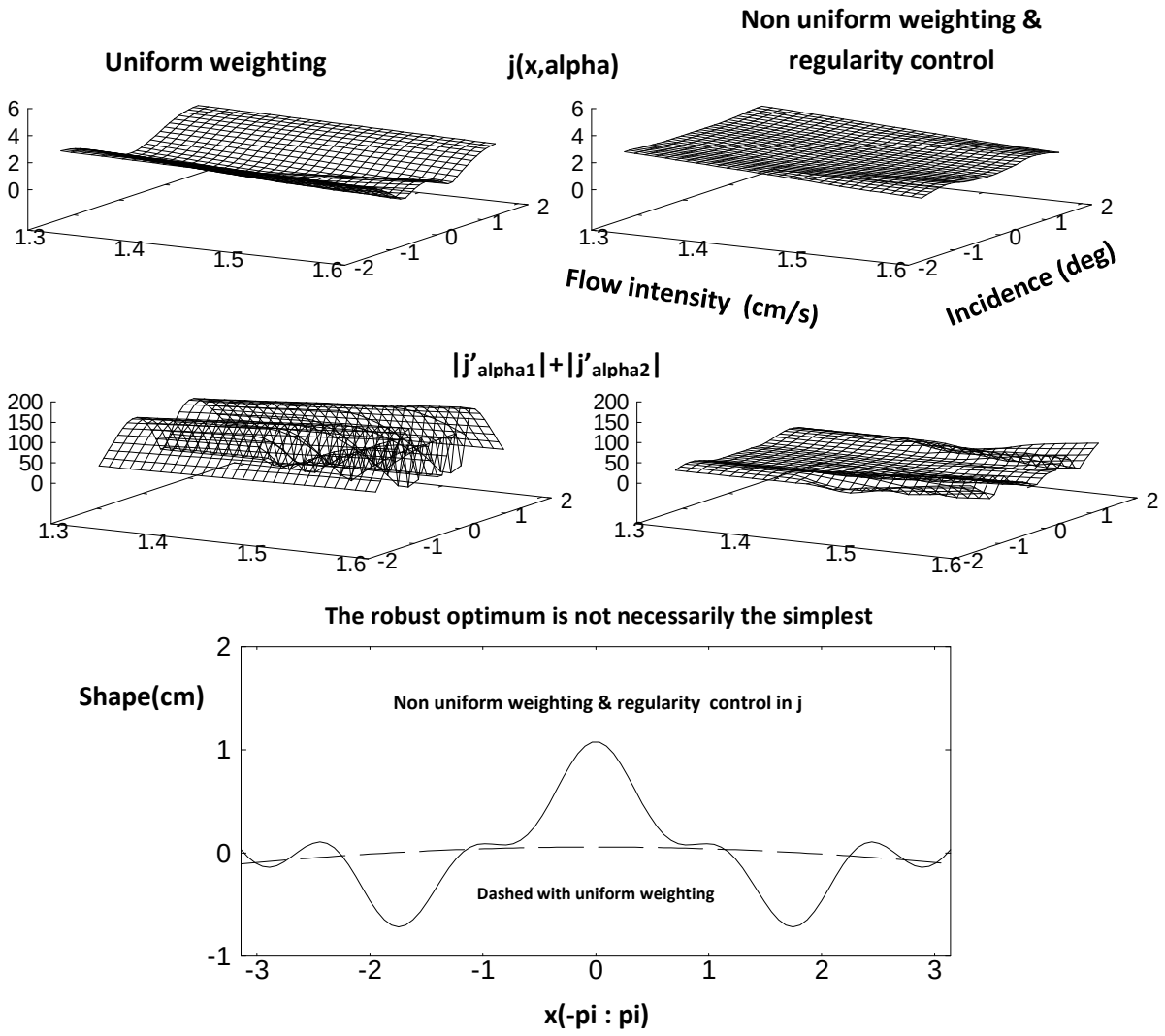

FIGURE 4. Robust shape optimization with uniform weighting and based on (12) with regularity control with $j(x,\alpha)$ given by (16). Upper: functional (16) vs. U_{in} inflow velocity and θ angle of attack of the design by uniform weighting and with the proposed corrections on a 30×30 sampling. Middle: $|\partial_{U_{in}} j(x_{opt},\alpha)| + |\partial_\theta j(x_{opt},\alpha)|$, a measure of robustness. Lower: the two final shapes (dashed by uniform weighting).

We have therefore

$$p(x,\alpha)u_\infty(\alpha)\cdot n(x) = \frac{1}{2}\rho_\infty(u_\infty(\alpha)\cdot n(x))^3.$$

Its derivative with respect to x is

$$(p(x,\alpha)u_\infty(\alpha)\cdot n(x))'_x = pu_\infty\cdot n' + p'_x(u_\infty\cdot n).$$

The first term in the right-hand-side is the incomplete sensitivity and the second term involves the linearization of the pressure. Adding the two, we have:

$$(p(x,\alpha)u_\infty(\alpha)\cdot n(x))'_x = \frac{3}{2}\rho_\infty u_\infty (u_\infty\cdot n)^2 n'.$$

On the other hand, the incomplete sensitivity alone gives:

$$(pu_\infty)n' = 1/2\rho_\infty u_\infty (u_\infty\cdot n)^2 n'.$$

The two derivatives realize the same optimality condition and only differ by a factor of 3 which is of no importance with a descent method. This has been successfully used in mono-point drag reduction shape optimization problems [1] but in multi-point optimizations the associated computational effort reduction will be even more significant.

The expression above can be rewritten as $pu_\infty.n = p|u_\infty|\cos(\frac{u_\infty}{|u_\infty|}.n)$. The incomplete gradient is therefore $p(u_\infty.n)_x = -p|u_\infty|\sin(\frac{u_\infty}{|u_\infty|}.n) = 0$ when n is aligned with u_∞. The incomplete sensitivity fails therefore for these area (e.g. area near the leading edge for instance for an airfoil at no incidence). On the other hand, it is exact if shape deformations are aligned with the local normal to the shape, if $\partial p/\partial n = 0$ boundary condition is applied along the shape. This remark is very important and permits to access at no cost the sensitivity of the aerodynamic coefficients with respect to shape deformations in the case of inviscid flows. We will use this in example 9.

8.2. **Sampling size analysis.** Figure 5 illustrates the relation between sampling and search capacity for problem 8. It shows the comparison of $\nabla_x(\mu_*+\beta\sigma_*)$ obtained at the same control point x for several different uniform samplings of size increasing from 4×4 to 30×30 with a uniform weighting with $\beta = 0$ and the non uniform one with both $\beta = 0$ and $\beta = 1$. The biggest impact comes when adding regularity control on $j(x,\alpha)$ through σ_*. The non uniform weighting and this regularity control can be added after one has performed a first optimization using a uniform weighting. One needs this first run to be able to define the weights and β. Indeed, β should be reduced if the functional is highly irregular in α in order for the constrained optimization problem not to be too stiff. The gradients differ in amplitude but indications on the search direction can be found after normalization. One sees that with the regular weighting approach one can degrade the sampling to 5×5 and still have the same search direction. With the non uniform weighting and controlling the regularity in the functional over the functioning parameter range even coarser samplings can be used. This is interesting as the non uniform weighting is a priori prescribed through (12) and does not need any tuning during optimization. Something which is not easy and this is one advantage of the present formulation. If a non uniform sampling is used, one expects being able to degrade the sampling to size $4 = 3+1$ (i.e. 2×2 or 3×1) to recover the theory in [11, 9]. We are back here to what said in section 4 on the relation between the sizes of the control space and the sampling.

9. Full aircraft robust shape optimization

This example concerns a shape optimization problem for a full aircraft in transonic cruise condition. For such configurations, an efficient mono-point optimization is already a difficult task as it involves several software. It also requires coherent

geometrical manipulation of the shape parameterization and the surface and volume meshes.

Let us briefly describe our shape optimization procedure [1, 7] where the direct dependency chain reads:

$$j(x,\alpha) : (x,\alpha) \rightarrow (q(x),\alpha) \rightarrow U(q(x),\alpha) \rightarrow j(x,q(x),U(q(x),\alpha)), \tag{18}$$

where x denotes a CAD-free parameterization (CAD: Computer Aided Design) [7] which does not require a priori local regularity assumptions on the shape as it is implicitly the case in CAD-based shape definitions. $q(x)$ denotes auxiliary unstructured mesh related geometrical quantities and $U(q(x),\alpha)$ flow variables depending also on extra parameters α not part of those involved in the definition of the shape. In example 8, $U(q(x),\alpha)$ was solution of an analytical model. In this problem $U(q(x),\alpha)$ denotes conservative flow variables solution of the Euler equations in conservation form.

Several sources of variability exist for these problems where large deviations exist between the prescribed shape and the shape during the flight. Shape deformations during the flight, which we suppose stationary, may be due to a change in the weight of the aircraft or due to variability in the flight conditions. In the former, the aircraft weight differs by several tons between departure and arrival and the shape deforms accordingly. This is even more visible on new generation of aircrafts using new materials allowing for large shape deformation. This shape variation should therefore be accounted for during the design through multi-point optimizations of aeroelastic shapes. The present approach eases handling this complexity. One example of the latter situation is when the aircraft cruises against transverse winds which are very common. In this case, the parameter α is the sideslip angle inducing fully 3D effects on the flow around the plane making necessary the consideration of a full aircraft during the design. Usually aircrafts are designed for a range of angle of incidence which permits to easily recover the lift coefficient thanks to the linear relationship between incidence and lift away from stall conditions. These designs are usually realized with the sideslip angle set to zero. We consider the sideslip angle α in $\mathbf{I} = [0, 10^o]$. Because the airplane geometry is symmetric spanwise, it is not necessary to consider a symmetric range for the transverse wind. Our aim is to reduce the sensitivity of the design to this wind.

In our mono-point design procedure, the derivative $\nabla_x j$ is computed by automatic differentiation in reverse mode using `tapenade` [8] and we have optimized by hand the reverse mode code for steady flow adjoint solutions in order to minimize memory requirements in reverse loops [1]. As we saw in algorithm (3) we need M evaluations of the state and functional sensitivities with respect to the shape. State evaluations are difficult to avoid but gradient evaluations can be avoided through incomplete sensitivity definition as explained in sections (5), (7) and (8.1). Our direct Euler code uses time marching to the steady solution with local time steps. The corresponding reverse (or adjoint) code can be easily used to produce the incomplete sensitivity by setting to one the number of reverse time iterations [1]. This is interesting as it permits to implement this approximation in existing adjoint based optimization platform with no extra coding. Also, we saw in section (8.1) that, due to the pressure boundary condition, in order for the incomplete sensitivity to be accurate we need shape deformations to be aligned with the normal to the shape. The parameter x represents shape deformations along the normal to the triangular faces of the surface

mesh as shown in figure 6. This search space has a dimension of about 5000 (i.e. $N = 5000$ here).

Let us consider a classical aerodynamic shape optimization which aims at minimizing the drag coefficient C_d:

$$j(x, q(x), U(x, \alpha), \alpha) = C_d = \frac{1}{2\rho_\infty \|\vec{u}_\infty(\alpha)\|^2} \int_{shape(x)} p(q(x), \alpha)(\vec{u}_\infty(\alpha).\vec{n}(q(x))d\gamma,$$

where superscript ∞ indicates inflow conditions. This minimization is usually performed under equality constraints on the lift and volume of the aircraft. Let us denotes these $C_1(x, q(x), U(x, \alpha), \alpha) = 0$ and $C_2(x, q(x), U(x, \alpha), \alpha) = 0$. These quantities are either geometric or enter the validity domain of incomplete sensitivity concept where one needs the functional to be of the form $j(x, q(x), U(q(x))) = \int_{shape} f(x, q(x))g(U(q(x)))$ which is product of geometry by state functions. Sensitivity evaluation of these constraints being cheap too, one prefers not to use weighting to define $J = j + \sum_{i=1,2} a_i C_i,\ a_i \in \mathbb{R}^+$ to be minimized, but rather to compute individual gradients and use projection to define a local admissible descent direction:

$$d = \nabla_x j - \sum_{i=1,2} < \nabla_x j, \nabla_x C_i > \frac{\nabla_x C_i}{\|\nabla_x C_i\|}.$$

Figure 7 shows the outcome of the Gram-Schmidt orthonormalization with two different sorting of the vectors in $Span(\nabla_x j(x, \alpha_k),\ \alpha_k \in \mathbf{I}_{200})$ for the initial shape. These curves show the evolution of the sampling size M necessary for the representation of all the gradients in $\mathbf{S}_{200}$ generated from a very fine sampling of the sideslip angle parameter range (this plays the role of $\mathbf{S}_\infty$). The subspace is never free and can always be generated by a subset of vectors. If $TOL = 10^{-3}$ the full subspace can be generated with only 5 gradient vectors. This number increases with TOL decreasing. The curve with $TOL = 10^{-4}$ illustrates an unsuitable situation where missing informations appear only after the addition of several new directions. Surprisingly, at some point it is best to increase the required accuracy as one eventually needs more vectors with $TOL = 10^{-5}$ than with $TOL = 10^{-6}$. In cases such behavior is observed, one should proceed to several designs with various TOL and see the deviations between the designs. The small differences in the details in high dimensional search directions are important when looking for very fine tuning of a design. This is something we are currently exploring. On the other hand, one could argue that the differences observed for $TOL < 10^{-4}$ fall below our overall numerical accuracy.

Figure 8 shows four gradient fields $\nabla_x j(x, \alpha_k)$ between 200 for the sideslip angles of $\alpha_k = 1, 3, 6$ and 10 degrees. One sees that the differences are mainly located along the cabin and vertical empennage (not shown here), more sensitive to the transverse wind. This explains why just a few gradients are suffisient to generate $Span(\nabla_x j(x, \alpha_k),\ \alpha_k \in \mathbf{I}_{200})$ for $TOL = 10^{-3}$ in (9).

In the sequel, we consider the case of $TOL = 510^{-4}$ only. Let us discuss the sampling size issue. Above we considered a uniform fine sampling of $[0, 10^o]$ with 200 points. In principle, one should have considered a sampling of the size larger by one than the size of the control space $N = 5000$. As we said, such a fine sampling is only necessary if all the associated gradients are linearly independent which is clearly not the case here. Indeed, the previous analysis by the Gram-Schmidt procedure showed that even 200 is too large and that for $TOL = 10^{-3}$, 5 gradients vectors on the initial

shape were enough to generate the whole search space defined by the 200 gradients meaning that $M = dim(Span(\nabla_x j(x, \alpha_k),\ \alpha_k \in \mathbf{I}_{200}) = 5,$ for $TOL = 10^{-3}$. Let us analyse the evolution of this dimension during optimization working with a sampling of size 30 (the conclusions would be the same with 200, but the simulations much more computer intensive).

Figure 9 shows the evolution of $M = dim(Span(\nabla_x j(x, \alpha_k),\ \alpha_k \in \mathbf{I}_{30}))$ with $TOL = 5.10^{-4}$ for $\mu_* + \beta\sigma_*$ with $\beta = 0$ and 1 and with uniform and non uniform weightings. This guarantees that all the informations on possible search directions are accounted for during optimization. With the non uniform weighting M is larger showing wider search space. In all cases, M the dimension of the free subspace search space remains below 8 making therefore 30 a safe choice for the sampling size.

Figure 10 shows the evolution of μ and σ^2 given by (3) of $j(x, \alpha_k),\ \alpha_k \in \mathbf{I}_{30}$ during the minimization of $j = \mu_* + \beta\sigma_*$ with $\beta = 0$ and 1 and with uniform (dash) and non uniform (continuous) weightings. The two final cumulative shape deformations are quite different as shown in figure 11 but provide the same average performance. The design with a non uniform weighting and with regularity control has less variability.

Figure 11 shows an upper-view of shape deformations for the minimization of $j = \mu_* + \beta\sigma_*$ with $\beta = 0$ and 1 and with uniform and non uniform weightings. Despite the differences in individual gradients in figure 8 were mainly located on the cabin, the final design based on gradients defined through (7) and (8) features sensible differences on the wings too. This indicates that for a robust design the wings, cabin and empennage should be designed together.

10. CONCLUDING REMARKS

In order to be easily integrated to existing shape optimization platforms using descent methods, robust shape optimization has been addressed through multi-point formulations. The aim was to provide an algorithm with the same time-to-solution characteristics than in the mono-point situation. The extra coding for this implementation is marginal as it only involves existing ingredients which need to be used independently in parallel. Gram-Schmidt orthonormalization permits to introduce confidence level on the quality of our search space in regard to the multi-point analysis. Hadamard incomplete sensitivity concept has been used together with the sensitivity of the functional with respect to functioning parameters to provide the gradient of a multi-point weighted sum without any linearization of the state equation. In order to make the multi-point analysis more efficient, the requested characteristics of the target of the optimization has been introduced in the definition of the functional through target-based weightings. The concepts have introduced through simple examples and eventually applied to the robust design of a full aircraft in transonic cruise condition over a range of transversal wind.

Acknowledgements The author would like to thank F. Gallard from Airbus for his feedbacks on the aircraft optimization problem.

References

[1] Mohammadi B., Pironneau O. *Applied Shape Optimization for Fluids* (2nd Edition), Oxford Univ. Press, 2009.

[2] Schillings C., Schmidt S., and Schulz V. Efficient shape optimization for certain and uncertain aerodynamic design. *Computers & Fluids*, 46/1, 78:87, 2011.

[3] Bungartz, H.-J. and Griebel, M. Sparse Grids, *Acta Numerica*, **13**, 147:269, 2004.

[4] Wan X., Karniadakis G.E., Multi-element generalized polynomial chaos for arbitrary probability measures, *SIAM J. Sci. Comput.* 28/3, 901:928, 2006.

[5] Melchers, R. E. *Structural Reliability Analysis and Prediction.* John Wiley and Sons, Chichester, 1999.

[6] Conn A., Scheinberg K., Vicente L.*Introduction to Derivative-Free Optimization*, SIAM, NY, 2002.

[7] Mohammadi, B. Pironneau, O. Shape Optimization in Fluid Mechanics, *Annual Revue of Fluid Mechanics*, 36/1, 255:279, 2004.

[8] Hascoet, L. Pascual, V. Tapenade 2.1 user's guide. INRIA Technical Report RT-300, 2004.

[9] Gallard, F. Mohammadi, B. Montagnac, M. Meaux, M. Robust parametric design by multipoint optimization. CERFACS Technical Report TR/CFD/12/33. 2012.

[10] Desideri, A. Multiple-gradient descent algorithm for multiobjective optimization. CRAS. Vol. 350/5, 313:318, 2012.

[11] Li, W. Huyse, L. and Padula, S. Robust Airfoil Optimization to Achieve Consistent Drag Reduction Over a Mach Range, *Structural and Multidisciplinary Optimization*, 24/1, 38-50, 2002.

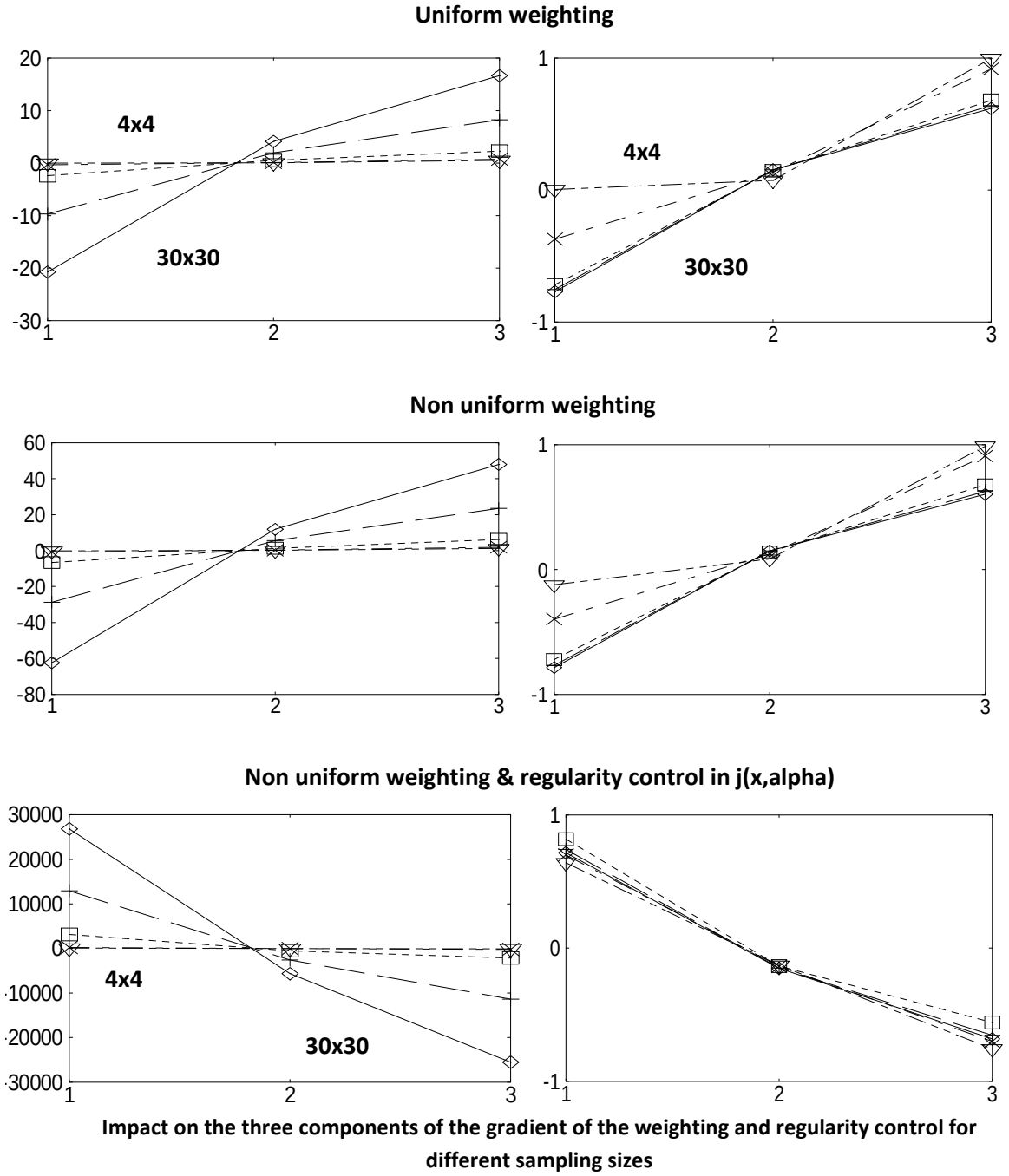

FIGURE 5. The three components of $\nabla_x(\mu_* + \beta\sigma_*)$ (left) and after normalization (right) for problem (8) evaluated with μ_* and σ_* defined on, respectively, a 4×4, 5×5, 10×10, 20×20 and 30×30 samplings of the functioning parameters (incidence and intensity of the flow). The lower-right picture shows that our a priori non uniform weighting together with regularity control in α on $j(x, \alpha)$ permits for $\nabla_x(\mu_*+\sigma_*)$ to be less sensitive to the choice of the sampling.

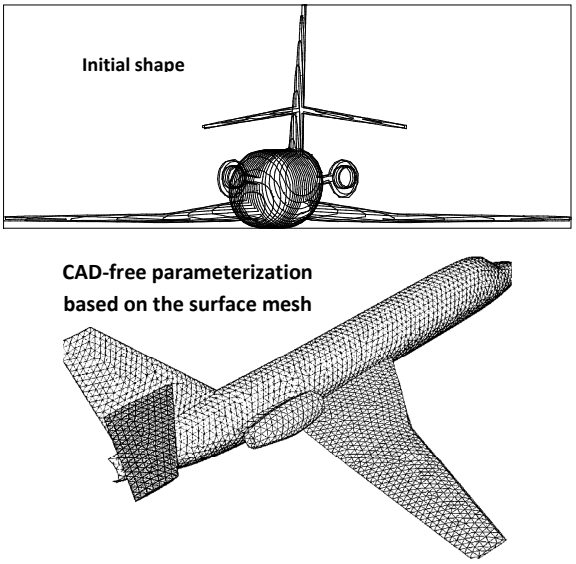

FIGURE 6. Initial shape and a view of the triangular surface mesh used to define shape deformations parallel to the local normal to the shape for problem 9. This in order for the incomplete sensitivity of aerodynamic coefficients to be exact for inviscid calculations thanks to the pressure boundary condition $\partial p/\partial n = 0$ over the shape.

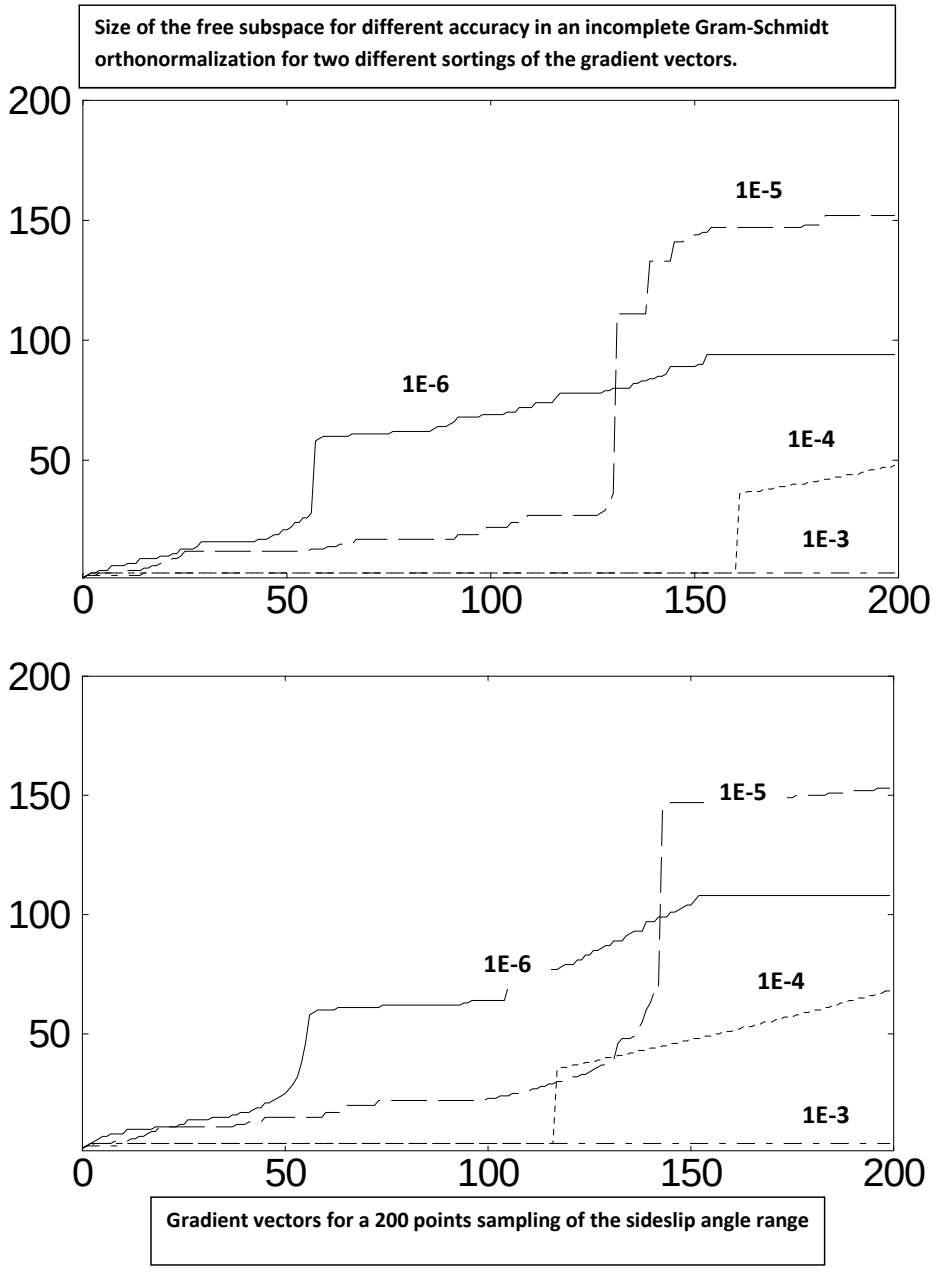

FIGURE 7. Gram-Schmidt orthonormalization with two different sorting of the gradient vectors for problem 9 for different requested accuracy TOL in algorithm 4.1. This indicates the size of the free subspace for different TOL. Hence, for $TOL = 10^{-3}$ the full gradient space can be generated with only 5 gradient vectors.

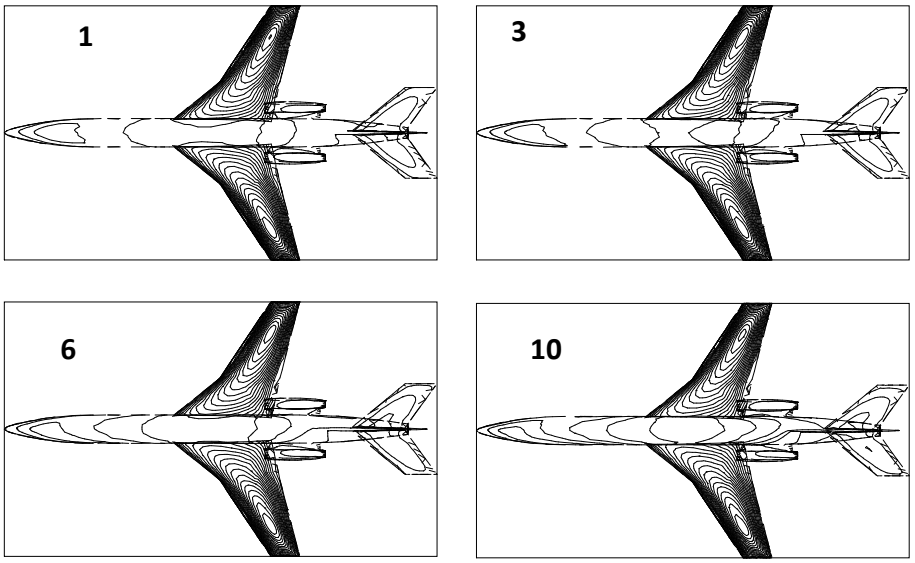

FIGURE 8. Four gradient fields $\nabla_x j(x, \alpha_k)$ for the sideslip angle $\alpha_k = 1, 3, 6$ and 10 degrees.

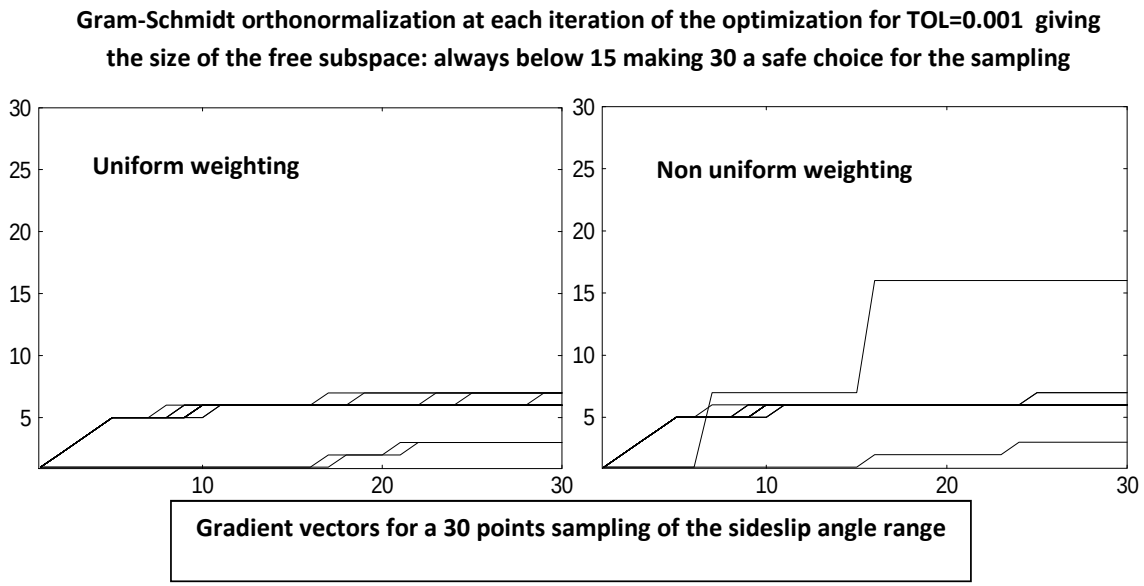

FIGURE 9. Gram-Schmidt orthonormalization during optimization and evolution of M, the size of the free subspace in $Span(\nabla_x j(x, \alpha_k),\ \alpha_k \in \mathbf{I}_{30})$.

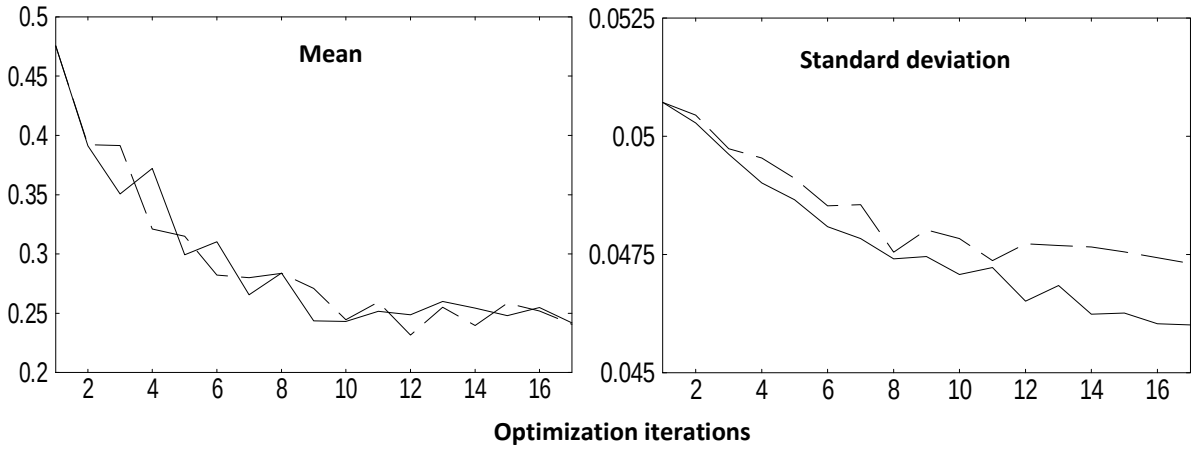

FIGURE 10. Evolution of the mean μ and standard deviation σ of $j(x, \alpha_k)$ $\alpha_k \in \mathbf{I}_{30}$ during the minimization of $j = \mu_* + \beta\sigma_*$ with $\beta = 0$ (dashed) and 1 (continuous) and with uniform (dashed) and non uniform (continuous) weightings. The two final shapes are quite different but provide the same average performance. The design with a non uniform weighting and with regularity control has less variability.

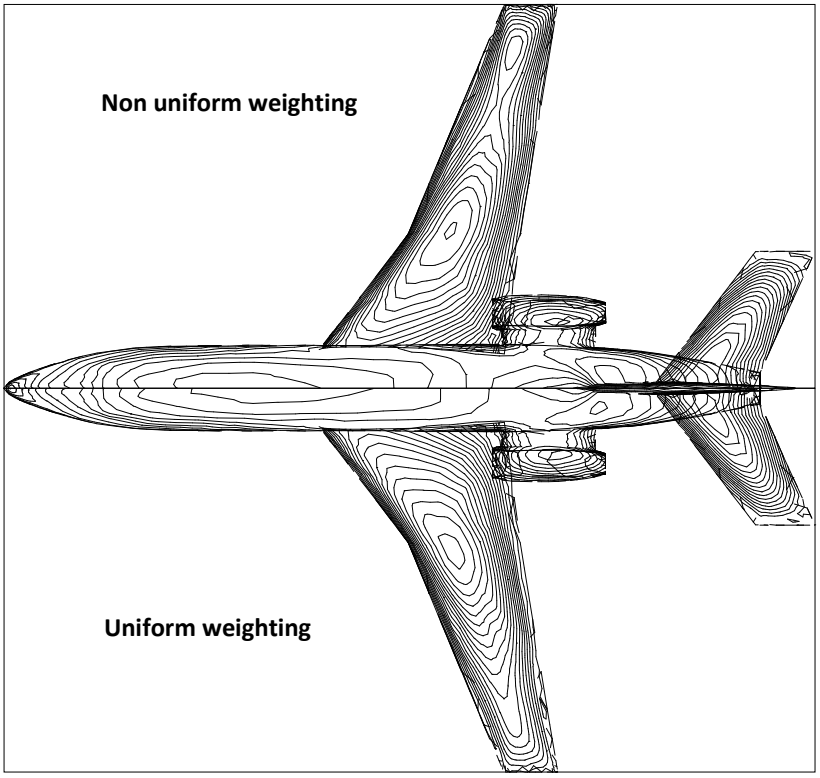

FIGURE 11. Upper-view of shape deformations for the minimization of $j = \mu_* + \beta\sigma_*$ with $\beta = 0$ and 1 and with uniform and non uniform weightings.