



Controlling First Four Moments for Robust Optimization

Bijan Mohammadi

► To cite this version:

Bijan Mohammadi. Controlling First Four Moments for Robust Optimization. Optimization and Engineering, 2017, 18 (3), pp.561-585. 10.1007/s11081-016-9337-y . hal-01366654

HAL Id: hal-01366654

<https://hal.science/hal-01366654>

Submitted on 19 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Controlling First Four Moments for Robust Optimization

Bijan Mohammadi

Montpellier University
Mathematics(IMAG), 34095 Montpellier, France
bijan.mohammadi@umontpellier.fr

ABSTRACT. The paper addresses the solution of robust moment-based optimization problems after a multipoint reformulation. The first four moments are considered (i.e. mean, variance, skewness and kurtosis) going beyond classical engineering optimization based on the control of the mean and variance. In particular, the impact on the design of a control of the third and fourth moments are discussed. The multipoint formulation leads to discrete expressions for the moments, linking moment-based and multipoint optimizations. The linearity of the sums in the discrete moments permits an easy evaluation of their gradients with respect to the design variables. Optimal sampling issues are analyzed and a procedure is proposed to quantify the confidence level on the robustness of the design. The proposed formulation is fully parallel and the time-to-solution is comparable to single-point situations. It is applied to three problems: an analytical least-square minimization problem, a shape optimization problem with a reduced-order model, and a full aircraft shape optimization 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. Today's industrial robust design mainly relies on reduced order modeling and intelligent sampling [1, 3, 4, 35] which either does not use high-fidelity simulations during design or uses lower accuracy than what would be affordable in a single-point optimization.

By robust design we mean a proposition which ensures similar performances over a given operation range. We will discuss the implication of this requirement on the moments of the performance functional. Our aim is to propose a methodology which permits to design a system having similar performances over a given range of its operating condition or functioning parameters. From a practical point of view, we would like this to be achieved modifying as less as possible an existing single-point optimization platform. In particular, we would like to avoid replacing the high-fidelity ingredients of the platform by low-complexity solvers. Finally, we would like the time-to-solution to remain comparable to the single-point situation. We will illustrate our approach on problems with increasing level of complexity and, in particular, on the optimization of the shape of an aircraft.

The literature on multipoint optimization is vast and exhaustive referencing is out of our scope. This formulation has been used, in particular, to extend single-point optimization to account for additional operating conditions. In this context,

Key words and phrases. mean, variance, skewness, kurtosis, adjoint-based parametric robust shape optimization, uncertainty.

aerodynamic shape optimization has been previously tackled by multipoint formulations together with high fidelity Reynolds Averaged Navier-Stokes (RANS) or Euler simulations [22, 37, 14].

Forward and backward uncertainty propagation [2, 5, 19] are obviously of great importance with a huge literature dedicated to uncertainty quantification (UQ). Forward propagation aims at defining, for instance, a probability density function (PDF) for a functional $j(\mathbf{x})$ knowing those of optimization variables \mathbf{x} [13, 15, 20]. This can be done, for instance, through Monte Carlo simulations or a separation between deterministic and stochastic features using Karhunen-Loeve theory (polynomial chaos theory belongs to this class) [10, 11, 35, 36]. Backward propagation aims at reducing models bias or calibrating models parameters knowing the PDF of j [6, 17, 34]. This can be seen as a minimization problem and Kalman filters [18] give, for instance, an elegant framework for this inversion assimilating the uncertainties on the observations.

In previous works, we showed how to quantify epistemic and aleatory uncertainties through geometric characterizations of global sensitivity spaces built using adjoint-based gradients of the functional available in existing optimization platforms [24]. In the presence of different modeling or solution methods, principal angles between these sensitivity subspaces permit to measure the deviation due to a change in the modeling [25]. We also showed how these ingredients can be used in the context of multipoint robust analysis of a system to define worst-case scenarios for its functioning. To this end we combined a multipoint search direction with the probabilistic features of the optimization variables through their quantiles [16, 27]. These ingredients permit to define the concept of directional uncertainty quantification and directional extreme scenarios (DES)[26]. Global sensitivity spaces can then be built for these extreme scenarios and the above geometric characteristics permit again to measure the impact of the variability of the variables of the problem. One major interest of this construction is to account for the variability of the variables in large dimension without a sampling of the space. One surprising consequence of this analysis is that beyond individual gradient accuracy (i.e. at each of the sampling point), what is important in multipoint problems is the global search space defined by the ensemble of the gradient vectors. This means that one might tolerate higher error levels in each of the gradient defined at the different sampling point than for a single-point optimization situation as what is important is for the global search space to remain unchanged. To go beyond directional extreme scenarios, we combined the previous ingredients with Ensemble Kalman Filters (EnKF) [18] which are efficient to account for the variability in the observations in inversion. The outcome is what we called Ensemble Directional Extreme Scenarios (EDES) which provides a more accurate and exhaustive sampling of the boundary of the uncertainty domain. We showed how to link the necessary ensemble size to the geometry of global sensitivity spaces in inversion [28].

Despite these approaches avoid the sampling of a large dimensional space and despite EnKF is, indeed an elegant solution for the inversion of uncertain data, the computation cost remains high and the procedures difficult to simply explain in engineering environments. To address this issue, we proposed a simplified low-complexity evaluation of the covariance matrix of the design variables using the mentioned extreme scenarios and an adjoint-based gradient of the functional [30].

This work follows what we started in [24] linking multipoint and moment-based optimizations. These are mostly used to address robustness issues in engineering because of their conceptual simplicity. The former can be efficiently formulated through the latter when only a few operating condition parameters describing the functioning of the system are present. This splitting between operating condition parameters and design variables is central to our discussion. The fact that only a few operating condition parameters are involved while the size of the vector of optimization variables can be very large is quite realistic. In this work we would like to go beyond monitoring the mean and the variance and see how much controlling the skewness and the kurtosis of a system performance would affect its design. This work therefore also addresses some aspects of optimization under uncertainty where the moments of the functional are computed after a multipoint formulation of the problem.

The paper starts with the formulation of robust parametric optimization problems and their solutions through moment-based and multi-point formulations. Section 3 presents the evaluation of the gradients of the moments and the construction of admissible search directions for a gradient-based optimizer. Section 4 shows how to use these ingredients in such algorithms. Then the paper presents the application of the approach to two model problems in sections 5 and 6 and to the optimization of a full aircraft shape in section 7.

2. ROBUST PARAMETRIC OPTIMIZATION

We are interested in a class of optimization problems where the cost function involves a few operating condition parameters α not considered as design variables:

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

The operating condition parameters therefore refer to the points in the multipoint formulation. \mathbf{x} is the vector of design variable belonging to \mathbf{O}_{ad} the optimization admissible domain. Typical situations of interest are with small n : a few operating conditions while the system can require several design variables.

This is a very general context and we visited it to address robustness issues in optimization with respect to \mathbf{x} and α [27, 25, 26, 30].

2.1. Moment-based optimization. We showed how to use multipoint optimization to address such optimization problem [24]. The aim is to remove the dependency in α during optimization. This can be done, for instance, minimizing a functional $J(\mathbf{x})$ encapsulating this dependency under a constraint on the second moment of j :

$$(2) \quad J = \mu = \mathbb{E}[j(x, \alpha)], \quad \text{such that } \sigma = \mathbb{E}[(j(x, \alpha) - \mu)^2] \leq \sigma_0,$$

where $\mathbb{E}[\cdot]$ is the expected value of j for the uncertain variable α .

In this work we would like to go beyond the first two moments and see how a control of the first four moments, for instance, of j would affect the design.

Going beyond first two moments is important when the PDF of j deviates from a pure Gaussian distribution. Indeed, even with interval-based (with uniform PDF) or Gaussian entries there is no reason the PDF of the solution of a simulation to remain uniform or Gaussian. In this work, we solely consider uniform PDF, but there is no limitation in applying the presented concepts to other PDFs.

The third and fourth moments, the skewness γ and the kurtosis κ , are defined as:

$$(3) \quad \gamma = \mathbb{E}\left[\left(\frac{j(x, \alpha) - \mu}{\sigma}\right)^3\right], \quad \kappa = \mathbb{E}\left[\left(\frac{j(x, \alpha) - \mu}{\sigma}\right)^4\right].$$

To ease their derivation with respect to \mathbf{x} we avoid their normalization by σ in the sequel. The skewness is an indication of the deviation from symmetry of the PDF and the kurtosis permits to quantify if it is tall and skinny or short and squat.

One can consider that a robust design should favor symmetry in the distribution which means lower absolute value of skewness. This is an assumption and other values than zero can be targeted for the absolute value of the skewness. However, the assumption of symmetry appears reasonable. Indeed, when driving a car on a straight line, one expects the car to have the same behavior for small and symmetrical disruptions in holding the steering wheel by the driver. In a Gaussian distribution we have $\gamma = 0$. Also, in a normal distribution the mean and median coincide and if a PDF is not too far from a normal distribution, the median will be near $\mu - \gamma\sigma/6$. Therefore, if $|\gamma| \rightarrow 0$ the PDF tends toward a normal distribution. This provides an inequality constraint on $|\gamma|$ as γ can be either positive or negative. For an unimodal PDF a reduction of the skewness comes when the mean and the mode of the distribution converge toward each other at given standard deviation.

Robust design means higher density near the mean which means higher kurtosis, but this is more subtle. Indeed, despite higher kurtosis means concentration of the probability mass around the mean, it could also imply thicker tails in the PDF. This means that more of the variance is the result of infrequent extreme deviations. We need therefore to define what we mean by robust design: acceptance of frequent modest deviations or acceptance of infrequent extreme ones. If operational security is a major concern the latter should be definitely avoided. We therefore consider that a reasonable requirement would be to have a design reducing the initial kurtosis value: $\kappa \leq \kappa_0$ together with a constraint on the variance σ .

In summary, we would like to consider these two constraints in addition to the necessity to control the deviation from the mean expressed by a constraint on the variance during optimization:

$$(4) \quad \min_{\mathbf{x} \in \mathbf{O}_{ad}} J(\mathbf{x}) = \mu(\mathbf{x}) \quad \text{such that} \\ C_1(\mathbf{x}) = \sigma(\mathbf{x}) - \sigma_0 \leq 0, \quad C_2(\mathbf{x}) = |\gamma(\mathbf{x})| - |\gamma_0| \leq 0, \quad C_3(\mathbf{x}) = \kappa(\mathbf{x}) - \kappa_0 \leq 0.$$

Equality constraints on the moments are particular cases of these. In cases, higher kurtosis are targeted (see discussion above), the last constraint becomes $C_3(\mathbf{x}) = -\kappa(\mathbf{x}) + \kappa_0 \leq 0$.

2.2. Discrete expressions. Monte Carlo simulations permit to recover these moments with an error decreasing as σ/\sqrt{M} with M the number of functional evaluations and this with a convergence rate independent of n . But, for small n , classical numerical integration over-performs Monte Carlo simulations in term of complexity based on the number of functional evaluations to recover the moments at a given accuracy. As we are interested by small values of n , this latter may therefore be preferred.

Both Monte Carlo trials and numerical integration lead to the introduction of weighted sums [24] over a M -point sampling \mathbf{I}_M of \mathbf{I} as estimators of the previous

moments (denoted with the same notation):

$$(5) \quad \mu = \sum_{\alpha_k \in \mathbf{I}_M} \omega_k j(\mathbf{x}, \alpha_k), \quad \sigma^2 = \sum_{\alpha_k \in \mathbf{I}_M} \omega_k (j(\mathbf{x}, \alpha_k) - \mu)^2,$$

$$(6) \quad \gamma = \sum_{\alpha_k \in \mathbf{I}_M} \omega_k (j(\mathbf{x}, \alpha_k) - \mu)^3, \quad \kappa = \sum_{\alpha_k \in \mathbf{I}_M} \omega_k (j(\mathbf{x}, \alpha_k) - \mu)^4.$$

A major difference between Monte Carlo and numerical integration is that in the former the sampling \mathbf{I}_M and the weights are chosen according to the PDF of α . In this paper we consider uniform PDF. We have therefore uniform sampling in each of the dimension of α and $\omega_k \sim 1/M$. Unbiased estimates use slightly different coefficients and also introduce corrections. For instance, $\omega_k = 1/(M-1)$ for σ , $\omega_k = M/((M-1)(M-2))$ for γ and so forth. However, because we are interested by the gradients of these quantities with respect to \mathbf{x} and gradient-based minimization algorithms, these changes will only have slight impacts on the amplitude of the gradient with no real incidence on the optimization history, especially if optimal descent step sizes are used. This discussion also shows that moment-based and multipoint optimization can be seen as a whole. If different PDFs are targeted than uniform, we need to redefine the sampling \mathbf{I}_M and the weights ω_k accordingly.

3. GRADIENTS AND ADMISSIBLE SEARCH DIRECTIONS

The linearity in the sums permits to access to the gradients of the moments with respect to the optimization variables \mathbf{x} from the gradient of the functional at the sampling point α_k (we recall that we took $\sigma = 1$ in Equation (3) to ease this linearization):

$$(7) \quad \begin{aligned} \nabla_{\mathbf{x}} \mu &= \sum_{\alpha_k \in \mathbf{I}_M} \omega_k \nabla_{\mathbf{x}} j(\mathbf{x}, \alpha_k), \\ \nabla_{\mathbf{x}} \sigma^2 &= \sum_{\alpha_k \in \mathbf{I}_M} 2\omega_k (j(\mathbf{x}, \alpha_k) - \mu) \nabla_{\mathbf{x}} j(\mathbf{x}, \alpha_k), \\ \nabla_{\mathbf{x}} \gamma &= \sum_{\alpha_k \in \mathbf{I}_M} 3\omega_k (j(\mathbf{x}, \alpha_k) - \mu)^2 \nabla_{\mathbf{x}} j(\mathbf{x}, \alpha_k), \\ \nabla_{\mathbf{x}} \kappa &= \sum_{\alpha_k \in \mathbf{I}_M} 4\omega_k (j(\mathbf{x}, \alpha_k) - \mu)^3 \nabla_{\mathbf{x}} j(\mathbf{x}, \alpha_k). \end{aligned}$$

To account for the non-differentiability of the constraint on γ we consider the following definition for $\nabla_{\mathbf{x}} |\gamma|$:

$$(8) \quad \begin{aligned} \nabla_{\mathbf{x}} |\gamma| &= \nabla_{\mathbf{x}} \gamma \quad \text{if } \varepsilon < \gamma \\ &= -\nabla_{\mathbf{x}} \gamma \quad \text{if } \gamma < -\varepsilon \\ &= 0 \quad \text{if } |\gamma| \leq \varepsilon, \end{aligned}$$

where $\varepsilon = \gamma_0/10$.

Knowing the individual gradients $\nabla_{\mathbf{x}} j(\mathbf{x}, \alpha_k)$ at sampling points α_k brings interesting information on the geometry of the problem. Indeed, they can be used to build a vector space and the analysis of this space provides important information on the complexity of the problem [24, 25, 26]. In the sequel we will use this information to quantify our confidence on the quality and sufficiency of the sampling \mathbf{I}_M . In the same way, analyzing the vector subspace generated by the gradients

of the moments permits to build admissible descent directions for a gradient based optimizer generating admissible minimizing sequences.

3.1. Sensitivity spaces. Let us introduce two vector spaces S_M and s_M .

$\nabla_{\mathbf{x}}\mu, \nabla_{\mathbf{x}}\sigma^2, \nabla_{\mathbf{x}}\gamma$ and $\nabla_{\mathbf{x}}\kappa$ are four vectors in $S_M = \text{Span}\{\nabla_{\mathbf{x}}j(\mathbf{x}, \alpha_k), \alpha_k \in \mathbf{I}_M\} \subset \mathbb{R}^N$. In applications of interest N is large. However, we showed that often $q = \dim(S_M) \ll N$ [25, 24, 26]. The evaluation of the dimension of S_M provides an a posteriori indicator and gives confidence bounds on the choice of the sampling size M which should be clearly larger than q but still much smaller than N . We show an example of such analysis in section 7.2.1.

We also consider the subspace $s_M = \text{Span}\{\nabla_{\mathbf{x}}C_{i=1,\dots,3}\} \subset \mathbb{R}^3 \subset \mathbb{R}^N$. Obviously $p = \dim(s_M) \leq 3$. Let us denote by $\{\mathbf{q}_{i=1,\dots,p}\}$ an orthonormal basis for s_M obtained, for instance, orthonormalizing the three gradient vectors by the Gram-Schmidt procedure. The gradients G of the constraints can therefore be expressed as linear combination of q_i : $G = (\nabla_{\mathbf{x}}C_{i=1,\dots,3}) = P^{-1}(\mathbf{q}_{i=1,\dots,p})$ with P being the matrix expressing the coordinates of \mathbf{q} in G .

3.2. Admissible search direction with equality constraints. Let us start building a descent direction d for problem (4) in the presence of equality constraints:

$$(9) \quad \min_{\mathbf{x} \in \mathbf{O}_{ad}} J(\mathbf{x}), \quad \text{such that} \quad C(\mathbf{x}) = (C_1(\mathbf{x}), C_2(\mathbf{x}), C_3(\mathbf{x})) = 0.$$

At first order d needs to be orthogonal to s_M .

Hence, using the local orthonormal basis $\{\mathbf{q}_{i=1,\dots,p}\}$, we consider d given by:

$$(10) \quad d = \nabla_{\mathbf{x}}\mu - \sum_{i=1}^p \langle \nabla_{\mathbf{x}}\mu, \mathbf{q}_i \rangle \mathbf{q}_i.$$

Denoting by Π the matrix of the projection operator $\langle \nabla_{\mathbf{x}}\mu, q \rangle$ we have:

$$(11) \quad d = \nabla_{\mathbf{x}}\mu - (\Pi P G)^t P G = \nabla_{\mathbf{x}}\mu - (G^t P^t \Pi P)^t G = \nabla_{\mathbf{x}}\mu + \Lambda^t G,$$

with $\Lambda^t = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$.

We have $d \rightarrow 0$ with the descent iterations converging. Stationarity in d therefore realizes the first order optimality condition for the Lagrangian $L = J + \Lambda^t C$.

3.3. Admissible search direction with inequality constraints. The solution of our minimization problem needs to verify the first order KKT conditions [33]. But, the optimality condition for the Lagrangian will involve only positive Lagrange multipliers: $\Lambda \in \mathbb{R}_+^3$ and $\nabla_{\mathbf{x}}L = \nabla_{\mathbf{x}}J + \Lambda^t \nabla_{\mathbf{x}}C = 0$ with the complementarity condition $\Lambda^t C = 0$ meaning that $\lambda_i = 0$ if $C_i \leq 0$ and $\lambda_i > 0$ if $C_i = 0$ (i.e. C_i is an active constraint). To define d we follow what put in place for the equality constraints, but only considering active constraints gradients in the definition of s_M which is not anymore a subspace but a convex cone:

$$(12) \quad s_M = \{\mathbf{x} \mid \mathbf{x} = \sum_{i=1}^3 \beta_i \nabla_{\mathbf{x}}C_i, \beta_i > 0 \mid C_i = 0\} \subset \mathbb{R}^3 \subset \mathbb{R}^N.$$

At the solution, $\nabla_{\mathbf{x}}J$ is orthogonal to this cone. Before working on the cone, let us start defining a local orthonormal basis $\{\tilde{\mathbf{q}}_{i=1,\dots,p}\}$ for \tilde{s}_M from Equation (12) but with $\beta_i \in \mathbb{R}$. This is therefore a subspace and the basis can be defined as previously with $p = \dim(s_M)$. Now, $\mathbf{q}_i = \pm \tilde{\mathbf{q}}_i$ and the sign chosen such that $\langle \mathbf{q}_{i=1,\dots,p}, \nabla_{\mathbf{x}}C_j \rangle \geq 0$, if $C_j = 0$ for $j = 1, \dots, 3$ (i.e. pointing inside the cone). Here, $\{\mathbf{q}_{i=1,\dots,p}\}$ are therefore

the generators of the cone s_M deduced from a basis of \tilde{s}_M . If the generators cannot be defined, the problem is found having no solution as at least two of the constraints are incompatible with their gradients parallel and pointing in opposite directions. These generators permit to define the admissible search direction d from Equation (10) but taking into account that we only remove the non admissible contribution:

$$(13) \quad d = \nabla_{\mathbf{x}}\mu - \sum_{i=1}^p \chi_i < \mathbf{q}_i, \nabla_{\mathbf{x}}\mu > \mathbf{q}_i,$$

with $\chi_i = 0$ if $< \mathbf{q}_i, \nabla_{\mathbf{x}}\mu > \geq 0$ and $\chi_i = 1$ if $< \mathbf{q}_i, \nabla_{\mathbf{x}}\mu > < 0$.

4. A MULTIPOINT DESCENT ALGORITHM

We consider the following iterative descent algorithm for our constrained minimization problem involving a direct simulation chain linking the independent variables and parameters (\mathbf{x}, α) to the state U solution of a state equation $F(U(q(\mathbf{x}), \alpha)) = 0$ and to a functional j is:

$$\left\{ \begin{array}{l} \text{Given } \mathbf{x}_0, 0 < \rho, \mathbf{I}_M, p_{max} \text{ and a stopping criteria,} \\ \text{optimization iterations } p = 1, \dots, p_{max} \\ \left\{ \begin{array}{l} 1-M \text{ parallel state equation solutions } F(U(q(\mathbf{x}_p), \alpha_l)) = 0, \alpha_l \in \mathbf{I}_M, \\ 2-M \text{ parallel evaluations of } j(\mathbf{x}_p, \alpha_l), \alpha_l \in \mathbf{I}_M, \\ 3-M \text{ parallel solutions of the adjoint state } V \text{ equation:} \\ \quad V^t F_U(U(q(\mathbf{x}_p), \alpha_l)) = j_U^t, \alpha_l \in \mathbf{I}_M, \\ 4-M \text{ parallel evaluations of } \nabla_{\mathbf{x}} j(\mathbf{x}_p, \alpha_l) = j_{\mathbf{x}} + (V^t F_{\mathbf{x}})^t, \alpha_l \in \mathbf{I}_M, \\ 5\text{-define } d \text{ the descent direction using Equation (13),} \\ 6\text{-minimization using } d: \text{ (e.g. } \mathbf{x}_{p+1} = \mathbf{x}_p - \rho d), \\ \text{Stop if a given stopping criteria is achieved.} \end{array} \right. \end{array} \right.$$

In multi-criteria problems steps 2, 3 and 4 also include the treatment of the different criteria. This introduces a different definition of the descent direction d as shown in section 7.

Despite the natural presence of parallelism due to the M independent evaluations of the state, functional and its gradient, computational complexity remains an issue. We have shown previously how to reduce this effort optimizing the sampling size [24] together with the use of incomplete sensitivity concept in the evaluation of the gradients which permits to avoid the solution of the M adjoint equations [31].

5. A FIRST EXAMPLE

Let us illustrate our ingredients on a simple example. The functional involves a least-square minimization:

$$(14) \quad j(\mathbf{x}, \alpha) = \frac{1}{2} \sum_{i=1}^N (x_i - \alpha)^2, \quad -0.5 \leq \alpha \leq 0.5, \quad -5 \leq x_{i=1, \dots, N} \leq 5, \quad N = 40.$$

Let us solve problem (4) with a gradient method. Here, α is a scalar and for a given α the optimality condition for $j(\mathbf{x}, \alpha)$ gives obviously $\mathbf{x}^* = (\alpha, \dots, \alpha) \in \mathbb{R}^N$. This would be the solution of a single-point optimization. Here α has a uniform PDF

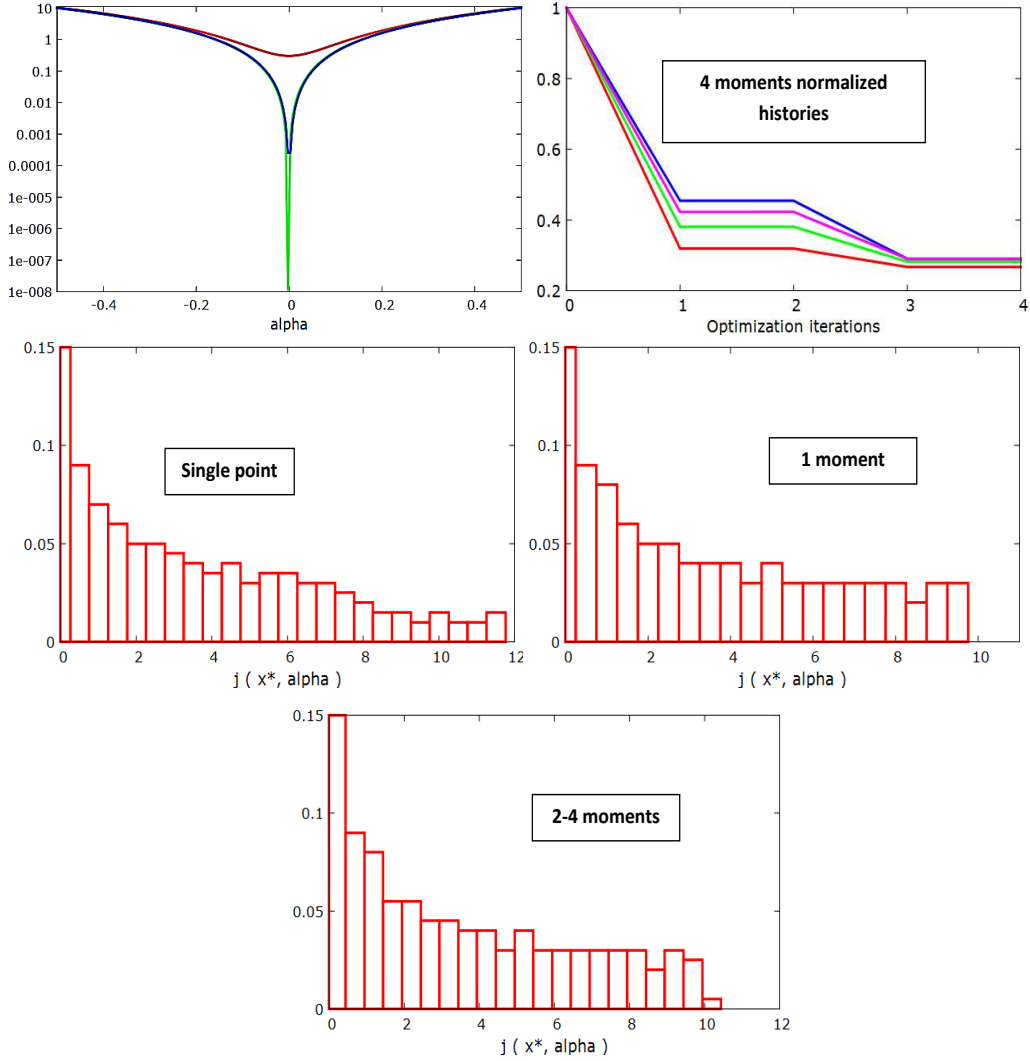


FIGURE 1. Functional (14): Upper-left: $j(\mathbf{x}^*, \alpha)$ given by Equation (14) vs. α for single-point optimization, mean-based and with moment constraints. Upper-right: normalized moments histories with moment constraints. Lower: histograms of $j(\mathbf{x}^*, \alpha_{k=1,\dots,M})$ for a single-point minimization and when controlling either one or more moments.

and we use a uniform sampling \mathbf{I}_M with $M = 40$. Minimizing the mean μ without any constraint on the other moments still produces a non robust optimum as the performance has large variability over the range of the operating condition α . On the other hand, asking for all four moments to decrease produces a much flatter functional over the range of α . Figure 1 shows the histories during optimization iterations of (μ/μ_0) , $(\sigma/\sigma_0)^{1/2}$, $(\gamma/\gamma_0)^{1/3}$ and $(\kappa/\kappa_0)^{1/4}$. The figure also shows the histograms of $j(\mathbf{x}^*, \alpha_{k=1,\dots,M})$ for these minimizations. In particular, one sees how controlling more moments affects the left peak. In this example controlling the first two moments gives nearly the same answer than with the first four moments. But, we will see that this is not always the case.

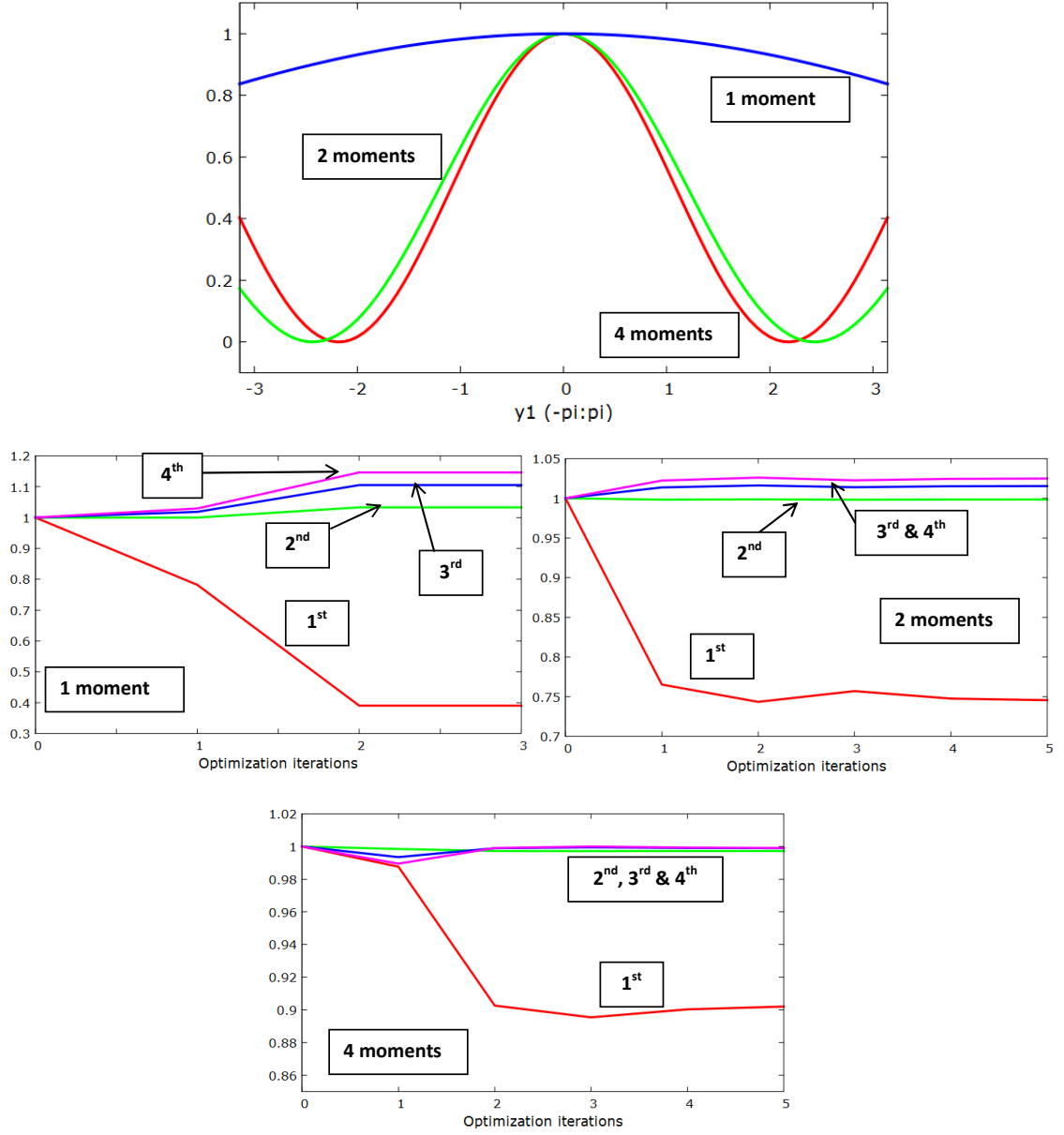


FIGURE 2. Robust shape optimization for functional (16). Optimal shapes obtained when controlling one or more moments and convergence histories for the first four moments.

6. SHAPE OPTIMIZATION WITH A REDUCED-ORDER MODEL

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

$$\begin{aligned}
 (15) \quad p(\Gamma(\mathbf{x}), \alpha) &= \frac{1}{2} \rho_{\infty} \|\mathbf{u}_{\infty}(\alpha)\|^2 \left(\frac{\mathbf{u}_{\infty}(\alpha)}{\|\mathbf{u}_{\infty}(\alpha)\|} \cdot \mathbf{n}(\Gamma(\mathbf{x})) \right)^2 \\
 &= \frac{1}{2} \rho_{\infty} (\mathbf{u}_{\infty}(\alpha) \cdot \mathbf{n}(\Gamma(\mathbf{x})))^2,
 \end{aligned}$$

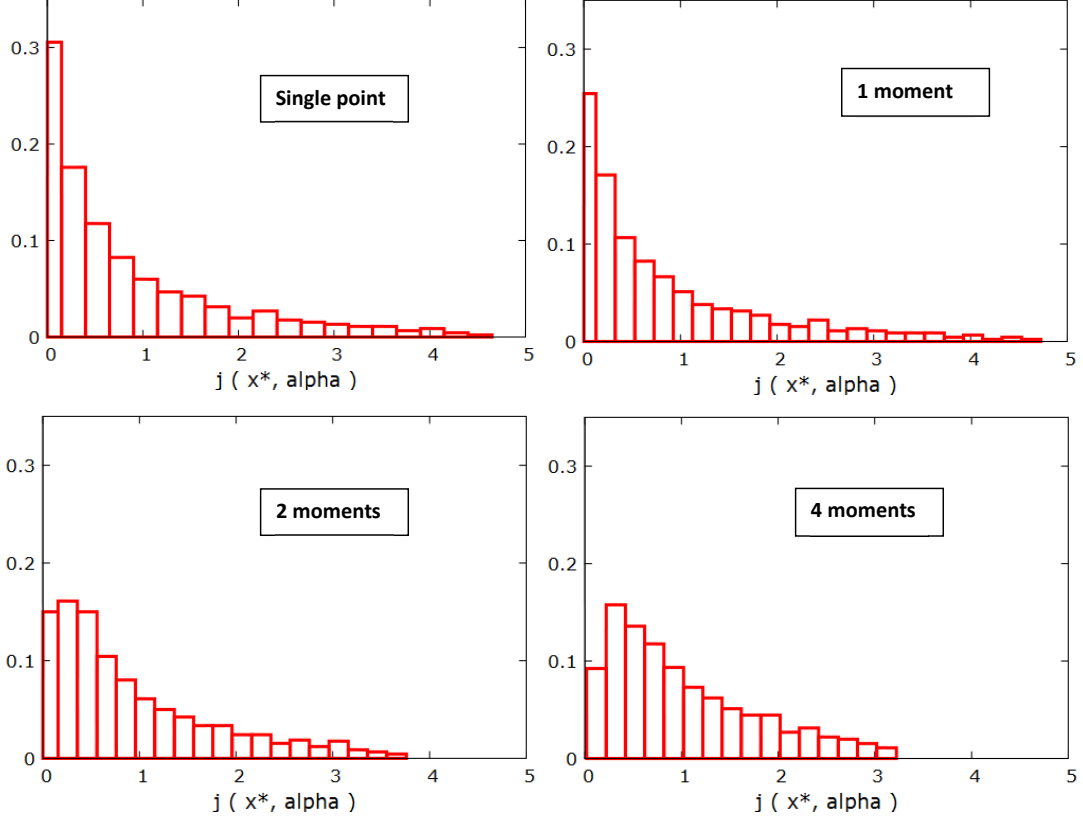


FIGURE 3. Histograms of $j(x^*, \alpha_{k=1,\dots,M})$ for the shapes given in figure 2.

where subscript ∞ denotes inflow quantity for the density and velocity and \mathbf{n} the local outward normal to the shape. This relation gives satisfactory prediction of the pressure distribution over blunt shapes (such as nose cone shapes) in a low-speed flow.

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

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

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 incidence (θ) defined by $\alpha = (U_{in}, \theta)$. Using the notations above, we have:

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

As in the previous example, we use a gradient method to solve problem (4). In a single-point optimization (i.e. at a given α) the shape will be flat and oriented in order for the normal \mathbf{n} to the shape to realize:

$$\mathbf{u}_\infty \cdot \mathbf{n} = \sqrt{2 \frac{p_{des}}{\rho_\infty}}.$$

However, while optimal for a single-point situation, a particular flat shape is not anymore optimal for a multipoint formulation based on a sampling of the range of $\alpha = (U_{in}, \theta) \in [1\text{cm/s}, 1.6\text{cm/s}] \times [-1.5^\circ, 1.5^\circ]$. We consider a 30×30 sampling of this two dimensional operating parameters range (i.e. the sampling size is $M = 900$). This means that at each iteration of optimization we need to compute 900 states and gradients of the functional at these sampling points.

Minimizing the first moment alone produces an arc shape as shown in figure 2. Now, let us consider this minimization together with equality constraints on higher moments. We see that, unlike in the previous example, going from 2 to 4 moments has a significant impact on the optimal shape. Figure 3 shows the histograms of $j(x^*, \alpha_{k=1,\dots,M})$ for these minimizations. One sees how controlling more moments reduces the dispersion of j . A fully robust design will only have one bar in its histogram. In that regard, compared to the single-point design, the situation clearly improves with two moments and more. With solely the first moment the robustness is slightly better as the weight of the tail in the histogram decreases. But, the improvement is not enough visible. This is expected as the 1-moment shape shown in figure 2 is close to the single-point flat shape. The evolution of the moments during optimization also shows the impact of adding two extra constraints requiring a conservation of the third and fourth moments. We will recover this behavior in the next example for a complex shape optimization problem.

7. FULL AIRCRAFT SHAPE OPTIMIZATION

Let us now apply the concepts presented in the previous sections to the optimization of the shape of an aircraft in transonic cruise condition. We have previously studied how to approach robust shape design for this configuration to account for the different sources of variability (e.g. due to a change in the weight of the aircraft during the flight because of fuel consumption or due to variability in the flight conditions) [24, 25, 26]. In all these works, the aim has been to upgrade existing deterministic single-point shape optimization platforms in order to minimize the necessary extra coding effort.

One important source of variability is with the flight conditions. For instance, when the aircraft cruises against transverse winds which are very common. In this case, the operating condition parameter α is the sideslip angle (i.e. $n = 1$) inducing fully 3D effects on the flow around the plane making necessary the consideration of a full aircraft during the design. Usually aircraft are designed for a range of angle of incidence and this 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. It is therefore necessary to reduce the sensitivity of the design with respect to this operating condition parameter.

Because the airplane geometry is symmetric spanwise, it is not necessary to consider a symmetric range for the transverse wind but, as already mentioned, we need to consider the whole aircraft as there is no spanwise symmetry in the flow for non zero sideslip angles. In the sequel we consider an aircraft flying at a Mach number of 0.8 and zero inflow incidence but with the sideslip angle α in $\mathbf{I} = [0, 10^\circ]$. These parameters fully describe a 3D inviscid flow around the aircraft.

7.1. Single-point shape optimization platform. We work in the framework of an existing shape optimization platform. We use, in particular, several of its

simulation codes for the shape parameterization and deformation, for the mesh deformation, for the flow calculations around the aircraft and for the shape adjoint sensitivity analysis of aerodynamic coefficients. This is a very standard and generic situation and one shall consider these as black-boxes. Also, the discussion remains valid if more sophisticated multi-physics scenarios need to be considered [23].

Let us briefly recall our direct dependency chain linking independent variables (α, \mathbf{x}) to the dependent variables $(q(\mathbf{x}), U(\alpha, \mathbf{x}))$ describing geometrical entities and state variables and to the cost function (here the drag coefficient C_d) and to the constraints $c_{i=1,\dots,3}$:

$$(17) \quad (\alpha, \mathbf{x}) \rightarrow \begin{cases} q(\mathbf{x}), \\ U(\alpha, q(\mathbf{x})), \\ (C_d, c_{i=1,\dots,3})(\alpha, \mathbf{x}, q(\mathbf{x}), U(\alpha, q(\mathbf{x}))). \end{cases}$$

It is important to identify all dependencies in order for the sensitivity analysis to be complete, especially when the operating conditions are not anymore single valued. The functional and constraints will be described in section 7.1.3.

7.1.1. Shape parameterization and geometrical entities. In relation (17) \mathbf{x} denotes a CAD-free parameterization [31, 32] which does not require a priori local regularity assumptions on the shape as it is implicitly the case in Computer Aided Design (CAD) based shape definitions. More precisely, \mathbf{x} represents shape deformations along the normal to the triangular faces of the surface mesh as shown in figure 4. For the problem discussed here this search space has a dimension $N \sim 5000$. This parameterization receives different denominations and belongs to the same class than node-based or free-form shape definitions. In all these approaches the regularity of the deformation needs to be monitored [21, 31]. This parameterization is intermediate in term of generality between CAD definitions of a shape and fully free topological optimization choices where both the regularity and topology of the shape are free. Examples of shape deformation produced by our optimization procedure for different regularity requirements are shown in figure 5.

Need for regularity control comes from the fact that, unlike with a CAD definition, the shape $\partial\Omega$ of an object Ω and a gradient-based deformation of $\partial\Omega$ do not belong to the same function space in terms of regularity and, actually, the second is always less regular [30, 31, 32].

This can be illustrated on a simple example with $J(\mathbf{x}) = \|A\mathbf{x} - b\|^2$ taking $\mathbf{x} \in H^1(\partial\Omega)$, $A\mathbf{x}$ and b in $L^2(\partial\Omega)$. The gradient $J'_\mathbf{x} = 2A^T(A\mathbf{x} - b)$ belongs to $H^{-1}(\partial\Omega)$. Therefore, any variation along $J'_\mathbf{x}$ will have less regularity than \mathbf{x} : $\delta\mathbf{x} = -\rho J'_\mathbf{x} = -\rho(2(A\mathbf{x} - b)A) \in H^{-1}(\partial\Omega)$. We therefore need to project (or filter or smooth) into $H^1(\partial\Omega)$. Now, suppose the shape is described in a finite dimensional parameter space, as for instance with a polynomial definition of a surface (this is like a CAD parameterization). When we consider as control parameters the coefficients of the polynomial, changes in those do not change the regularity as the new shape will always belong to the same polynomial space. Sobolev inclusions give the key for the choice of the regularity operator with the CAD-free parameterization [31]. In our case, because we are using a piecewise linear discretization, a second-order elliptic system is sufficient. More precisely, we use a local Laplace-Beltrami operator over the shape [31].

This capacity to monitor the regularity of the shape is also interesting as often the optimal solution is not reachable by the current CAD parameterization of the shape. Hence, after an optimization with the CAD-free parameterization and using different level admissible regularity for the shape, one can decide which realization is more suitable and also whether it is interesting or not to enrich the current CAD definition of the shape.

To keep the discussion focused on moments control, in the sequel we only consider one level of regularity during optimization corresponding to the lower-right picture in figure 5.

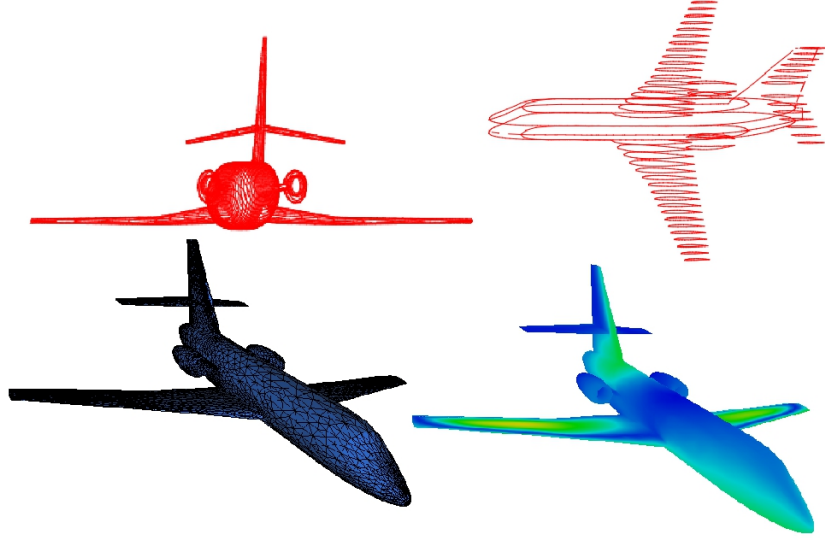


FIGURE 4. CAD-free shape parameterization (lower-left) and by-section definitions (upper) of the shape for geometric constraints enforcement. Lower-right shows a typical distribution of $\nabla_{\mathbf{x}}C_d - < \nabla_{\mathbf{x}}C_d, \pi > \pi$ described in section 7.1.4 for this CAD-free parameterization.

$q(\mathbf{x})$ denotes the auxiliary unstructured mesh related geometrical quantities (surfaces, volumes, normals, etc). When the shape is modified, this change must be propagated through the mesh keeping it admissible and we need to recalculate all related geometrical quantities. Admissible and positive mesh deformation is achieved by a 3D torsional spring analogy method [9].

7.1.2. Flow solver. In relation (17) $U(\alpha, q(\mathbf{x})) = (\rho, \rho\vec{u}, \rho E)^t$ denotes the flow variables in conservation form solution of the Euler equations where, T being the temperature, the total energy is given by $E = C_v T + \|\vec{u}\|^2/2$ and the pressure by the state law $p = \rho R T$ with R the perfect gas constant.

The details of the implementation of the flow solver are available in [31]. It is based on a finite volume Galerkin method on unstructured tetrahedral meshes [8]. Other choices are possible for the flow solver and the literature on numerical methods for compressible flows is huge. This is not central to our discussion. We target steady solutions and use time marching with local time steps to reach these. The

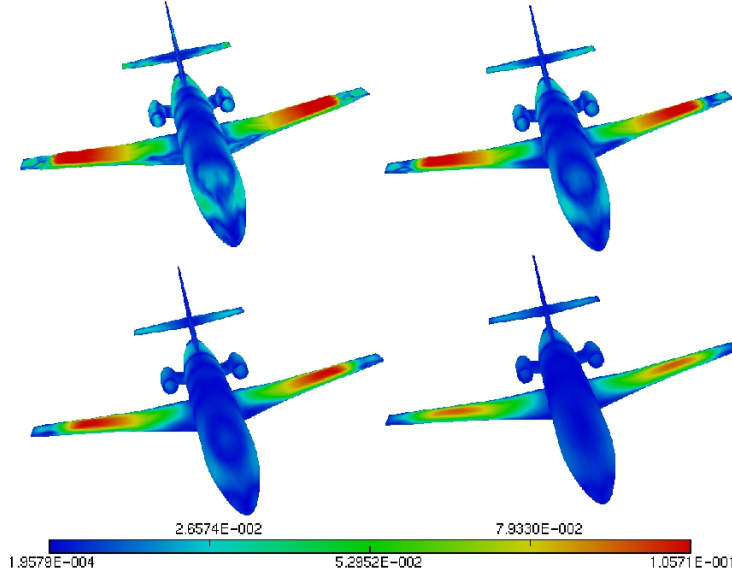


FIGURE 5. Regularity control in CAD-free shape parameterization: examples of shape deformation produced by our optimization procedure for different regularity requirements. These are isovals of the norm of the vector of shape deformation along the normal to the surface.

time integration procedure is explicit and is based on a low-storage Runge-Kutta (RK) scheme. To improve computational efficiency we only use partial convergence for the state equations. In particular, the sufficient level of convergence retained is when the flow solver iterations only modify the third digits in the aerodynamic coefficients. This is achieved with about 100 RK iterations starting from a uniform solution distribution. During optimization a new calculation for a new shape is always started from the previously available solution making us to proceed with typically only 20 RK new iterations [25, 26, 30].

The mesh used here has around one million elements which is obviously insufficient to reach full mesh resolution for the flow around an aircraft, even in a context of inviscid modeling. It is however important that the approach uses the ingredients of a generic high fidelity platform and does not remove or simplify any of its ingredients as often it is the case in uncertainty quantification procedures using reduced order models.

Also, as mentioned in the introduction, we should consider that in practice our modeling capability and our computational resources will always be limited. Suppose we have access to different models, numerical methods or levels of refinement. Each will provide a couple of vector spaces S_M and s_m introduced in section 3.1. Analysis of their respective dimensions and principal angles between corresponding spaces [26, 25] permits to estimate if an increase in the modeling complexity or numerical accuracy has an impact on the search space: if unchanged, an increase in the complexity is useless. And we can go further assuming the computed state (here the pressure over the shape) as uncertain with different levels of uncertainty over different regions of the shape. The assumption can be based on what is expected

from the modeling or observed (experimental or in flight). It is, indeed, well known that for a civil transport the flow distribution is quite stable in the cockpit and over the first and business class sitting area where the flow is nearly potential. On the other hand, flow variability increases spanwise (easy to see from wings tips oscillating) and also toward the tail of the aircraft (flying coach once makes this easy to understand). These are due, among others, to separated turbulent flows instabilities and fluid-structure interactions. These remarks permit to define a rough geometric distribution of expected state variability distribution. In [30] we showed how to use this information together with the gradients of the functional to build the covariance matrix of \mathbf{x} after optimization. The diagonal of this matrix gives, in particular, the uncertainty we have on the design due to the uncertainty or lack of modeling and state resolution.

7.1.3. Optimization problem. Aircraft performance analysis concerns its payload and range. These are directly linked to the aerodynamic coefficients of the aircraft called the lift (conditioning the payload) and drag (conditioning the fuel consumption) coefficients. The drag C_d coefficient is given by:

(18)

$$C_d(\alpha, \mathbf{x}, q(\mathbf{x}), U(\alpha, q(\mathbf{x}))) = \frac{2}{\mathbf{S}\rho_\infty(\alpha)\|\mathbf{u}_\infty(\alpha)\|^2} \int_{\text{shape}(\mathbf{x})} p(\alpha, q(\mathbf{x}))(\mathbf{u}_\infty \cdot \mathbf{n}(q(\mathbf{x})))d\gamma,$$

where subscript ∞ indicates inflow conditions and \mathbf{S} a constant reference area (taken here as one square meter). The lift coefficient C_l is evaluated with formula (18) where \mathbf{u}_∞ is replaced by \mathbf{u}_∞^\perp in the boundary integral.

The lift coefficient often appears through an inequality:

$$c_1(\alpha, \mathbf{x}, q(\mathbf{x}), U(\alpha, q(\mathbf{x}))) = C_l^{\text{target}} - C_l(\alpha, \mathbf{x}, q(\mathbf{x}), U(\alpha, q(\mathbf{x}))) \leq 0,$$

or equality constraint

$$c_1(\alpha, \mathbf{x}, q(\mathbf{x}), U(\alpha, q(\mathbf{x}))) = |C_l^{\text{target}} - C_l(\alpha, \mathbf{x}, q(\mathbf{x}), U(\alpha, q(\mathbf{x})))| = 0,$$

with C_l^{target} a target performance. Let us consider this second situation.

Structural efficiency and necessity of useful free volume also implies the consideration of geometric criteria such as a constraint on the volume V of the aircraft or its by-section definition. As for the lift coefficient, this gives a constraint of the form:

$$c_2(q(\mathbf{x})) = |V^{\text{target}} - V(q(\mathbf{x}))| = 0.$$

The volume of an object Ω (here the aircraft) is expressed through the boundary integral formula: $V = \int_\Omega 1 = \int_\Omega \frac{1}{3} \nabla \cdot (\mathbf{X}) = \int_{\partial\Omega} \mathbf{X} \cdot \mathbf{n}$, where $\mathbf{X} = (x_1, x_2, x_3)^t$ is the local coordinate over the shape.

The last geometric constraint concerns the local wing by-section thickness which is prescribed. We define by-section definitions of the shape as shown in figure 4. In this construction the number of sections n_s is free (here $n_s = 50$) and can be adapted to account for the complexity of the geometry. Each node in the parameterization is associated to a section Σ_i , and for each section, we define the maximum thickness Δ_i . This last operation requires the projection of the upper-surface nodes over the lower surface for each section. This constraint is expressed as:

$$c_3(q(\mathbf{x})) = \sum_{i=1}^{n_s} |\Delta_i(q(\mathbf{x})) - \Delta_i^{\text{target}}| = 0.$$

An alternative solution which is much simpler to implement is to only enforce a local volume constraint in each section Σ_i using the volume formula above: $V(\Sigma) = \int_{\Sigma_i} 1 = \int_{\Omega} 1 \chi_{\Sigma_i} = \int_{\Omega} \frac{1}{3} \nabla \cdot (\mathbf{X}) \chi_{\Sigma_i} = \int_{\partial\Omega} \mathbf{X} \cdot \mathbf{n} \chi_{\partial\Sigma_i}$, where χ is an indicator function ($\chi = 1$ if the point is in section Σ_i and $\chi = 0$ otherwise). Testing if a point is in Σ_i is easy and only requires an interval-based coordinate check, spanwise in this situation.

In summary, we consider the following optimization problem involving aerodynamic coefficients and geometric constraints:

$$\begin{aligned}
 (19) \quad & \underset{\mathbf{x}}{\text{minimize}} && C_d(\alpha, \mathbf{x}, q(\mathbf{x}), U(\alpha, q(\mathbf{x}))) \\
 & \text{subject to} && c_1(\alpha, \mathbf{x}, q(\mathbf{x}), U(\alpha, q(\mathbf{x}))) = 0, \\
 & && c_2(q(\mathbf{x})) = 0, \\
 & && c_3(q(\mathbf{x})) = 0, \\
 & && F(U(q(\mathbf{x}), \alpha)) = 0.
 \end{aligned}$$

7.1.4. Gradient of the functional and constraints. The minimization algorithm we consider uses the gradients of functional and constraints with respect to the shape. In our approach, all the sensitivities are computed by automatic differentiation (AD) in reverse mode using **tapenade** [12]. Details on how AD in direct and reverse modes works and, in particular, how **tapenade** generates a computer program for $\nabla_{\mathbf{x}} j$ from the program computing $j(\mathbf{x})$ in similar situations are given in [31, 32, 30].

The constraints can be accounted for by introducing penalty or weighting terms in the cost function: $j = C_d + \sum_{i=1,3} a_i c_i$, $a_i \in \mathbb{R}^+$.

The weighting approach can however be avoided using the formulation presented in section 3 to account for the constraints C_i on the second, third and fourth moments of the functional. More precisely, we consider a locally admissible gradient orthogonal to $\mathbf{A} = \text{Span}(\nabla_{\mathbf{x}} c_i, i = 1, \dots, 3)$ with $\dim(\mathbf{A}) \leq 3$. Let us denote by π an orthonormal basis of this subspace obtained by the Gram-Schmidt procedure applied to the gradients of the constraints. The admissible search direction is given by:

$$(20) \quad \delta_k = \delta(\mathbf{x}, \alpha_k) = \nabla_{\mathbf{x}} C_d - \langle \nabla_{\mathbf{x}} C_d, \pi \rangle \pi,$$

where \langle, \rangle indicates the scalar product over subspace \mathbf{A} . This is therefore similar to the construction given in Equation (10) where $\pi = \{\mathbf{q}_{i=1, \dots, 3}\}$ and with the constraints C_i replaced by c_i . In the presence of inequality constraints $c_i \leq 0$ instead of equality we proceed as in section 3.3 and build the admissible direction based on the KKT optimality conditions.

The constructions above require $\nabla_{\mathbf{x}} C_d$, $\nabla_{\mathbf{x}} C_l$, $\nabla_{\mathbf{x}} V$ and $\nabla_{\mathbf{x}} \Delta_i$. The two former require the adjoint of the state equation and we take advantage of the capability for multi-right-hand-side adjoint calculation of **tapenade** in reverse mode to access these gradients without necessitating the solution of two separate adjoint problems. Our direct Euler code uses time marching to the steady solution with local time steps. An optimization of the reverse mode of AD comes from the fact that, our situations of interest being stationary in time, there is no need to store the forward intermediate states before backward integration: only the final steady state is enough [7, 29, 31].

In the sequel the developments of section 3 will be followed with the gradient $\nabla_{\mathbf{x}} j$ replaced by the search direction δ_k .

7.2. Multipoint optimizations. Let us now use the ingredients above in our moment control problem described in section 2. We require all moments to decrease. However, as described in section 2.1, other strategies can be adopted, especially for the skewness and the kurtosis. In the sequel we consider a $M = 100$ points sampling of \mathbf{I} . From a complexity point of view, this means that at each iteration of optimization, we need to solve 100 direct and adjoint state equations to access the state $U(\mathbf{x}, \alpha_k)$ and the gradients $\nabla_x J(\mathbf{x}, \alpha_k)$ for the different $\alpha_k \in \mathbf{I}_{100}$. This represents the major cost of the optimization (see table 1).

7.2.1. The search space S_M . Let us analyze the global search space S_M introduced in section 3.1 and generated by the individual search directions $\delta_{k=1,\dots,M}$ at the sampling \mathbf{I}_M of the operating condition parameter range: $S_M = \text{Span}\{\delta(\mathbf{x}, \alpha_k), \alpha_k \in \mathbf{I}_M\} \subset \mathbb{R}^N$. In particular, we are interested by the dimension of this space. As shown in [24], this analysis will give an a posteriori confidence bound on M . Figure 6 shows the outcome of the Gram-Schmidt orthonormalization of the vectors in $\{\delta_{k=1,\dots,100}\}$ at each iteration of optimization. The subspace \mathbf{S}_{100} is not free and can always be generated by a subset of vectors with a maximum of 35 independent directions over the 100 involved. This maximum represents a worse case in term of sensitivity as in most situations the space can be generated with less than 15 independent vectors. This provides some confidence on the chosen sampling size.

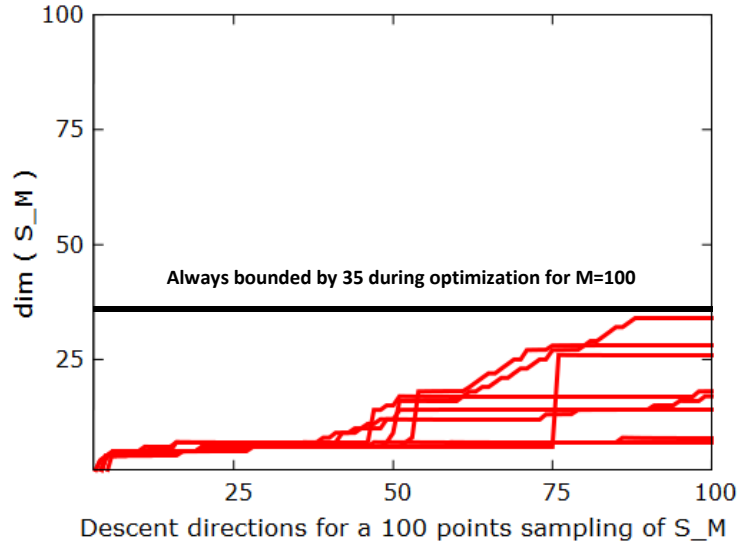


FIGURE 6. Histories of Gram-Schmidt orthonormalization of $\{\nabla_x j(x, \alpha_k), \alpha_k \in \mathbf{I}_{100}\}$ during optimization. The dimension of the global search space always remains below 35 to be compared to $M = 100$.

7.2.2. Comparing the four strategies. Figure 7 shows final shape deformations for our four strategies minimizing zero (i.e. single-point optimization), one, two and four moments. The single-point situation is for $\alpha = 0$. One sees that the final shapes are very different.

TABLE 1. Relative computation time in one optimization iteration

Task	Single-point	Multipoint (parallel M points)
Direct state (CFD)	1	M
Gradient (Adjoint)	7	$7 \times M$
Descent direction by (13)	10^{-4}	10^{-4}
Gradient based minimization	10^{-5}	10^{-5}

The factor of 7 is the observed extra calculation complexity for the automatically differentiated adjoint code generated by **Tapenade**. Building the descent direction involves Gram-Schmidt orthonormalizations. The optimization step uses constant step size and no linesearch is applied. The computational time for the Euler CFD code on a 2.5GHz core per mesh node and per RK time iteration is about 10^{-5} sec including the I/O steps.

Considering the normalized histories of the first four moments during optimization shown in figure 8, one sees that controlling these moments is something achievable. We voluntarily consider only a few iterations as in practice the amount of calculation effort is bounded and in particular there is no a priori knowledge of the necessary effort. We want to see the impact of the different optimization problem formulations on the outcome of the design.

One should however insist on the fact that these are very computer intensive simulations as each optimization iteration requires $M = 100$ direct state equations and M adjoint solutions with multiple right-hand-side (to access $\nabla_{\mathbf{x}}C_d$ and $\nabla_{\mathbf{x}}C_l$). These evaluations permit to build the multipoint-based descent direction $\delta_k = \delta(\mathbf{x}, \alpha_k)$ using Equation (20). These directions are then used to access the gradients of the different moments as described in section 3. Table 1 shows the relative computation time for each of these tasks during one optimization iteration.

Let us finish this discussion looking at the different histograms for the expression $j(\mathbf{x}, \alpha) = C_d + \sum_{i=1,3} c_i$ where we have used unit weighting coefficients in the penalty term. This expression is only used in this post-processing step and not to define the search directions. Figure 9 shows the frequencies of belonging to each of the categories versus a normalized expression for $j(\mathbf{x}_{opt}, \alpha) \leftarrow \frac{j(\mathbf{x}_{opt}, \alpha) - j_{min}^0}{j_{max}^0 - j_{min}^0}$ where $j_{max}^0 = \sup\{j(\mathbf{x}^0, \alpha_k), \alpha_k \in \mathbf{I}_M\}$ and $j_{min}^0 = \inf\{j(\mathbf{x}^0, \alpha_k), \alpha_k \in \mathbf{I}_M\}$ for the initial shape \mathbf{x}^0 .

Going from single-point to moment-based optimization, with more and more moments involved, changes the geometry of the optimization problem and this has an obvious impact on the convergence histories. This has been also observed in the optimization with the model problem as shown in figure 2. We start with the 4-moment formulation and the optimization algorithm stops when the changes observed in the moments are found negligible. Then we reduce the number of moments considered and proceed with the same number of iterations. Again, the optimization algorithm can stop before if the changes observed in the moments are found negligible. We should keep in mind that in real applications it is unlikely we can proceed up to a full convergence. One usually asks for an a priori number of iterations and this is somehow the philosophy adopted here: we would like to see the differences we have in the solution by the different approaches at a given calculation complexity (i.e. after a given number of optimization iterations). We recall that the extra cost of going from one to four moments is negligible as indeed, once the gradients at the sampling

points have been calculated, the extra cost only concerns linear combinations of these gradients and an orthonormalization of up to four vectors.

The modeling used here is based on the inviscid Euler equations. If one moves to more sophisticated physics (e.g. including viscous effects, turbulence, etc.) it is clear that the distributions will be different as the aerodynamic coefficients vary. Also we expect viscous effects to increase the impacts of transverse winds on the performance of the aircraft making even more important the necessity for robust design.

Finally, it is interesting to notice that despite the final histograms are quite close (see figure 9), the shapes vastly differ (see figure 7). This puts in evidence non uniqueness issues and the fact that going from 2 to 4 moments control requirement might lead to more than small variations of the design. We recover here the conclusions with the model problem of section 6.

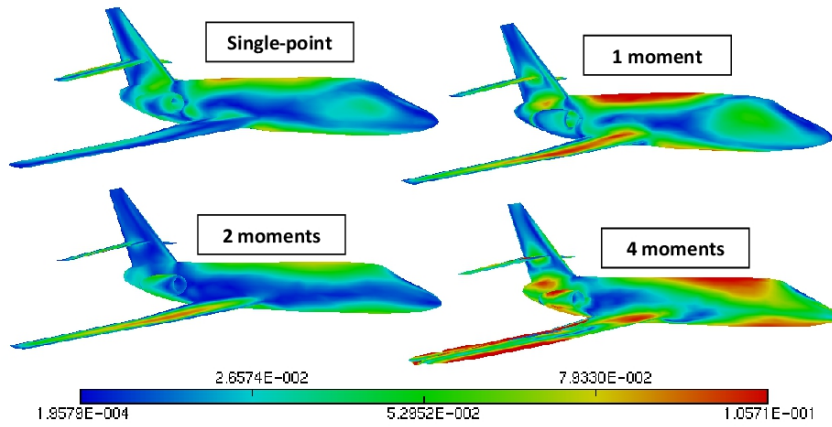


FIGURE 7. Total shape deformations from the initial shape with different moment control strategies. These are isovalues of the norm of the vector of shape deformation along the normal to the surface. The maximum deformation reaches 10 cm.

8. CONCLUDING REMARKS

Robust parametric optimization has been addressed as a problem of functional moments control expressed through a multipoint formulations. This permits to address situations where operating parameters are not anymore single valued but defined through their PDF. In this work only uniform PDFs have been considered with the operating parameters known through their ranges of variation. However, the approach can be extended without limitation to other PDFs. This would only require the definition of the sampling \mathbf{I}_M and the weights ω_k in Equations (5) and (6) according to these PDFs.

The first four moments of the functional have been considered in order to go beyond the classical mean-variance based optimization. Subtleties of what should be the target for the skewness and the kurtosis of the probability density function of the performance of the system for a robust design have been discussed.

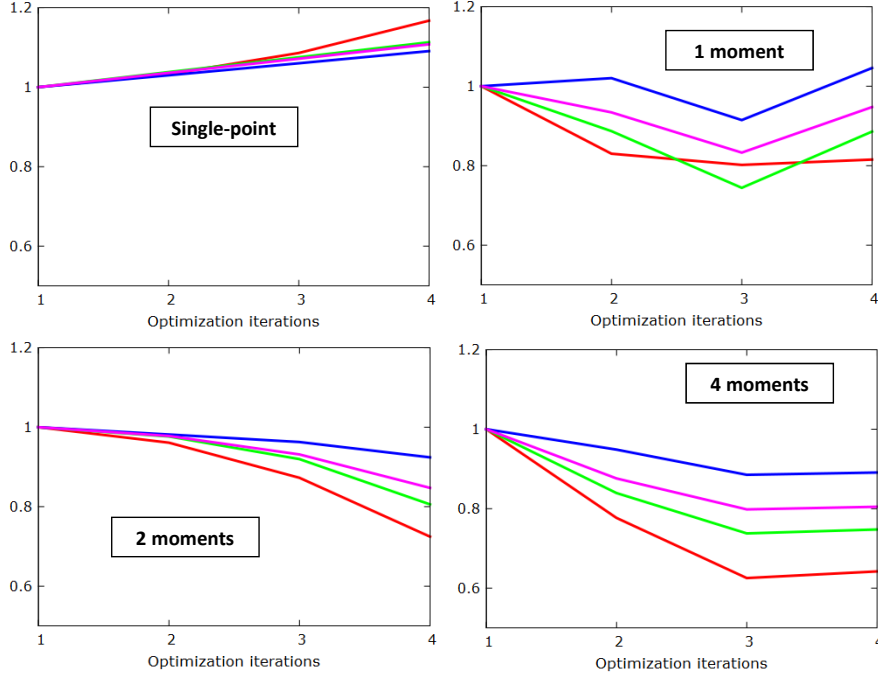


FIGURE 8. Histories of (μ/μ_0) , $(\sigma/\sigma_0)^{1/2}$, $(\gamma/\gamma_0)^{1/3}$ and $(\kappa/\kappa_0)^{1/4}$ for four optimization strategies.

It has been shown that controlling third and fourth moments can have major and non intuitive impacts on the design while the time to solution remains comparable to the single-point situation as all the extra calculations can be carried out in a fully parallel and independent manner. The approach requires quite small additional coding and computational effort compared to when solely considering one moment. Indeed, the cost of the calculation is in getting the multipoint states $U(\mathbf{x}, \alpha_k)$ and also $\nabla_x J(\mathbf{x}, \alpha_k)$ for the different α_k . Once we have these vectors, gradients of the second, third and fourth moments just involve linear combinations following Equations (7). And we need a Gram-Schmidt orthonormalization of up to four vectors to define the search direction d using relation (11) or (13), which again costs nothing. Therefore, once the multipoint design platform is in place, the extra cost of an optimization iteration remains negligible when going from one to four moments as the consideration of extra moments does not require any new state or adjoint evaluations. Also, depending on the application and the expert choices other requirements could have been chosen for each of the moments than those applied here. And more moments can be considered as well at nearly no extra cost. But, then one needs to be able to analyze the impact of their control on a design. In all cases, if one knows the kind of behavior one wants for any given moment, it can be included as well in the problem.

Finally, a discussion is proposed on how to provide some quantitative confidence level on the pertinence of the multipoint sampling and consequently on the quality of the search directions built from multiple gradients evaluation in constrained environments.

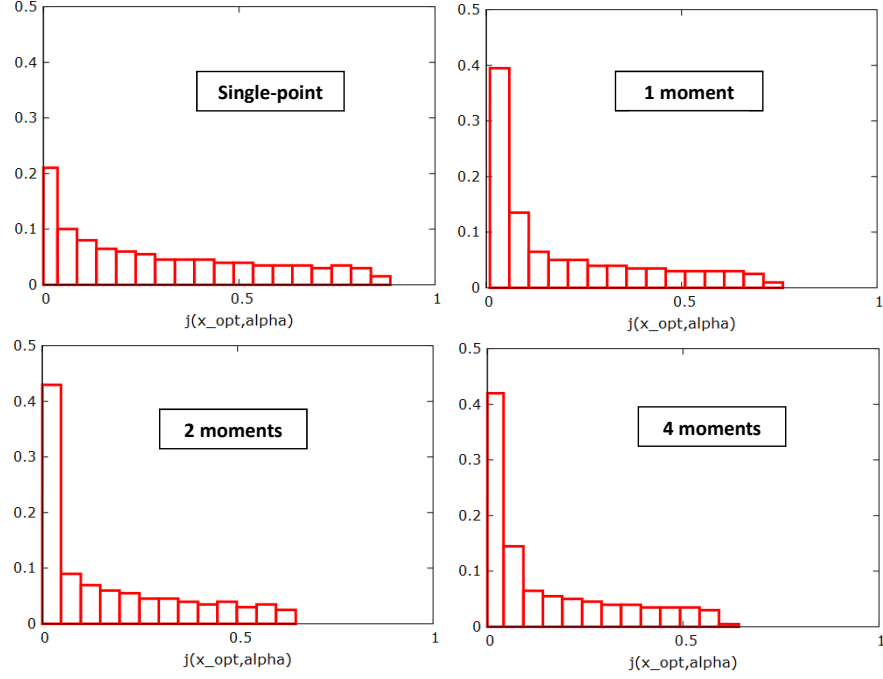


FIGURE 9. Histograms of the functional $j(\mathbf{x}, \alpha_k \in \mathbf{I}_M) = C_d + \sum_{i=1,3} c_i$ at the optimal solution for the single-point optimization and the different moment control strategies.

Acknowledgements

The author would like to thank G. Rogé from Dassault Aviation for his feedback and suggestions.

The multiple right-hand-side adjoint solver has been obtained using **Tapenade** AD tool developed at INRIA-Sophia Antipolis by L. Hascoet and his team.

REFERENCES

- [1] K. Scheinberg A. Conn and L. Vicente. *Introduction to Derivative-Free Optimization*. SIAM, NY, 2002.
- [2] AIAA. Guide for the verification and validation of computational fluid dynamics simulations. *AIAA*, G-077, 1998.
- [3] H.-J. Bungartz and M. Griebel. Sparse grids. *Acta Numerica*, 13:147–269, 2004.
- [4] S. Schmidt C. Schillings and V. Schulz. Efficient shape optimization for certain and uncertain aerodynamic design. *Computers & Fluids*, 46(1):78–87, 2011.
- [5] J. Caers. *Modeling Uncertainty in the Earth Sciences*. Wiley-Blackwell, 2011.
- [6] G. Casella and R. Berger. *Statistical Inference, 2nd Ed.* Duxbury Press, London, 2001.
- [7] B. Christianson. Reverse accumulation and implicit functions. *Optimization Methods and Software*, 9/4:307–322, 1998.
- [8] A. Dervieux. Steady euler simulations using unstructured meshes. *VKI Lecture series, Revised version published in Partial Differential Equations of hyperbolique type and Applications*, *World Scientific*, 1985/04:23–64, 1985.
- [9] C. Farhat and C. Degand. A three-dimensional torsional spring analogy method for unstructured dynamic meshes. *Computers & Structures*, 80/3:305–316, 2002.
- [10] R. Ghanem and A. Doostan. On the construction and analysis of stochastic models: characterization and propagation of the errors associated with limited data,. *J. of Comput. Phys.*, 217:63–81, 2006.

- [11] R. Ghanem and P. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer Verlag, New York, 1991.
- [12] L. Hascoet and V. Pascual. Tapenade user’s guide. In *INRIA Technical report*, pages 1–31. INRIA, 2004.
- [13] P.G. Hoel. *Introduction to Mathematical Statistics*. John Wiley, London, 1971.
- [14] J. J. Alonso M. J. Rimllnger J. J. Reuther, A. Jameson and D. Saunders. Constrained multipoint aerodynamic shape optimization using an adjoint formulation and parallel computers: part ii. *Journal of Aircraft*, 36/1:61–74, 1999.
- [15] R. O. Onez J. T. Spooner, M. Maggiore and K.M. Passino. *Stable Adaptive Control and Estimation for Nonlinear Systems: Neural and Fuzzy Approximator Techniques*. John Wiley, New York, 2002.
- [16] Ph. Jorion. *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill, New York, 2006.
- [17] M. Ghil K. Ide, P. Courtier and A. Lorenc. Unified notation for data assimilation: operational, sequential and variational. *Journal of the Meteorological Society of Japan*, 75/1B:181–189, 1997.
- [18] R.E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, 82:35–45, 160.
- [19] K.C. Kapur and L.R. Lamberson. *Reliability in Engineering Design*. John Wiley & Sons, New York, 1997.
- [20] H.R. Lindman. *Analysis of Variance in Complex Experimental Designs*. Freeman, New York, 1974.
- [21] R. Wuchner M. Firl and K. Bletzinger. Regularization of shape optimization problems using fe-based parametrization. *Structural and Multidisciplinary Optimization*, 47/4:507–521, 2013.
- [22] D. W. Zingg M. Nemec and T. H. Pulliam. Multipoint and multi-objective aerodynamic shape optimization. *AIAA J.*, 42/6:1057–1065, 2004.
- [23] J. R. R. A. Martins and A. B. Lambe. Multidisciplinary design optimization: A survey of architectures. *AIAA J.*, 51(9):2049–2075, 2013.
- [24] B. Mohammadi. Reduced sampling and incomplete sensitivity for low-complexity robust parametric optimization. *Int. J. Num. Meth. Fluids*, 73/4:307–323, 2013.
- [25] B. Mohammadi. Principal angles between subspaces and reduced order modeling accuracy in optimization. *Structural and Multidisciplinary Optimization*, 50/2:237–252, 2014.
- [26] B. Mohammadi. Uncertainty quantification by geometric characterization of sensitivity spaces. *Compt. Meth. Appl. Mech. Eng.*, 280:197–221, 2014.
- [27] B. Mohammadi. Value at risk for confidence level quantifications in robust engineering optimization. *Optimal Control: Applications and Methods*, 35/2:179–190, 2014.
- [28] B. Mohammadi. Ensemble kalman filters (enkf) and geometric characterization of sensitivity spaces for uncertainty quantification in optimization. *Computer Methods in Applied Mech. & Eng.*, 290:228–249, 2015.
- [29] B. Mohammadi. Parallel reverse time integration and reduced order models. *J. of Computational Mathematics*, 2:17–33, 2015.
- [30] B. Mohammadi. Backward uncertainty propagation in shape optimization. *Int. J. for Numerical Methods in Fluids*, 103(4)-DOI: 10.1002/flid.4077:307–323, 2016.
- [31] B. Mohammadi and O. Pironneau. *Applied Shape Optimization for Fluids (2nd Edition)*. Oxford Univ. Press, Oxford, 2009.
- [32] B. Mohammadi and O. Pironneau. Shape optimization in fluid mechanics. *Annual Revue of Fluid Mechanics*, 36/1:255–279, 2004.
- [33] J. Nocedal and S. Wright. *Numerical Optimization*. Springer, NY, 2006.
- [34] A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. SIAM, N.Y., 1987.
- [35] X. Wan and G.E. Karniadakis. Multi-element generalized polynomial chaos for arbitrary probability measures. *SIAM J. Sci. Comput.*, 28/3:901–928, 2006.
- [36] D. Xiu. *Numerical Methods for Stochastic Computations: A Spectral Method Approach*. Princeton University Press, 2010.
- [37] G. K. W. Kenway Z. Lyu and J. R. R. A. Martins. Aerodynamic shape optimization investigations of the common research model wing benchmark. *AIAA J.*, 53/4:968–985, 2015.