Comput. Methods Appl. Mech. Engrg. 200 (2011) 216–232

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

Shape optimization of an airfoil in a BZT flow with multiple-source uncertainties

P.M. Congedo [a,*], C. Corre [a], J.-M. Martinez [b]

[a] Institut Polytechnique de Grenoble, Laboratoire des Ecoulements Géophysiques et Industriels, 1023 rue de la piscine, 38400 Saint Martin d'Hères, France
[b] CEA – Saclay, Département de Modélisation des Systèmes et Structures, F-91190 Gif-sur-Yvette, France

ARTICLE INFO

Article history:
Received 8 March 2010
Received in revised form 12 August 2010
Accepted 13 August 2010
Available online 20 August 2010

Keywords:
Uncertainty quantification
Robust optimization
Polynomial chaos
Sparse grid
Real gas
Thermodynamic model

ABSTRACT

Bethe–Zel'dovich–Thompson fluids (BZT) are characterized by negative values of the fundamental derivative of gasdynamics for a range of temperatures and pressures in the vapor phase, which leads to non-classical gasdynamic behaviors such as the disintegration of compression shocks. These non-classical phenomena can be exploited, when using these fluids in Organic Rankine Cycles (ORCs), to increase isentropic efficiency. A predictive numerical simulation of these flows must account for two main sources of physical uncertainties: the BZT fluid properties often difficult to measure accurately and the usually fluctuating turbine inlet conditions. For taking full advantage of the BZT properties, the turbine geometry must also be specifically designed, keeping in mind the geometry achieved in practice after machining always slightly differs from the theoretical shape. This paper investigates some efficient procedures to perform shape optimization in a 2D BZT flow with multiple-source uncertainties (thermodynamic model, operating conditions and geometry). To demonstrate the feasibility of the proposed efficient strategies for shape optimization in the presence of multiple-source uncertainties, a zero incidence symmetric airfoil wave-drag minimization problem is retained as a case-study. This simplified configuration encompasses most of the features associated with a turbine design problem, as far the uncertainty quantification is concerned. A preliminary analysis of the contributions to the variance of the wave-drag allows to select the most significant sources of uncertainties using a reduced number of flow computations. The resulting mean value and variance of the objective are next turned into metamodels. The optimal Pareto sets corresponding to the minimization of various substitute functions are obtained using a genetic algorithm as optimizer and their differences are discussed.

1. Introduction

Dense gases (DGs) are defined as single-phase vapors operating at temperatures and pressures of the order of magnitude of those of their thermodynamic critical point. The study of the complex dynamics of compressible dense gas flows is strongly motivated by their potential technological advantages as working fluids in energy-conversion cycles such as Organic Rankine Cycles (ORCs). Specific interest has developed in a particular class of dense gases, known as the Bethe–Zel'dovich–Thompson (BZT) fluids [1], which exhibit non-classical gasdynamic behaviors in a range of thermodynamic conditions above the liquid/vapor coexistence curve, such that the fundamental derivative of gasdynamics $\Gamma = 1 + \frac{\rho}{a}\left(\frac{\partial a}{\partial \rho}\right)_s$ with ρ the fluid density, a the isentropic sound speed and s the specific entropy, becomes negative. At these conditions, the well-known compression shocks of the perfect-gas (PFG) theory violate the entropy inequality over a certain range of temperatures and pressures in the vapour phase and are therefore not admissible [2]. The BZT properties are generally encountered in fluids possessing large heat capacities and formed by complex, heavy molecules, such as some commercially available heat transfer fluids. The non-classical phenomena typical of BZT fluids have several practical outcomes: prominent among them is an active research effort to reduce losses caused by wave drag and shock/boundary layer interactions in turbomachines and nozzles [3–6], with particular application to ORCs used to generate electric energy in low-power applications. The use of BZT fluids [5,6] as ORC working fluids is potentially interesting because the shock formation and the consequent losses could be ideally avoided if turbine expansion could happen entirely within or very close to the inversion zone where $\Gamma < 0$. In fact, as pointed out in [7,8] and recently investigated in [9,10], rarefaction shock waves are physically admissible in the inversion region.

The point of interest in the present paper is that a dense gas flow in a turbine cascade is characterized by a significant uncertainty on the physical parameters and on the operating conditions at the turbine inlet [11]. Indeed, the ORCs are mainly used in

* Corresponding author.
E-mail addresses: pietro.congedo@hmg.inpg.fr (P.M. Congedo), Christophe.corre@hmg.inpg.fr (C. Corre), jean-marc.martinez@cea.fr (J.-M. Martinez).

0045-7825/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cma.2010.08.006

biomass and geothermal applications – development in solar and heat recovery applications are also expected – where the "renewable" heat sources display a non-negligible level of variability. Besides, the thermophysical properties of dense gases are themselves characterized by a strong uncertainty [12,13]: the experimental data on the critical parameters and on the caloric properties are known with an uncertainty that can exceed 5%. The determination of critical-point data for dense gases is indeed delicate because such gases may decompose, totally or partly, at temperatures close to the critical one; in this case, critical point values just rely on estimates. Note the accuracy of an equation of state can theoretically be improved by increasing the number of expansion terms, as done with the virial expansion for the Martin–Hou equation for instance. However, increasing the number of constants in the model requires a larger number of thermodynamic inputs and introduces new sources of uncertainty. As demonstrated in [14], a somehow paradoxal situation occurs where increasing the complexity of the thermodynamic model to improve its (deterministic) accuracy, eventually leads to larger modeling uncertainties because of a larger number of uncertain parameters. An analysis of thermodynamic uncertainties was performed in [14] using a non-intrusive polynomial chaos (PC) approach, retained for its flexibility. This feature was considered crucial when dealing with evolving flow solvers since regularly modified to accommodate new equations of state.

When designing a turbine specifically adapted to BZT fluids, a meaningful numerical prediction of the performance must necessarily take into account these uncertainties on the thermophysical properties but also on the inlet boundary conditions which are typical of dense gases applications. Moreover, a robust shape optimization must also include the more generic uncertainty introduced by the machining of the physical blade. Papers devoted to stochastic robust optimization are mostly found in the field of Multidisciplinary Optimization (MDO), with a well- developed line of research on structural robust design optimization, where uncertainties are found on the loads, constraints or material properties [15–19]. Several contributions deal with aeronautics problem and uncertainties on the operating and off-design conditions [20–23]. To the best of our knowledge, there are few papers [24,25] devoted to robust optimization with a PC-based approach to take into account the uncertainties. Besides, previous works on dense gas turbine [26] and airfoil shape [27] optimizations were purely deterministic studies which it will be interesting to reinvestigate at the light of the uncertainty quantification methodology developed in the present paper.

Building on the approach followed in [14], a straightforward extension to shape optimization with uncertainties not only on thermophysical properties but also on inlet conditions and geometry could be proposed through the coupling of the previously used PC approach and the CFD dense gas solver for this increased number of uncertainties. However, such a direct strategy would soon become unrealistic due to the excessive amount of computational time it would require. The present paper explores ways to make tractable the problem of shape optimization with multiple sources of uncertainty. The simplified drag minimization problem for a BZT flow over an isolated airfoil is retained as case of study and two ways of reducing the cost of this optimization when thermophysical, inlet boundary and geometrical uncertainties are taken into account are proposed and analyzed: (i) identifying the most significant uncertainties through a preliminary analysis of the contributions to the drag variance on a baseline configuration, (ii) deriving surrogate models for the mean drag and its variance, on which the optimization loop is applied.

The paper is organized as follows. Section 2 is devoted to the description of the various numerical tools needed in the study – CFD tools for the dense gas flows simulations, tools for uncertainty quantification – and details the coupling between CFD, shape optimization and uncertainty quantification. Section 3 describes the preliminary analysis performed to screen the most influential uncertain parameters contributing to the variance of the airfoil drag. Section 4 compares different strategies to derive metamodels for the mean drag and its variance and analyzes the set of optimal solutions provided by the simultaneous minimization of these objectives. The closing Section 5 summarizes the conclusions that can be drawn from the present work regarding the topic of shape optimization with multiple-sources of uncertainty and proposes new lines of study to be followed for designing optimal shapes for dense gas flows.

2. Methodology and tools

2.1. CFD tools for dense gas flows

2.1.1. Conservation laws and thermodynamic models

The two-dimensional Euler equations completed by a real-gas thermodynamic model are solved using a cell-centered finite volume discretization on a general unstructured grid dividing the flow domain into a finite number of triangles or quadrangles. The time rate of change of the cell-averaged state vector w is balanced with the area-averaged (inviscid) numerical fluxes, computed across each cell face using a HLLC scheme. The Harten, Lax, and van Leer (HLL) scheme developed in [28] is attractive because of its robustness and easy extension to dense gas flows. This approximate Riemann solver relies indeed on the characteristic speeds of two pressure waves, which depend on the local velocity and speed of sound; adapting the definition of the speed of sound to the equation of state describing a dense gas makes the scheme ready for use on the flows of interest in the present study. The HLL solver is known to be overly diffusive because it neglects the contact surface when reducing the exact Riemann problem to two pressure waves. This limitation was lifted by Toro et al. [29] where the modified three-wave solver named HLLC was derived, with an explicitly presented contact. The definition of the associated wave speed does itself depend on the wave speeds previously generalized to dense gas flows. The HLLC scheme has been extended in [30] to deal with low-Mach number flows by a simple modification of the signal velocities; the extension of this preconditioned HLLC scheme to low-speed real gas has been performed in [31]. Though the low-Mach preconditioning is not activated in the present transonic flow computations, the solver described in [31] is retained in this study. Second-order space accuracy is ensured by a MUSCL-type reconstruction process on the conservative variables [32], extended to unstructured grid finite-volume formulation. The HLLC numerical flux on each cell face makes use of linearly reconstructed states in the two cells sharing the face. These states are computed from the cell-centered values of the conservative variables and the cell-centered gradient estimates. Rather than using Green-Gauss formulae to compute the cell-gradient, as proposed in [33], least-squares formulae with a prescribed support for each cell are applied. The details of the cell-gradient calculation can be found in [34,35]. To allow an oscillation-free capture of flow discontinuities, the reconstruction process includes slope limiters as initially proposed in [33] and revisited in [36]. Following a procedure inspired from [37], the second-order HLL scheme is coupled with a simple first-order Rusanov type implicit stage which allows the use of large time-steps hence fast convergence to steady state. The numerical flux through the boundary edges is computed using an inflow/outflow characteristic-based condition at the far-field boundary and a mirror boundary condition to define the ghost-cell states at the airfoil surface.

2.1.2. Thermodynamic model: Peng–Robinson equation of state

Various strategies are followed in the literature to describe gases close to saturation conditions, or in the supercritical region, or in the two-phase regime [38–44]. Equations of state based on theoretical and analytical criteria, such as the van der Waals, Redlich–Kwong, Peng–Robinson, Martin–Hou equations (see [45], for more details) depend on a limited number of thermodynamic inputs (e.g. critical temperature and pressure, acentric factor, etc.), which are measured with a significant experimental error in the case of dense gases. In the present study, the Peng–Robinson–Strijek–Vera (PRSV) cubic equation [46] is retained as thermal equation of state not only because it provides a good level of accuracy for a reduced computational cost [47,48] but also for its robustness when dealing with uncertainties on thermophysical properties. In [14] this equation was indeed proved to yield a lower variance in the flow prediction when the dense gas physical properties were considered to be affected by uncertainties, with respect to other equations such as Martin–Hou's. Note the highly complex and accurate Span–Wagner equation of state [13] is not considered in this study because it does not explicitly depend on measurable physical parameters hence does not lend itself to an uncertainty quantification analysis. The PRSV equation of state can be expressed in the following non-dimensional form:

$$p_r = \frac{T_r/Z_c}{1/\rho_r - b_r} - \frac{a_r}{1/\rho_r^2 + 2b_r/\rho_r - b_r^2}, \tag{1}$$

where $p_r = p/p_c$, $T_r = T/T_c$, $\rho_r = \rho/\rho_c$ are, respectively, the reduced pressure, temperature and density, *i.e.* quantities normalized by their values at the critical point. The coefficients appearing in (1) are defined by:

$$\begin{cases} a_r = \left(0.457235/Z_c^2\right)\alpha(T_r), \\ b_r = 0.077796/Z_c, \end{cases} \tag{2}$$

where Z_c is univocally determined by the solution of a cubic equation. The function $\alpha(T_r)$ reads:

$$\alpha(T_r) = \left[1 + m\left(1 - T_r^{0.5}\right)\right]^2 \tag{3}$$

with the coefficient m depending on the acentric factor ω:

$$m = 0.480 + 1.57\omega - 0.176\omega^2. \tag{4}$$

The value of the acentric factor itself can be either derived from experiments or computed using the expression reported in [45]:

$$\omega = \frac{3T_e/T_c}{7(1 - T_e/T)}\log_{10}(p_c) - 1, \tag{5}$$

where T_e is the normal boiling temperature. For the calculation of all caloric properties, the thermal equation is supplemented with the ideal gas contribution to the specific heat at constant volume, approximated here by a power law of the form:

$$c_{v\infty}(T_r) = c_{v\infty}(T_c)(T_r)^{nexp}, \tag{6}$$

where the exponent *nexp* and the critical-point ideal specific heat are material-dependent constants. The coefficient *nexp* is obtained by minimizing the error with respect to experimental data, and is provided in literature for different fluids (see for example [4]). Then, the caloric equation of state associated to the chosen thermal equation of state is completely determined via the compatibility relation:

$$\begin{aligned} e &= e(\rho_r, T_r) \\ &= e_0 + \int_{T_0}^{T_r} c_{v\infty}(T')dT' - \int_{\rho_0}^{\rho_r}\left[T_r\left(\frac{\partial p_r}{\partial T_r}\right)_{\rho_r} - p_r\right]\frac{d\rho'}{\rho'^2}, \end{aligned} \tag{7}$$

where quantities with a prime subscript are auxiliary integration variables, and the subscript 0 indicates a reference state. The thermal non-dimensional PRSV model and its associated caloric Eqs. (1), (6) and (7) eventually depend on the set of parameters input_$T = \{\omega, nexp, c_{v\infty}(T_c)/R\}$.

2.1.3. Baseline configuration and sources of uncertainty

The steady transonic inviscid flow of a dense gas over a symmetric profile is considered as baseline configuration for investigating a shape optimization process taking into account uncertainties; this choice, over for instance a turbine cascade geometry, is mainly motivated by the objective of cost reduction when performing this preliminary study. The initial profile is a sonic arc for a perfect gas (with $\gamma = 1.4$) flow, that is the shape around which the flow is nowhere supersonic even when the inflow Mach number is getting close to unity. Note it is possible to theoretically derive the shape of a sonic arc in the case of a perfect gas or a real gas described by the Van der Waals equation of state, as reported in [49]. The fluid considered is the heavy fluorocarbon PP10 ($C_{13}F_{22}$), with critical properties reported in Table 1. The thermodynamic (reduced) conditions of the free-stream are $p_r = 0.985$, $\rho_r = 0.622$ with a freestream Mach number $M_\infty = 0.95$ (there is no incidence, only a half-airfoil being computed). The set of operating conditions is denoted from now on as input_O = $\{M_\infty, p_r, \rho_r\}$. The saturation curve for PP10 constructed by means of the PRSV equations is reported with the iso-Γ curves in Fig. 1: it can be observed the chosen set for Input_O lies near the inversion region ($\Gamma \approx 0.3$) yet not too close to ensure the thermodynamic state remains in the monophasic region even when uncertainties introduce fluctuations in this thermodynamic state. This configuration has also been previously studied by the authors: in [2] for a similar optimization process but without taking into account the uncertainty and in [14] for analyzing the sole thermodynamic uncertainties independently from any optimization process. In accordance with now

Table 1
Mean values adopted for the fluid considered, the PP10.

Property	P_c (atm)	T_c (K)	Z_c	T_c (K)	n	$C_{v,\infty}(T_c)/R$	ω
	16.2	630.2	0.2859	467	0.5255	78.37	0.4833

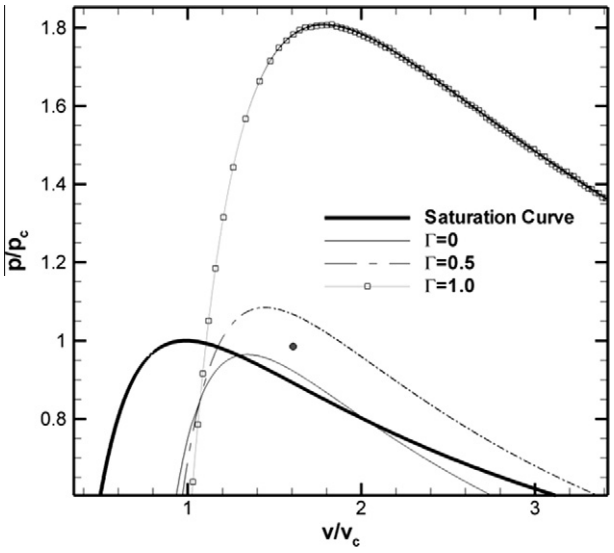

Fig. 1. Saturation curve and iso-gamma for the PP10, computed by means of PRSV equation, filled point the chosen operating condition.

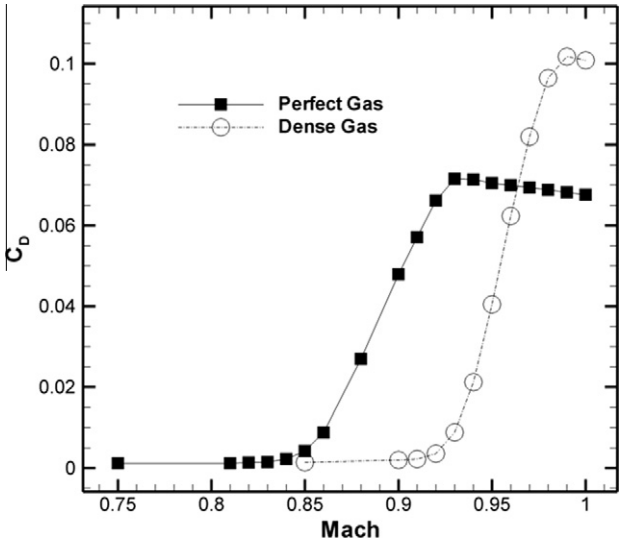

Fig. 2. Drag coefficient of the sonic arc at zero incidence as a function of Mach number for a perfect and dense-gas flow.

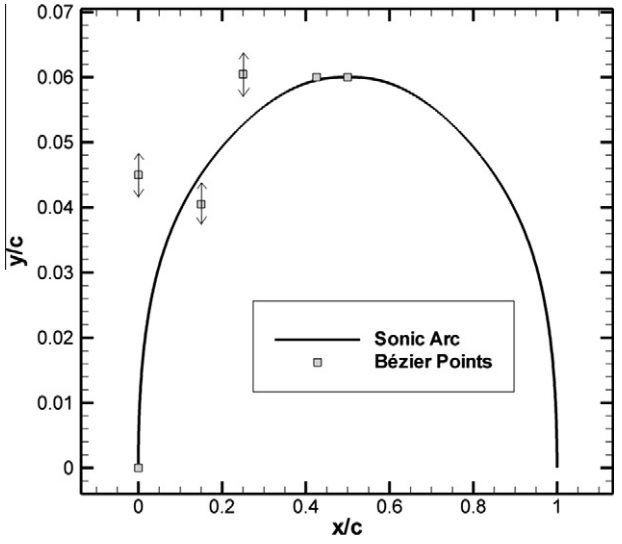

Fig. 3. Baseline profile with its interpolating Bézier points.

well-established conclusions from previous studies [3,4,43,49], the use of a BZT gas allows a significant increase of the critical Mach number as well as the divergence Mach number: for the present sonic arc geometry, the drag divergence occurs for a freestream Mach number of about 0.83 in the case of the perfect gas flow and 0.91 in the case of the dense gas flow (Fig. 2).

All the flow computations performed in the study use a half C-grid made of 100×32 cells, with a mean height of the first cell closest to the wall equal to 0.001 chords and an outer boundary located at 10 chords from the airfoil surface. This grid represents a reasonable trade-off between accuracy and computational cost, given the number of CFD runs required by the stochastic solver. Some considerations about the effect of the computational grid on the quality of the stochastic simulations are reported in [14] where stochastic simulations have been performed with uncertainties on the physical properties of the dense gas only: it was demonstrated, for a flow configuration similar to the one studied here, the proposed level of grid refinement is sufficient to provide a reasonable representation of the flow solution expectancy and variance.

The sonic arc is defined by a Bézier polynomial [2] using six points of fixed position along the airfoil unit chord as indicated in Fig. 3. The arc being symmetric with respect to $x/c = 0.5$, the geometry is constructed on $[0,0.5]$ and then reflected on $[0.5,1.0]$. The first and last point are, respectively, fixed to $(0,0)$ and $(0.5,0.06)$, the airfoil thickness normalized by the chord being chosen equal to 0.12. The position of the fifth control point is fixed to $(0.426,0.06)$ to ensure the tangent to the profile at mid-airfoil ($x/c = 0.5$) is horizontal. The y-coordinates of the three internal points are left free and form a set of geometric parameters denoted as input_G = $\{y_1,y_2,y_3\}$.

The parameters defining the set input_O of operating conditions are supposed to display a 1% variation with respect to their prescribed average values. This choice allows to deal with a significant uncertainty while remaining in a monophasic region; for a stronger variation of input_O parameters, the flow could enter the liquid–vapour mixture region, which must be avoided since the flow solver does not include two-phase flow models. The physical properties of the PRSV model defining the set input_T are affected by a very strong uncertainty (see [10]), reflected in a 3% variation with respect to their mean values, which are reported in Table 1 and correspond to the heavy fluorocarbon PP10 [12]. Note it has been checked again that all the possible thermodynamic states for the chosen uncertainty ranges always remain in the monophasic region. The geometric tolerances are taken into account by injecting uncertainties on the Bezier parameters of input_G: a maximal variation of 0.7% is considered. A summary of the mean values and uncertainty intervals associated with the three sources of uncertainties and the nine uncertain parameters is provided in Table 2.

2.2. Tools for uncertainty quantification (UQ)

Non intrusive polynomial chaos are used to deal with the quantification of the effect of uncertainties on the dense gas flow over the airfoil. The number of CFD simulations needed to perform this quantification is minimized by making use of a Sparse Grid approach based on Smolyak's construction. The key ingredients of the methodology followed are described in this section.

2.2.1. Polynomial chaos

Polynomial chaos (PC) expansions are derived from the original theory of Wiener on spectral representation of stochastic processes using Gaussian random variables. PC expansions have been used for UQ by Ghanem and Spanos [50] and extended by Xiu and Karniadakis [51] to non-Gaussian processes. Any well-behaved process y (e.g. a second-order process) can be expanded in a convergent (in the mean square sense, see Cameron and Martin [52]) series of the form:

$$y(x,t;\xi) = \sum_{\alpha} y_\alpha(x,t)\Psi_\alpha(\xi), \tag{8}$$

where ξ is a set of nx independent random variables $\xi = (\xi_1,\xi_2,\ldots,\xi_{nx})$ and α a multi-index $\alpha = (\alpha_1,\alpha_2,\ldots,\alpha_{nx})$ with each component $\alpha_i = 0, 1,\ldots$. The multivariate polynomial function Ψ_α is defined by a product of orthogonal polynomials $\Phi_i^{\alpha_i}(\xi_i)$ in relation to the probability density of the random variable ξ_i, i.e. $\Psi_\alpha(\xi) = \prod_{i=1}^{nx}\Phi_i^{\alpha_i}(\xi_i)$.

A one-to-one correspondence exists between the choice of stochastic variable ξ_i and the polynomials $\Phi_i^{\alpha_i}$ of degree α_i. For instance if ξ_i is a normal/uniform variable, the corresponding $\Phi_i^{\alpha_i}$ are Hermite/Legendre polynomials of degree α_i; the degree of Ψ_α is $|\alpha|_1 = \sum_{i=1}^{nx}\alpha_i$. The multivariate polynomial functions Ψ_α are orthogonal with respect to the probability distribution function of the vector ξ of standard independent random variables ξ_i, $i = 1, 2,\ldots,nx$. Coefficients $y_\alpha(x,t)$ are the PC coefficients or stochastic

Table 2
Summary of the mean values, maximal variation and probability density function (PDF) for input_G, input_O and input_T.

		Mean values	Max variation	PDF
Input_G $\{X_1,X_2,X_3\}$	Decoupled	$\{0.044993, 0.040517, 0.060488\}$	0.7%	Uniform/Gaussian
	Coupled			Uniform
Input_O $\{M_\infty, p_r, p_r\}$	Decoupled	$\{0.95, 0.985, 0.622\}$	1%	Uniform/Gaussian
	Coupled			Gaussian
Input_T $\{\omega, n, c_{v,\infty}(T_c)/R\}$	Decoupled	$\{0.5255, 78.37, 0.4833\}$	3%	Uniform/Gaussian
	Coupled			Gaussian

modes of the random process y. Defining the scalar product by the expectation operator yields:

$$y_\alpha(x,t) = \langle y(x,t), \Psi_\alpha \rangle \|\Psi_\alpha\|^{-2}. \tag{9}$$

For practical use, the PC expansions have to be truncated in term of degree polynomial no:

$$y(x,t,\xi) = \sum_{|\alpha|_1 \leq no} y_\alpha(x,t)\Psi_\alpha(\xi). \tag{10}$$

The number of multivariate polynomials Ψ_α, *i.e.* the dimension of the expansion basis, is related to the stochastic dimension nx and the degree no of polynomials and is given by the formula $(nx+no)!/(nx!\ no!)$.

Several approaches can be used to estimate PC coefficients. The first approach is based on a Galerkin projection of the model equations; it leads to a set of coupled equations and requires an adaptation of the deterministic code in order to be applied. Alternative non-intrusives approaches are based on Monte Carlo simulations or quadrature formulae to evaluate PC coefficients (see for instance [53]). When the number d of variables is large, quadrature formulae based on tensor product of a 1D formula require too many numerical evaluations and Sparse Grids integration based on Smolyak's construction [54] are preferred. For these non-intrusive approaches, PC coefficients are therefore evaluated from a set of points (ξ^i, ω_i) and weights by formulae of the form:

$$y_\alpha(x,t) = \|\Psi_\alpha\|^{-2} \sum_{i=1}^{n} y(x,t;\xi^i)\ \Psi_\alpha(\xi^i)\ \omega_i. \tag{11}$$

From this PC expansion of the random process $y(x,t)$, it is easy to derive its mean and variance and to deal with the sensitivity analysis using the analysis of variance (ANOVA). Before detailing this sensitivity analysis, let us introduce the Sparse Grid Method, very useful to reduce the number of computer simulations required by the UQ process.

2.2.2. Sparse Grid Method

PC coefficients are computed from multidimensional integration based on quadrature formulae. To reduce the number of computer simulations a method based on Smolyak's construction is used. Smolyak's cubature formulae, also called Sparse Grid, allow us to estimate the integration of a function of d variables over a domain Ω

$$If^{(d)} := \int_\Omega f^{(d)}(\xi)d\xi. \tag{12}$$

The construction of the Smolyak's cubature is based on the tensor-product of one-dimensional quadrature formulas of level l, for the univariate function $f^{(1)}$:

$$If^{(1)} \approx Q_l^1 f := \sum_{i=1}^{n_l^1} w_{li} f(x_{li}), \tag{13}$$

using the weight w_{li} and abscissas x_{li} of a sequence Γ_l^1 of n_l^1 points. Let us introduce the difference quadrature formula:

$$\Delta_k^1 f := (Q_k^1 - Q_{k-1}^1)f \quad \text{with } Q_0^1 f := 0, \tag{14}$$

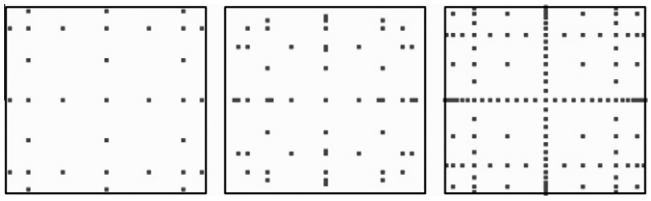

Fig. 4. Sparse Grids corresponding to Kronrod Patterson (on the left) (33 points), Gauss Legendre (in the middle) (55 points, not nested) and Clenshaw Curtis (on the right) (65 points) sequences for $d=2$ (dimension) and $l=5$ (level equal to polynomial degree of exactness), uniform distribution with variables varying in $[0,1]$.

which is defined on the union of the grids $\cup_{k=1}^{l} \Gamma_k^1$. Smolyak's construction for d-dimensional functions is given by

$$Q_l^d f := \sum_{|k|_1 \leq l+d-1} (\Delta_{k_1}^1 \otimes \cdots \otimes \Delta_{k_d}^1) f, \tag{15}$$

where $|k|_1 = \sum_{i=1}^{d} k_i$. To minimize functions evaluations (number of points in the union of the grids), an interesting case is when the set of one-dimensional quadrature are nested, i.e. $\Gamma_{l-1}^1 \subset \Gamma_l^1 \Rightarrow \Gamma_{l-1}^d \subset \Gamma_l^d$. If $n_l^1 = O(2^l)$ the order of n_l^d is $O(2^l l^{d-1})$ to be compared with product rules for which the order is $O(2^{ld})$. The efficiency of Sparse Grid increases with the dimension of the problem. The polynomial degree of exactness of the Smolyak construction is $l+d-1$ if the one dimensional quadratures are exact for polynomial of degree l. Examples for Sparse Grid corresponding to delayed Clenshaw Curtis, Gauss Legendre and Fejer rules for $d=2$ and $l=5$ are shown in Fig. 4. These Sparse Grids are computed by the toolbox Nisp [55], which includes the SmolPack library of Petras to use delayed Clenshaw Curtis sequences [56]. The size of the grids depends of the level (l), the number of stochastic variables (d) and the used rule. For instance to compute exactly the coefficients of a polynomial function of degree 3 with 9 uncertainties, the size of the sample with the SmolPack is 871 against 1177 with Gauss formulae (for which grids are non nested), hence a 26% reduction in CFD evaluations. For a polynomial function of degree 4 with 9 uncertainties, 3463 and 5965 runs are, respectively, needed by SmolPack and Gauss formulae, corresponding to a 42% reduction offered by the Sparse Grid approach.

2.2.3. Sensitivity analysis from polynomial chaos

From the PC expansion (16), it is easy to derive the mean E and variance V of the random process $y(x,t)$. To simplify notations, an orthonormal basis is chosen, i.e. $\|\Psi_\alpha\| = 1\ \forall\alpha$; the following equalities hold:

$$\begin{aligned} E(y(x,t)) &= y_0(x,t). \\ V(y(x,t)) &= \sum_{\alpha}^{y_\alpha^2(x,t)}. \end{aligned} \tag{16}$$

Another interesting property of PC expansion is to make easier sensitivity analysis [46]; so-called global sensitivity analysis [57] is based on the analysis of variance (ANOVA). Let us introduce the

generic model $Y(\xi) = f(X(\xi))$ where X is a set of nx independent random variables defined from probabilistic transformation $X_i = X_i(\xi_i)$. The variance decomposition of the response is:

$$V(Y) = \sum_{u \subseteq U} \sigma_u^2(X_u), \tag{17}$$

where $U = (1,2,\ldots,nx)$ is the set of random variables indexes and σ_u^2 is the variance introduced by interactions of random variables $X_u \subseteq X_U$. The variance decomposition is based on $2^{nx} - 1$ terms σ_u. From a PC expansion it is easy to compute all variances terms σ_u^2 of X_u equivalent to ξ_u, i.e. $\sigma_u^2 = \sum_{\alpha \geqslant u} y_\alpha^2$, with the notation $\alpha \geqslant u \Rightarrow \alpha_i \geqslant u_i \ \forall\ i = 1, 2, \ldots, nx$.

The associated sensitivity measure of X_u, named Sobol's index [58], is written as the correlation ratio:

$$S_u = \frac{V[E(Y|X_u)]}{V(Y)} = \frac{\sum_{v \leqslant u} \sigma_v^2}{V(Y)}. \tag{18}$$

For a single variable, the first order Sobol's index is obtained by:

$$S_i = \frac{\sigma_i^2}{V(Y)}. \tag{19}$$

If the generic model is linear, i.e. $Y = \sum_{i=1}^{nx} \alpha_i X_i$, the first Sobol's index S_i is equal to the classical linear correlation coefficient, so the Sobol's indexes allow us to deal with global sensitivity analysis for non-linear models.

2.3. Coupling CFD/ shape optimization and UQ

2.3.1. Problem formulation

Our aim is to find an optimal shape for an isolated symmetric airfoil which provides a robust minima for the drag coefficient. Robust means this shape simultaneously minimizes the mean value and the variance of the drag coefficient computed when taking into account the physical uncertainties of the problem gathered in the previously defined sets input_T, input_O and input_G. Note this bi-objective problem possesses a set of solutions, forming the global Pareto front and corresponding to various trade-off between a low mean value and a low variance. In mathematical terms, the problem to solve is expressed as:

$$\min_{y_1, y_2, y_3} \left(mean(C_D), \sigma(C_D)\right) \tag{20}$$

with $\{y_1, y_2, y_3\}$ varying in the solution space S, made of intervals defined as $[\overline{y_i} - \delta y_i; \overline{y_i} + \delta y_i]$ with $\overline{y_i}$ the reference mean value for the geometric parameter y_i (these reference mean values are those used to reproduce the baseline sonic arc) and δy_i 10% of the corresponding mean value $\overline{y_i}$. The drag coefficient C_D is a function of input_T = $\{\omega, nexp, c_{v\infty}\ (T_c)/R\}$, input_O = $\{M_\infty, p_r, \rho_r\}$ and input_G = $\{y_1, y_2, y_3\}$. With the non-intrusive statistical approach considered in this paper, the stochastic simulation allowing to compute the mean and the variance of C_D reduces to perform series of CFD evaluations

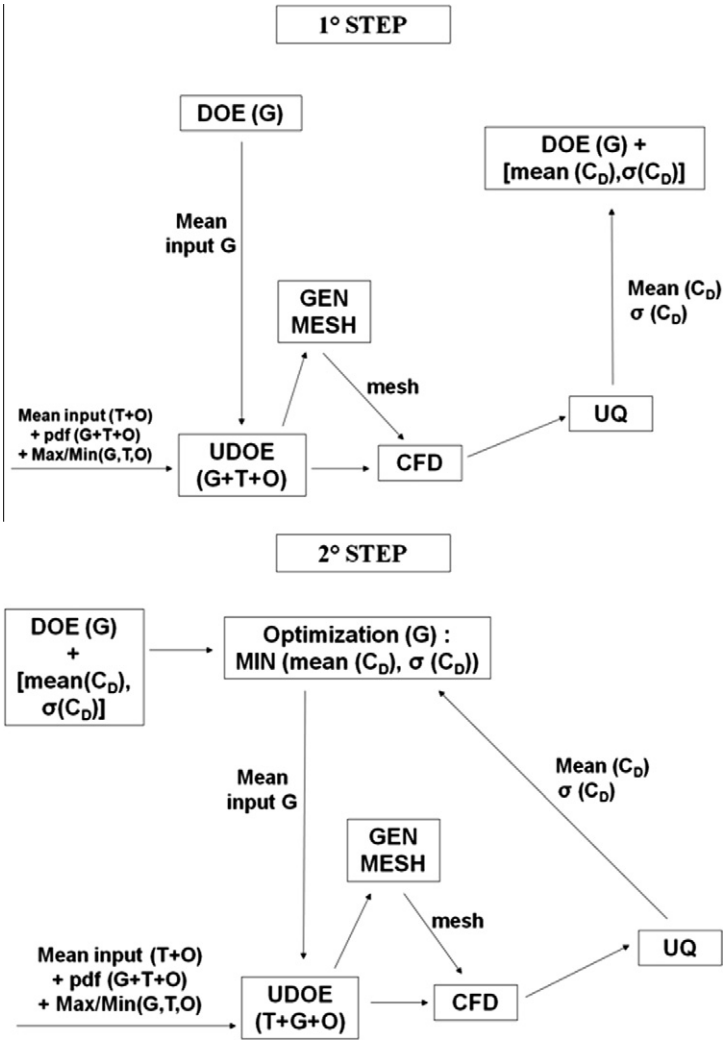

Fig. 5. Fully coupled approach for optimization problem.

using specific sets of values for input_T, input_G and input_O, which are chosen as reported in Section 2.2. For a given shape in the previously prescribed domain of geometric definition, a single evaluation corresponds to a particular combination of input_G, input_O and input_T; the global set containing all the possible variations of the uncertain inputs is called Uncertain Design of Experiment (UDOE). Three input data define the UDOE: (i) the mean values of input_G, input_T and input_O parameters (the mean values of input_T and input_O are fixed throughout the optimization problem), (ii) the maximal variations associated with these parameters (which have been specified in Section 2.1), (iii) the density probability distribution associated with each of these parameters. The UDOE is a set of combinations for input_G, input_T and input_O, determined by the quadrature formulae used to compute the PC coefficients. For every combination, a CFD evaluation is performed giving a computed C_D value; when such a value has been obtained for each element of the UDOE, the mean value and variance of the drag coefficient are estimated using the UQ tools described in the previous section.

A shape optimization procedure including uncertainties is made of two steps schematized in Fig. 5. The first step is the generation of a DOE for the variables of the optimization problem; in the present case a set of initial mean values for input_G is generated in the solution space S. Using the UQ tools described in the previous section, a UDOE is built for each value of input_G in the initial DOE; a grid is then generated for each distinct value of input_G and the flow is computed on each of this grid using the CFD code with the corresponding values of input_T and input_O in the UDOE. The mean value and variance of the drag coefficient are computed from the set of values associated with the UDOE and once this computation has been performed for all the values of input_G in the initial DOE the first generation for solving (20) is available. The second step corresponds to the optimization process itself. In the present work the NSGA algorithm [59] is applied to obtain the optimal Pareto set associated with (20). The main tuning parameters of the algorithm are the population size, the number of generations, the crossover and mutation probabilities pc, pm and the so-called sharing parameter σ used to take into account the relative isolation of an individual along a dominance front. Typical values for pc, pm are, respectively, 0.9 and 0.1; values of σ are retained following a formula given in [59] that takes into account the population size and the number of objectives. Using the values of the objective functions $mean(C_D)$, $\sigma(C_D)$ and selection, cross-over, mutation genetic operators, a new population of mean values for the input_G parameters is generated in S. For each member of the population (viz. airfoil shape) a UDOE is built, which differs from the ones previously used by the values of input_G only since the mean values of input_T and input_O are fixed once for all in the present problem. A grid is generated for each new shape and a CFD run is performed for the corresponding input_T, input_O in the UDOE. The quantities $mean(C_D)$, $\sigma(C_D)$ are then computed using the available UQ tools and a new population of potentially improved individuals is obtained which is evolved by applying the genetic operators. This procedure is repeated until a convergence criterion is satisfied (typically the variation of the Pareto set is required to decrease below a prescribed threshold).

2.3.2. Fully coupled approach

In a fully coupled approach, the above optimization procedure is directly applied, which means the global cost of the process is given by the unit cost of a CFD run yielding a value for C_D, multiplied by the total number of CFD runs performed. This last number is equal to the number of different input_G present in the initial DOE (also equal to the size of the evolving population of airfoil shapes) multiplied by the number of generations for the genetic optimization process multiplied by the number of different input sets forming the UDOE, so that:

$$\text{total cost} = (\text{unit CFD cost}) \times (\text{size of initial population}) \times (\text{number of generations}) \times (\text{size of the UDOE}).$$

With an initial population made of 20 individuals evolving during 30 generations and a 15-min CFD run to achieve steady-state for a dense gas flow over an airfoil, the cost of the optimization without uncertainty would be 150 h or a bit more than 6 days; if nine uncertainties are taken into account with a third-order PC, the size of the UDOE will be $4^9 \approx 200$ thousand sets of inputs hence a totally unacceptable total cost, even with the use of massive parellelization. To achieve reasonable computational times, specific strategies are needed to reduce the number of evaluation for the uncertainty quantification and/or the number of individuals computed throughout the optimization.

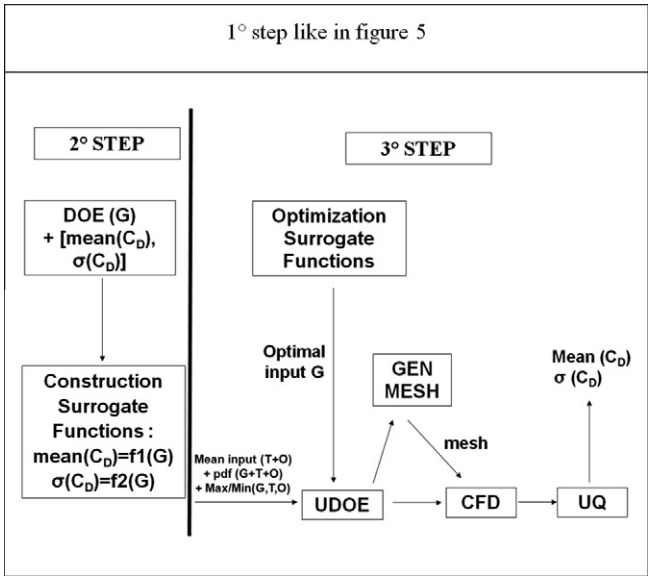

Fig. 6. Approach using surrogate models for optimization problem.

2.3.3. Approach using surrogate models

As a first way to reduce the computational cost of a fully coupled approach, a preliminary analysis on the baseline sonic arc is performed which allows to assess the contribution of each uncertain parameter to the drag coefficient variance (ANOVA-type sensitivity analysis) so as to retain only the most influential ones in the UDOE used within the optimization process. A second way to achieve this CPU reduction is the massive use of surrogate models. The optimization process relying on surrogate models is also decomposed in two steps (see Fig. 6): the first step is identical to the one described for the fully coupled approach but with a size of UDOE which can be reduced thanks to the previous sensitivity analysis and with a size of DOE that can be different (larger in particular) because it will only slightly impact the global cost of the process; in the second step, the DOE of mean values for input_G or DOE (G) with their associated computed values for the mean and variance of C_D is used in order to build a surrogate function for each of these objectives: mean $(C_D) = f_1(\text{input\_G}) = f_1(\{y_1, y_2, y_3\})$, variance $(C_D) = f_2(\text{input\_G}) = f_2(\{y_1, y_2, y_3\})$ where $\{y_1, y_2, y_3\}$ corresponds to a (mean) value in the S solution space. These simple mathematical functions (approximately) account for the uncertainties effects with a very modest computational cost. A Pareto optimal set for the simultaneous minimization of f_1 and f_2 is readily obtained. For every member of this optimal set, the UQ tools can then be used

to compute a correct final estimate for mean (C_D) and variance (C_D), following the approach previously described in Section 2.3.1 and used for every population individual at each generation in the fully coupled approach of Section 2.3.2. Since the final population produced by NSGA is usually entirely distributed along the global Pareto front, the total cost of the approach can be estimated as:

$$\text{total cost} = (\text{unit CFD cost}) \times (\text{size of initial DOE}) \times 2 \times (\text{size of the UDOE}),$$

where the factor 2 corresponds to the initial computations of Mean (C_D) and Variance (C_D) and to the final a posteriori computations. The cost of optimization for the surrogate functions f_1, f_2 is negligible with respect to the unit CFD cost. Note however the size of the UDOE must be reduced for the calculation to remain tractable hence the total number of uncertain parameters must be reduced as much as possible and the preliminary ANOVA analysis remains a compulsory step.

3. Preliminary Anova-based screening

3.1. General strategy

When computing the dense gas flow over the baseline sonic arc, nine parameters are considered uncertain, which are classified into three distinct sources: the three design parameters input_G describing the geometry, the three parameters input_O defining the operating condition for the flow and the three parameters input_T closing the gas thermodynamic model. The mean values, maximal variations and pdf type for each parameter are summarized in Table 2. The deterministic pressure field for the baseline sonic arc (corresponding to the mean value of input_G) with operating conditions input_O and thermophysical properties input_T fixed equal to their mean values is computed with the CFD solver and plotted in Fig. 7. The computed drag coefficient in this reference case is $C_D = 4.05E-02$. The analysis of the contribution of the nine uncertainties to the variance of this drag coefficient is first performed following a decoupled analysis: a single source of uncertainty (input_T or input_G or input_O with three parameters in any case) is taken into account, the two other sources being held constant equal to the respective mean values of their six parameters. The contribution of the three-parameter source of uncertainty is analyzed using the ANOVA approach described in Section 2.2.3, which yields the individual contribution of each uncertain parameter as well as the possible contribution of interacting parameters.

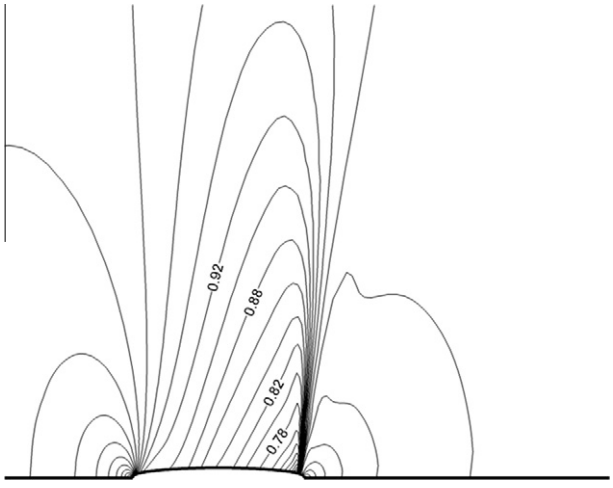

Fig. 7. Deterministic pressure field for the baseline sonic arc.

This analysis is carried out with a third-order PC, the coefficients of which are computed with a UDOE of size $4^3 = 64$. This UDOE is generated assuming successively a Gaussian and a Uniform pdf for the uncertain parameters.

In a second step, all the sources of uncertainty (input_T and input_G and input_O) are simultaneously taken into account in order to assess possible interactions between the geometry, the operating conditions and the thermophysical parameters that might contribute to the variance of C_D. If the UDOE is generated using again a third-order PC, its size will grow up to $4^9 = 262{,}144$. In view of this extremely high computational cost, it is decided to resort to the Sparse Grid Method (SGM) described in Section 2.2.2 to perform this fully coupled analysis. To improve the reliability of this analysis, the SGM is applied not only with a third-order polynomial (yielding a UDOE of size 871) but also with a fourth-order one (yielding a UDOE of size 3463), allowing an assessment of the convergence and accuracy of the statistical analysis.

3.2. Results of the decoupled analysis

When input_T is considered as the sole source of uncertainty, the variance of C_D is 5.8E−06 (resp. 5.7E−06) for a Gaussian (resp. Uniform) pdf. The contributions of the parameters $c_{v\infty}(T_c)/R$, ω and *nexp* to this variance, as computed using the ANOVA approach, are plotted in Fig. 8. It can be first observed the results depend only very weakly on the pdf type (0.4% difference between Gaussian and Uniform for $c_{v\infty}(T_c)/R$, 0.9% for ω). Uncertainty is mainly caused by the parameter $c_{v\infty}(T_c)/R$ (≈76.5% of the variance explained) and the parameter ω (≈23.4% of the variance explained) while the parameter *nexp* is non-influent. The weight of the caloric component with respect to the thermal component when dealing with an uncertain equation of state agrees with previous observations [14]. Since the first order effects of $c_{v\infty}(T_c)/R$ and ω explain 99.8% of the variance, the stochastic model for the drag coefficient displays an "additive" form, i.e. the interactions are not influent (that is why they are not reported in Fig. 8). The contours of the pressure coefficient maximal standard deviation (square root of the variance) in the computational domain are plotted in Fig. 9, along with the mean distribution and error bar for the pressure coefficient distribution along the baseline sonic arc. It can be concluded the effect of uncertainty on input_T is concentrated in the vicinity of the shock location.

When input_O is considered as the sole source of uncertainty, the variance of C_D is 3.7E−04 (resp. 3.9E−04) for a Gaussian (resp. Uniform) pdf. The contributions of the operating Mach number, pressure and density to this variance, as computed using the ANOVA approach, are plotted in Fig. 10. The difference between Gaussian and Uniform distribution remains slight: 1.2% for the Mach, 7% for the pressure, and 5.1% for the density. The operating Mach

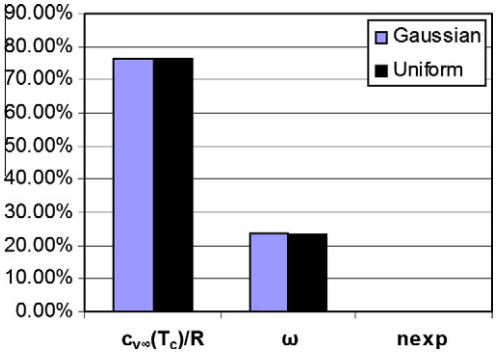

Fig. 8. Analyse of variance for the stochastic dense gas simulation when uncertainties on thermodynamic models (input_T) are taken in account.

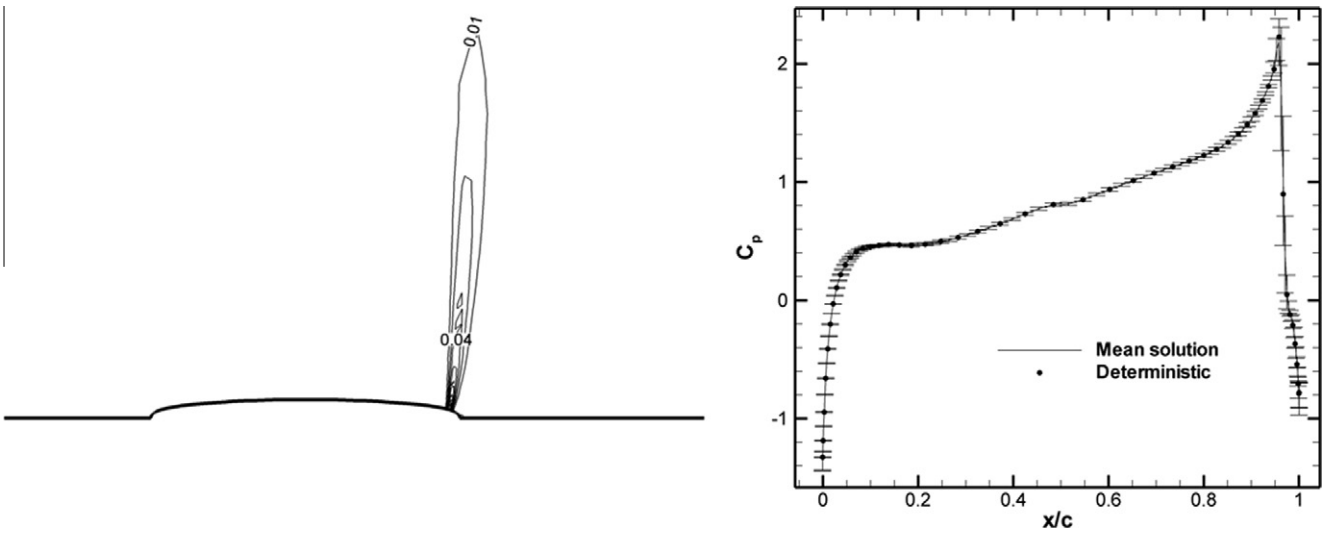

Fig. 9. Standard deviation of the pressure field (max = 0.095) on the left, mean solution and error bar for the Cp at the wall on the right, uncertainties on thermodynamic model (input_T).

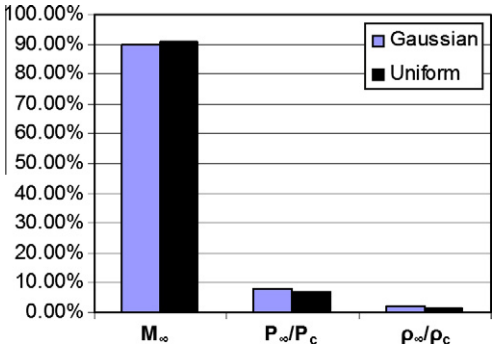

Fig. 10. Analyse of variance for the stochastic dense gas simulation when uncertainties on operating conditions (input_O) are taken in account.

number is clearly the most influential parameter since it explains 89.9% of the variance for the gaussian pdf. The strong influence of this parameter is easily understood from Fig. 2: when the operating Mach number varies in the range [0.9405,0.9595] the dense gas flow is precisely in the drag divergence region where the shock location is highly sensitive to the inflow conditions. The operating pressure and density contributes respectively to 7.8% and 1.9% of the variance (for Gaussian pdf). The first order effects on Mach, pressure and density explain 99.9% of the variance, clearly indicating the stochastic model for the drag coefficient displays an additive form, with no influence of interactions between parameters. The contours of the pressure coefficient maximal standard deviation are plotted in Fig. 11, along with the mean distribution and error bar for the pressure coefficient distribution along the baseline sonic arc: the effect of uncertainty on input_O is concentrated in the vicinity of the shock location. There is also an effect of uncertainty near the trailing and leading edge, even if much reduced with respect to the shock. This effect is not observed when uncertainties on input_T and input_G are considered.

When input_G is considered as the sole source of uncertainty, the variance of C_D is 2.1E−08 (resp. 2.3E−08) for a Gaussian (resp. Uniform) pdf. The contributions of the parameters y_1, y_2, y_3 to this variance, as computed using the ANOVA approach, are plotted in Fig. 12. The first order effects of y_1, y_2, y_3 explain 99% of the variance, hence an additive form for the stochastic model of drag coefficient can be concluded. The parameter y_1 is found to be a bit more influential than y_2 and y_3 but the effect of these latter parameters cannot be neglected. The contributions depend more significantly on the pdf but the keypoint remains the lack of dominant parameter, which can be related to the reduced number of geometrical parameters considered for the shape parameterization, making significant the contribution of each parameter. The contours of the pressure coefficient maximal standard deviation are plotted in Fig. 13, along with the mean distribution and error bar for the pressure coefficient distribution along the baseline sonic arc: the effect of uncertainty on input_G is important in the vicinity of the shock location but more distributed along the airfoil than with input_O and input_T.

Table 3 summarizes the computed mean value, variance and coefficient of variation (the ratio of the standard deviation to the mean) of the drag coefficient when considering successively input_T, input_O and input_G as sole source of uncertainty, with an associated Gaussian or Uniform pdf. It is essential to observe the variance associated with input_O is two orders of magnitude larger than the one associated with input_T and four orders of magnitude larger than the one associated with input_G. This same hierarchy is also observed in Figs. 9, 11, 13 when analyzing the contour levels of the pressure coefficient variance or the error bars of the wall pressure distribution: the influence of geometrical uncertainties appear negligible with respect to the other uncertainties, with the uncertainties on operating conditions clearly dominant. The computation of the coefficient of variation is important to estimate how uncertainties affect the prediction of numerical simulation when compared for example to the discretization error. As showed in Table 3, uncertainties associated to input_T and input_O affect the drag coefficient computation much more than the discretization error.

3.3. Results of the coupled analysis

The three sources of uncertainties, corresponding to nine uncertain parameters, are now simultaneously taken into account using the SGM. A uniform pdf is assumed for the geometrical parameters of input_G, corresponding to a worst case scenario where all the parameters are equally important; a Gaussian distribution is assumed for input_T and input_O since the previous analysis has shown the difference between a Gaussian and Uniform pdf remains

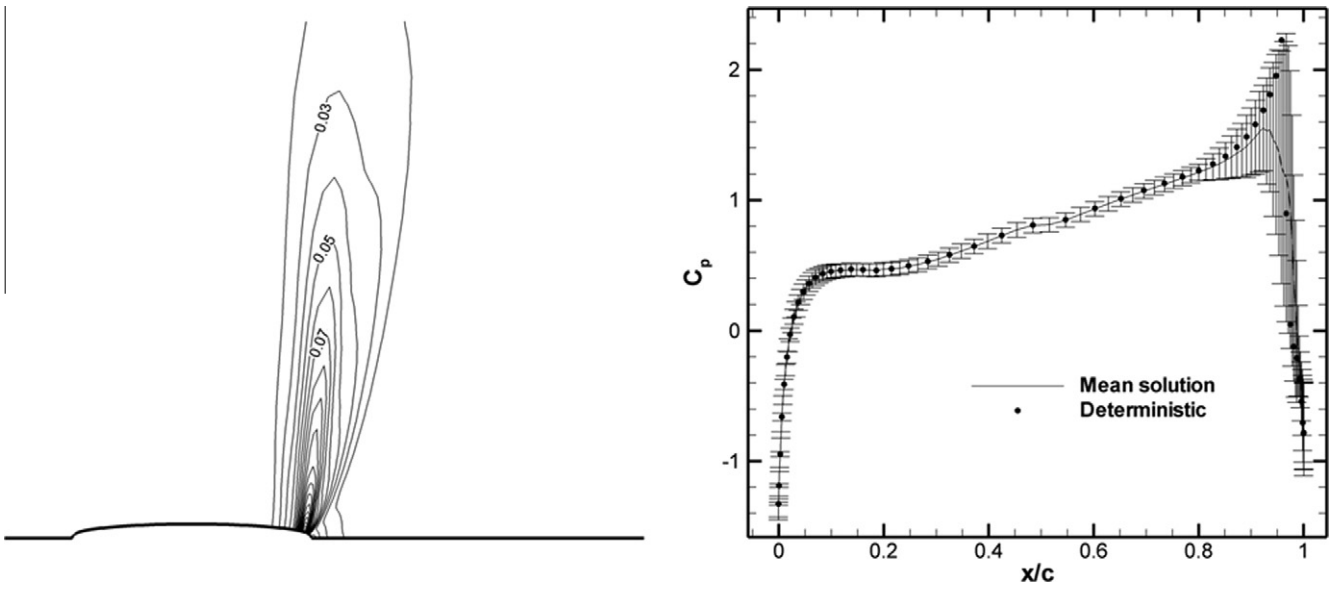

Fig. 11. Standard deviation of the pressure field (max = 0.203) on the left, mean solution and error bar for the Cp at the wall on the right, uncertainties on operating conditions (input_O).

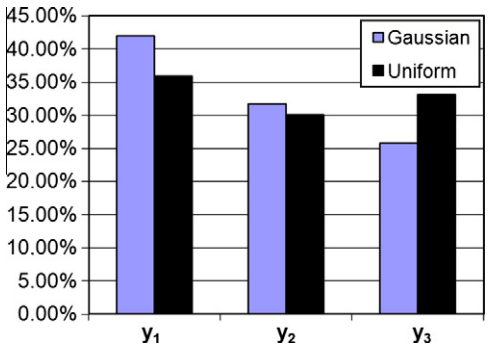

Fig. 12. Analyse of variance for the stochastic dense gas simulation when uncertainties on geometry (input_G) are taken in account.

slight. The ANOVA analysis is performed for a third-order and fourth-order polynomial; the hierarchy of the most influential parameters and the contribution of each uncertainty to the global variance is analyzed. The results obtained with the two plans are very similar: the hierarchy is identical with minor differences in the decomposition of the variance as reported in Fig. 14. The fourth-order analysis is displayed in Fig. 15 and yields the following comments: the uncertainty on the operating Mach number is the most influential with a contribution to more than 89% of the variance; the first-order effects on the operating Mach number, pressure and density explain 99.5% of the variance. The drag coefficient stochastic model is of "additive" form with negligible interaction effects. Going back to the third-order/ fourth-order comparison, the difference for the three most dominant parameters (input_O) does not exceed 10% so that the results of the ANOVA analysis can be considered as converged. The uncertainties on

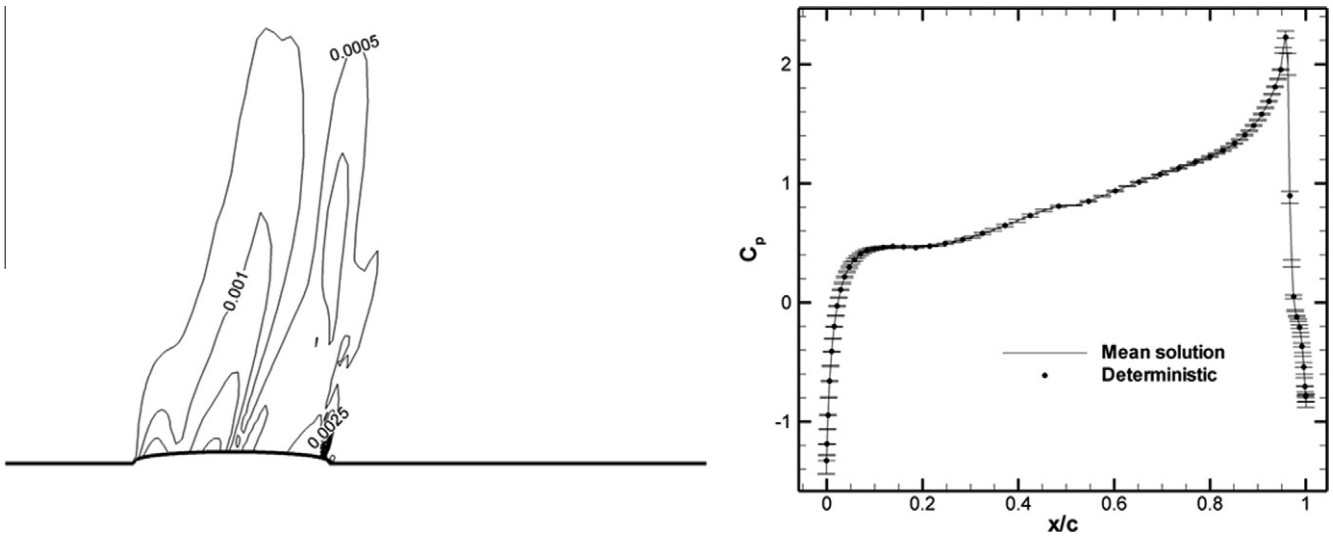

Fig. 13. Standard deviation of the pressure field (max = 0.013) on the left, mean solution and error bar for the Cp at the wall, uncertainties on the geometry (input_G).

Table 3
Mean, variance and coefficient of variation for each phase of the decoupled analysis (isolated uncertainties on input_T, input_O and input_G, respectively).

	Phase 1	Phase 2	Phase 3
Mean (Gaussian, uniform)	(4.05E−02; 4.05E−02)	(4.13E−02; 4.14E−02)	(4.05E−02; 4.05E−02)
Variance (Gaussian, uniform)	(5.76E−06; 5.74E−06)	(3.70E−04; 3.90E−04)	(2.12E−08; 2.29E−08)
Coefficient of variation (Gaussian, uniform)	(5.93%; 5.92%)	(46.57%; 47.7%)	(0.36%; 0.37%)

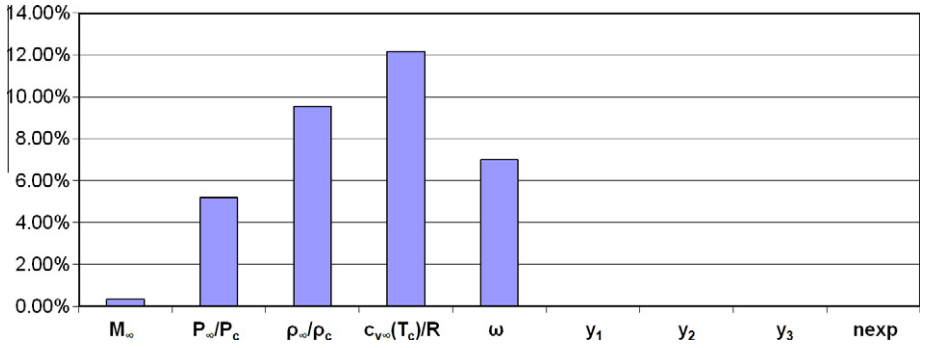

Fig. 14. Analyse of variance for the stochastic dense gas simulation when uncertainties on geometry (input_G), operating conditions (input_O), and thermodynamic models (input_T) are taken in account (coupled analysis), difference (%) on the contribution to the variance between the order 3 and the order 4 with the Sparse Grid Method.

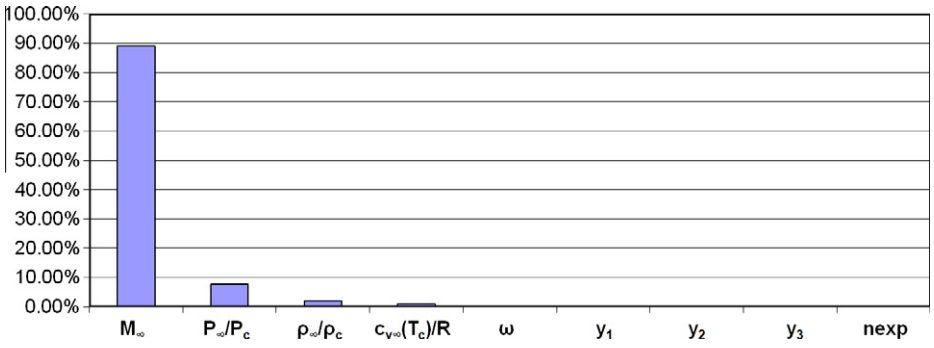

Fig. 15. Analyse of variance for the stochastic dense gas simulation when uncertainties on geometry (input_G), operating conditions (input_O), and thermodynamic models (input_T) are taken in account (coupled analysis, Sparse Grid Method order 4).

the geometry (input_G) and the thermodynamic model (input_T) produce negligible effects with respect to the operating conditions.

3.4. Conclusions

The decoupled and coupled analysis of the contribution to the variance for the set of uncertain parameters involved in the dense gas flow computation allow to draw important conclusions regarding the selection of parameters that need to be taken into account when performing a robust shape optimization for this dense gas flow. In the present case, geometrical uncertainties are much less influential than the other sources of uncertainties hence input_G will be considered as deterministic in the optimization performed in the next section, with geometrical parameters rigorously equal to their mean values, varying in the previously defined solution space S. The influence of the uncertainty on the thermophysical properties input_T remains also limited, with respect to that on the operating conditions input_O. Note this hierarchy is likely to depend on the choice of equation of state: the PRSV model has been found in [14] to be less sensitive than other models to uncertainties on its parameters and the present conclusion is consistent with these previous findings. In the forthcoming optimization process, input_T will also be considered as deterministic, with its three parameters fixed equal to the mean values given in Table 2.

4. Shape optimization under uncertainty

4.1. Optimization using surrogate functions

The preliminary analysis described in the previous section showed that the number of uncertainties that should be evaluated during the optimization procedure could be reduced to the three uncertainties on the operating conditions (input_O). Since input_T is fixed, it no longer appears in the schematic view of the optimization process displayed in Fig. 15. Similarly, the set of mean values for input_G given by the initial DOE (G) is directly used to generate a population of airfoil shapes, without impacting the UDOE, the sole input of which is given by the mean values for input_O, hence denoted UDOE (O). The remainder of the first step is identical to the one described in Fig. 6, with the generation of (mean value, variance) couples for the drag coefficient associated with each element of DOE (G). The second step is also left unchanged with the design of surrogate functions f_1(input_G), f_2(input_G) for mean (C_D) and variance (C_D). The third step of a posteriori UQ assessment for optimal shapes undergoes the same simplification as the start of the first step: the mean value and variance of the drag coefficient for an optimal shape are computed using the UQ tools on UDOE (O). The proposed strategy of optimization taking into account uncertainties involves three types of

choices: the method for generating DOE (G), the method for deriving the surrogate functions and the UQ method for generating UDOE (O). Three distinct strategies corresponding to three different combinations of methods have been considered in order to assess to what extent the optimal values obtained for input_G are independent from these choices.

In strategy 1, a Quasi Monte Carlo (QMC) geometric DOE (G) with 20 individuals is used given by the toolbox Nisp [55]; the UDOE (O) is built from a third-order Sparse Grid Method, using Nisp again. QMC is similar to the usual Monte Carlo simulation but uses quasi-random sequences instead of (pseudo) random numbers; these sequences are deterministic and fill the space more uniformly than random sequences (known as low discrepancy sequences, i.e. the Sobol sequences). The surrogate functions f_1, f_2 are built from the values (mean (C_D), variance (C_D)) associated with DOE (G) using a multi-layered perceptron (MLP) neural network, with a cross validation method enabling to select 4 hidden units so that each neural network has a $3 \times 4 \times 1$ architecture.

In strategy 2, a QMC geometric DOE (G) with 40 individuals and a UDOE (O) also based on a QMC distribution are considered. The surrogate functions are derived using the same MLP than for strategy 1.

In strategy 3, DOE (G) is generated using a Box Wilson DOE (based on fractional factorial design) of 15 individuals and UDOE (O) is generated using the same Sparse Grid Method as strategy 1. The surrogate functions are derived using the same MLP than for strategy 1.

The set of optimal solutions obtained using these three strategies are displayed in Fig. 16, both in the objective plane (f_1,f_2) and in the parameter planes (y_1,y_2) and (y_1,y_3). Let us recall the solution space for $\{y_1, y_2, y_3\}$ is defined as $[0.04049, 0.04949] \times [0.03646, 0.04457] \times [0.05444, 0.06654]$; the Pareto fronts in the parameter space are very similar: they correspond to the upper limit of variation for y_1 and y_2 with the third parameter y_3 varying on $[0.0544, 0.0599]$ for strategy 3, on $[0.0575, 0.0600]$ for strategy 1 and $[0.0540, 0.0570]$ for strategy 2.

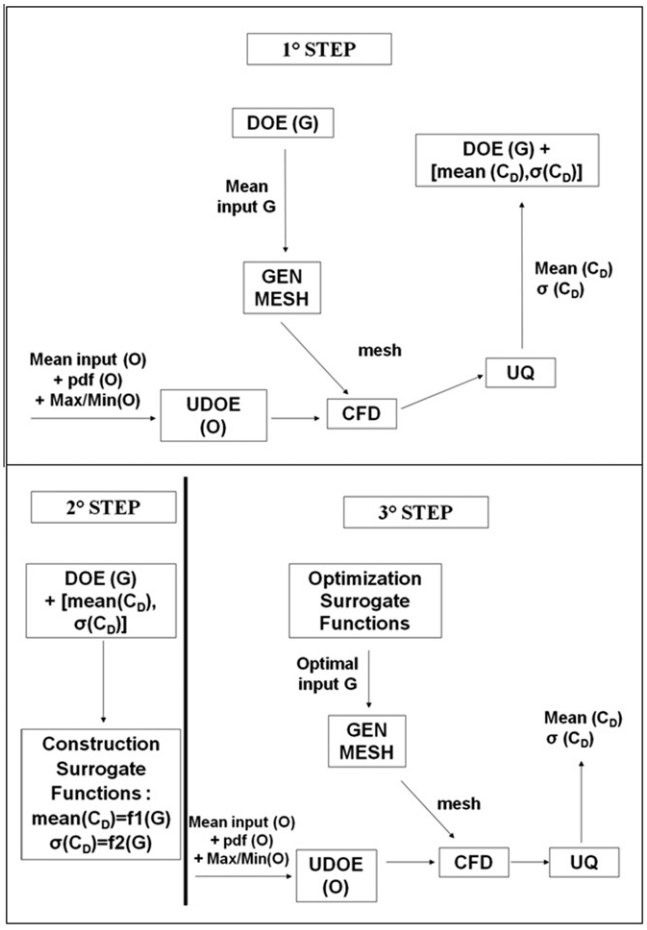

Fig. 16. Approach using surrogate models for optimization problem after preliminary analysis.

4.2. A posteriori UQ analysis for selected optimal shapes

Three individuals are selected on each Pareto front in the objective plane, corresponding respectively to the minimum of f_1, the minimum of f_2 and a point at mid-distance between the previous ones. In the final step of the optimization, an a posteriori UQ is performed for these individuals, that is (f_1,f_2) are replaced by (mean (C_D), variance (C_D)) computed using an UDOE (O) based on the Sparse Grid Method with a fourth-order polynomial (83 CFD evaluations are performed for each geometry to quantify the effects of the three uncertainties input_O) or on the Quasi Monte Carlo method (with 64 CFD evaluations for each geometry). The use of two distinct strategies for computing the mean and variance of the drag coefficient aims at offering some cross-validation for the conclusions that will be drawn regarding the proposed optimal shapes. These choices of UDOE (O) are also applied to the baseline sonic arc to compute the mean value and variance associated with this initial geometry. Besides, a classical "deterministic" minimization of the drag coefficient is also performed, where input_T and input_O are equal to their mean values given in Table 2 and input_G vary in the parameter space S. This single-objective minimization of the drag coefficient is performed using the same genetic algorithm as the one used for the bi-objective minimization of (mean (C_D), variance (C_D)). The geometric parameters defining this "classical" optimum are reported in Fig. 17 along with the Pareto fronts for strategy 1, 2 and 3 in the parameter space. It is interesting to note this classical optimum seems to belong to these Pareto fronts; however, it is mandatory to carry out the analysis in the plane of the real (and not surrogate) objective functions, which is done in Figs. 18 and 19, respectively, using SGM and QMC for computing (mean (C_D), variance (C_D)). The nine individuals (three on each Pareto front for strategy 1, 2 and 3) are numbered in Figs. 17–19, in order to see exactly how the Pareto points move in the objective plane in Figs. 18 and 19 with respect to the optimization of the surrogate functions in Fig. 17. Though all the optimal shapes produced by Strategy 1, 2 and 3 as well as the deterministic optimization dominate or are not dominated by the baseline configuration, it must be noted the surrogate functions (f_1,f_2) introduce a significant error level with respect to the "exact" objective functions. In particular, the shapes corresponding to a Pareto set in the (f_1,f_2) plane are no longer forming such a Pareto front in the (mean (C_D), variance (C_D)) plane. A keypoint however is that a global Pareto front can be obtained by gathering shapes produced by the optimization strategies with uncertainties and the deterministic shape. This latter solution correspond to a minimal value for mean (C_D) along the Pareto front but a maximal value for variance (C_D), where the deterministic C_D is taken into account in the deterministic optimization process. Strategy 2 and 3 provide in particular the same robust optimum corresponding to a minimal value of variance (C_D) along the Pareto front but to a maximal value for mean (C_D), remaining however below the mean value provided by the baseline configuration. The same conclusions can be drawn from both UDOE (O) (generated either by SGM or by QMC): the maximum difference on the mean drag prediction is 6% and 4.5% on the variance prediction. Moreover, the global Pareto fronts featured in Figs. 18 and 19 are composed of the same individuals (1, 4, 5, 7 and the classical optimum), though one is concave while

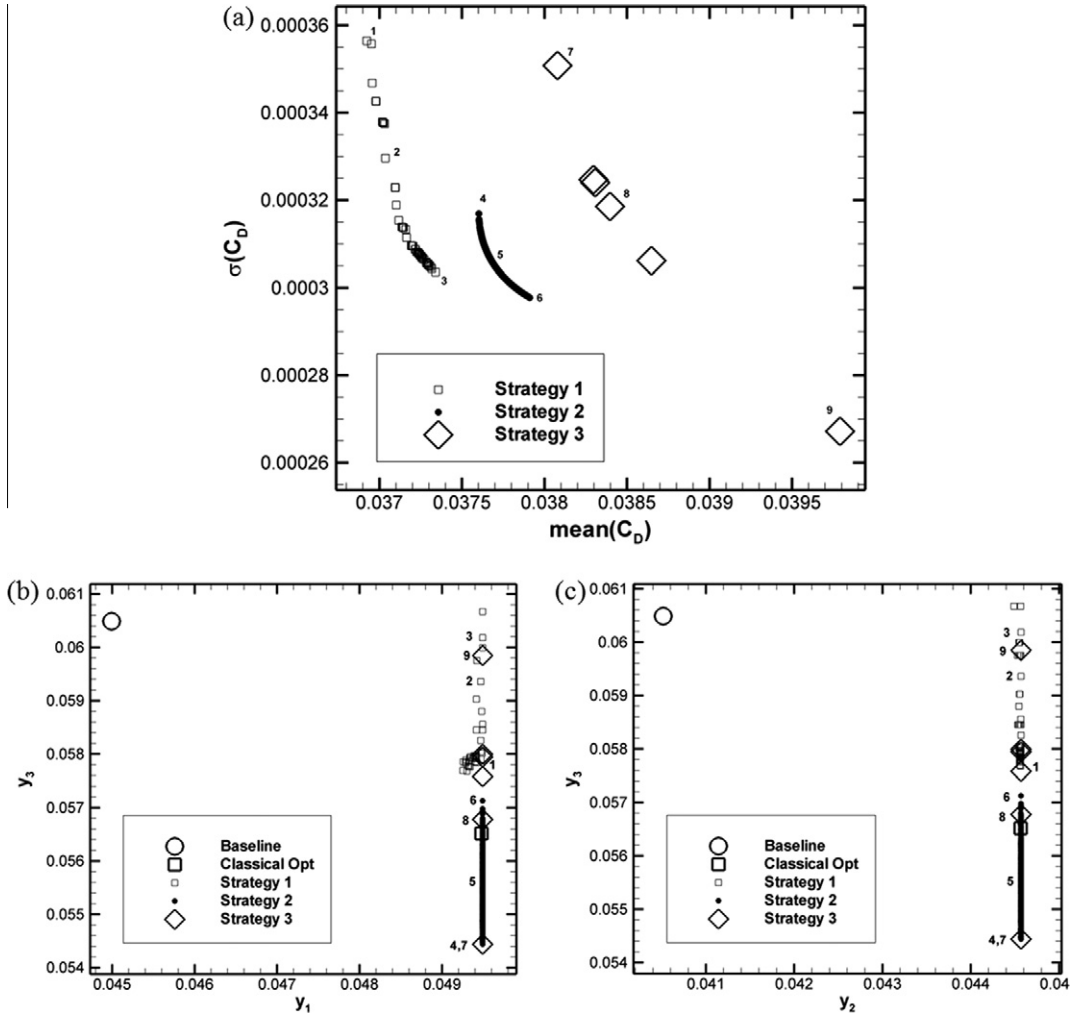

Fig. 17. Mean and variance of the drag coefficient (a), and position of Pareto front in the parameters plan y_1–y_3 (b) and y_2–y_3 (c), of the optimal individuals issued from the different strategies.

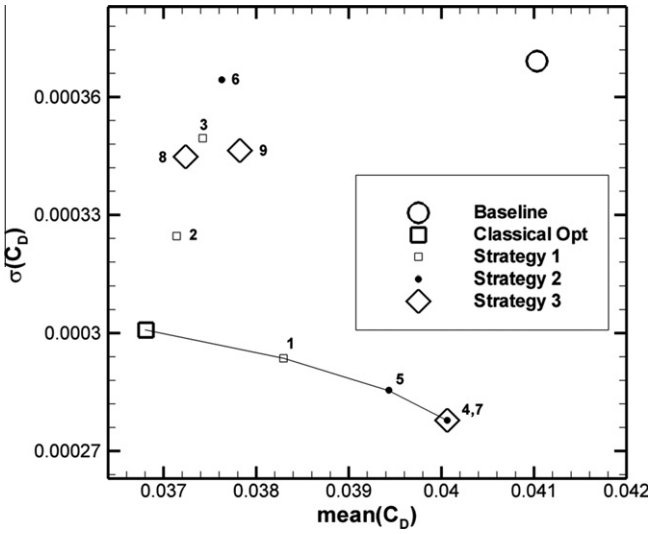

Fig. 18. Comparison of different optimal individuals by using an UDOE based on the Sparse Grid Method, connected points represent the global Pareto front.

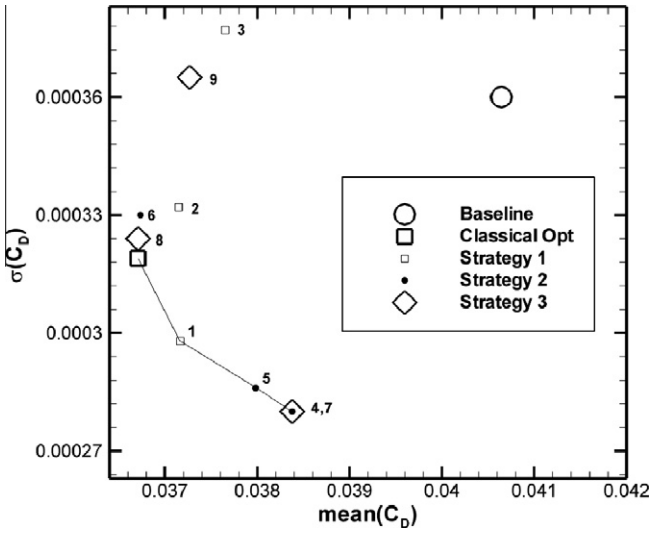

Fig. 19. Comparison of different individuals by using an UDOE based on the Quasi Monte Carlo Method, connected points represent the global Pareto front.

the other is convex. This difference in the mean and variance drag prediction can be explained by considering the specific behaviour of dense gas flows. When uncertainties on thermodynamic model and operating conditions are considered at the same time, the free-stream Mach number can be close to the critical Mach number for some particular uncertain inputs. In this case, values of C_D can be

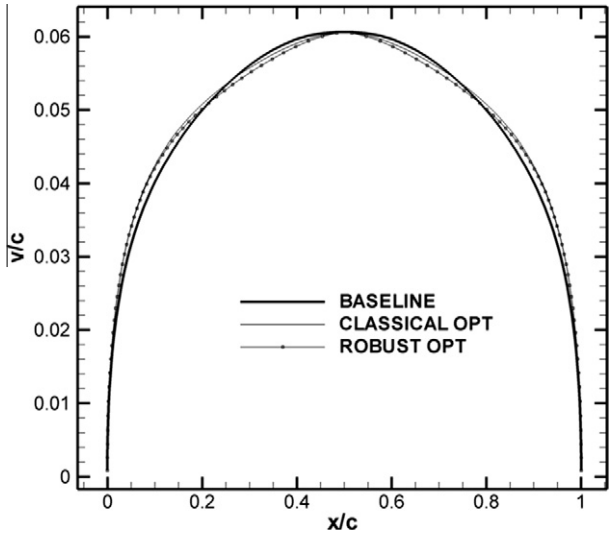

Fig. 20. Some geometries of interest: baseline sonic arc, optimal shape produced by the classical (deterministic) optimization and optimal shape yielded by the robust optimal design.

strongly reduced, that makes statistics hard to compute. Then, mean and variance computations depend on the size of UDOE, that is why two strategies can give slightly different results. A Monte Carlo computation would be necessary in order to have a reference results for mean and variance, but it is unfeasible given the computational cost.

Several remarks can be made by comparing the different strategies. Concerning strategies 1 and 2, enriching the DOE makes the surrogate functions more predictive, then the approximation of mean and variance (as function of parameters y_1,y_2 and y_3) is improved. In fact, with strategy 2 two individuals belonging to Pareto front are obtained (individuals 4 and 5), as shown in Figs. 18 and 19. Seeing the complexity in mean and variance computations, two strategies (1 and 3) can give slightly different results if few design points are considered. Geometries of the Pareto front are very similar, with differences on y_3 parameter, while y_1 and y_2 are the same. The parameter y_3 is equal to 0.0544 for individuals 4 and 7, to 0.0556 for 5, to 0.0565 for the classical optimum and to 0.0578 for individual 1. Along the Pareto front, lower y_3 is linked with an increase of mean (C_D). Geometry of individual 5 is very similar to geometries 4–7 (difference up to 0.4%, where difference between two geometries A and B is computed using the formula $\left(\sum_N |y_{A_i} - y_{B_i}|/y_{A_i}\right)/N$ and N is the number of discretization points equally spaced along x (see Fig. 20)), while geometries 1-6-8 are similar to the classical optimum (differences up to 0.3%). Even if y_3 of individual 8 is close to those ones of individual 1 and of the classical optimum, individual 8 does not belong to the Pareto front. This means there is a non-linear behaviour in the proximity of y_3 of the classical optimum. This fact could be confirmed by means of Monte Carlo in order to have a better estimate for statistical quantities. The shapes corresponding to the classical optimum and the robust optimum (individuals 4 and 7) are plotted in Fig. 20, along with the baseline sonic arc (owing to the upstream/downstream and upper/lower symmetries of the airfoil only the upper upstream quarters of the airfoils are displayed). The mean and variance of the pressure contours are computed for each of these shapes using the UDOE (O) based on the Sparse Grid Method with a fourth-order polynomial: the superior stability offered by the robust optimum can be clearly observed in Fig. 21, with a 36% reduction of the maximum standard deviation with respect to the baseline sonic arc (against a 26% reduction offered by the classical optimum).

The computational costs of the various optimization processes (including or not the uncertainties) investigated in this work are reported in Table 4, with the unit cost taken as the CPU time needed to perform a single CFD run over an airfoil shape. The extra-cost associated with the detection of robust optimal shapes, ensuring a minimization of the variance of the deterministic objective, varies between two and four times the cost of a standard single objective.

5. Conclusions and perspectives

The predictive numerical simulation of a dense gas flow through an ORC turbine must take into account two types or sources of physical uncertainties: the physical properties of the fluid and the operating conditions at the turbine inlet. When designing the geometry of the turbine so as to optimize its efficiency, the geometrical uncertainties on the manufactured shape must also be accounted for. Shape optimization including the quantification of the effect of uncertainties means solving a multi-objective problem where, typically, the mean value and the variance of the initial objective are computed using uncertainty quantification tools. In this paper, the feasibility of performing such an optimization for realistic computational costs has been assessed on a simplified configuration (drag minimization under uncertainties for the BZT flow over an isolated symmetric airfoil) and a general procedure applicable to more complex problems has been proposed. Note that the present work did not investigate the uncertainty on the strongest BZT effects in the inversion region where rarefaction shock waves may occur because such effects are not relevant for the dense gas turbine design constituting the final long-term objective of the study. A polynomial chaos non-intrusive approach has been retained to take into account the different sources of uncertainties, with a Sparse Grid Method to further reduce the number of CFD evaluations needed for the UQ analysis. The global cost of the uncertain optimization has been reduced by using a three-step strategy: in a first preliminary step, the contributions of the uncertainties to the variance of the objective have been analyzed for the baseline configuration so as to retain only the parameter with the most important effects in the shape optimization process; in the second step, surrogate functions or metamodels have been derived for the mean value and variance of the objective and a fast minimization of these functions was performed to obtain a set of optimal geometric parameters; in the third and last step, selected individuals along the pseudo-(since defined in the plane of surrogate functions) Pareto fronts are a posteriori analyzed using the UQ tools. The mean value and variance of the drag coefficient associated with these shapes are compared with these same quantities for the baseline airfoil and for the so-called classical optimal shape obtained by performing a drag minimization without taking into account any uncertainty. A reliable Pareto front is then obtained in the (mean (C_D), variance (C_D)) plane, which includes at its upper left end (minimum mean value, maximum variance along the front) the classical optimum and at its lower right end (maximum mean value, minimum variance along the front) a robust optimum which reduces the drag of the baseline airfoil while minimizing the variance of this drag when uncertainties exist on the operating conditions. The price to pay for obtaining such a complete description of optimal shapes minimizing the drag coefficient with nine uncertainties initially taken into account is the sum of the number of CFD evaluations needed for the preliminary ANOVA screening (about 6000) and the number of CFD evaluations (about 6000) needed to build the surrogate functions once the most influential parameters have been selected and to perform the a posteriori UQ analysis on the potentially optimal shapes. Since the deterministic optimization, without

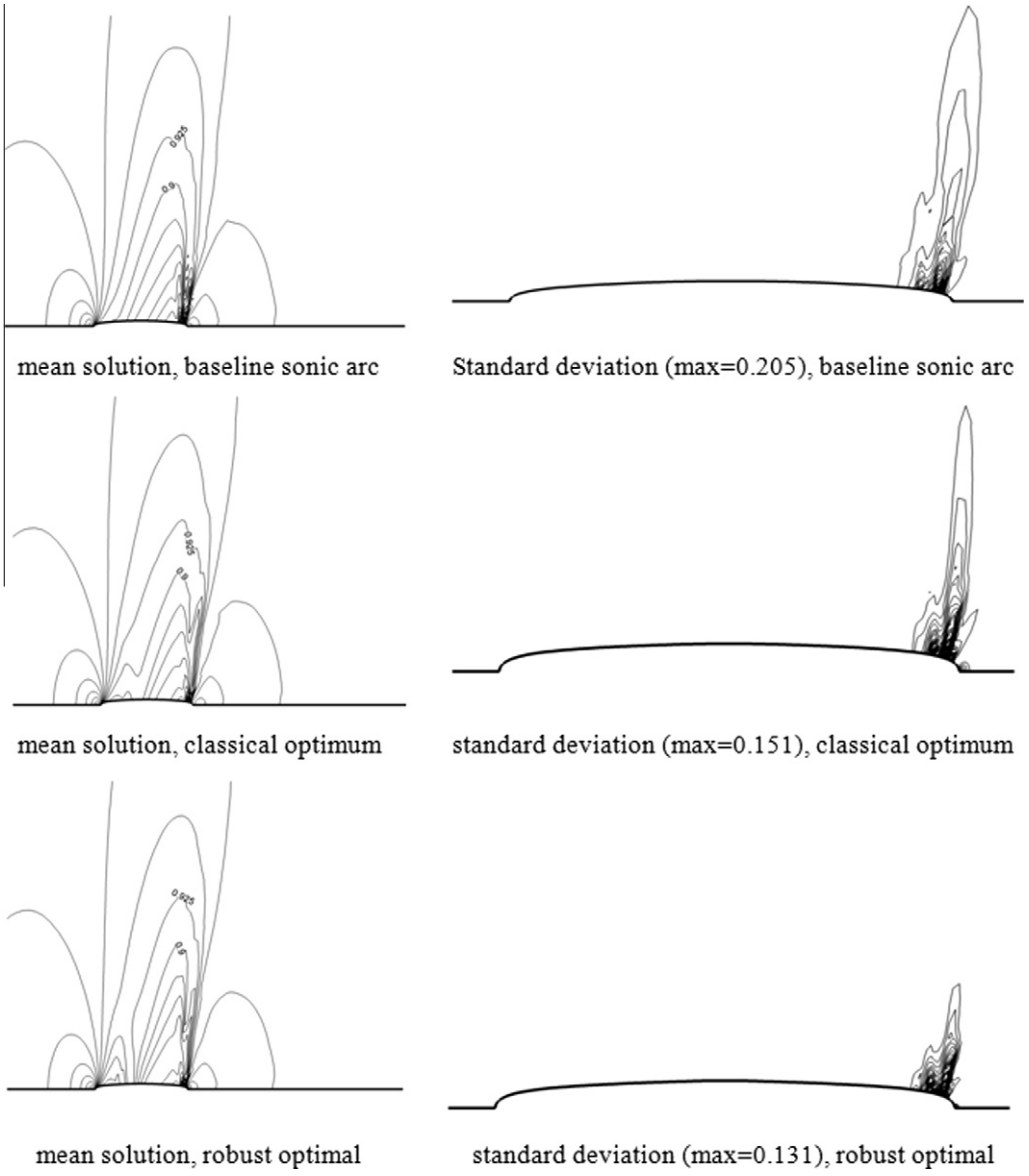

Fig. 21. Mean solution and standard deviation for the base sonic airfoil, individual of the classical optimization and optimization issued from the approach using surrogate models, uncertainty computed with Sparse Grid Method.

Table 4
Global number of CFD evaluations for the different strategies.

	Global number of CFD evaluations
Classical optimization	600
Strategy 1	1909
Strategy 2	2752
Strategy 3	1494

uncertainty, requires typically 600 CFD evaluations, the cost of an optimization including uncertainty can be estimated, in the present study, to 20 times the cost of a classical optimization, which remains perfectly acceptable. For the present study, with a unit CFD cost of 15 min, the overall computational cost for obtaining a set of robust optimal shapes is about 125 days, translated into a few days of processing when the fully independent CFD evaluations are performed on a parallel machine. The computer system available for the study allowed a typical number of eight parallel CFD computations, reducing the total optimization time to a bit more than two weeks. In future work, the present efficient approach and the fully coupled approach described in Section 2.3.2 will be tested on an even simpler quasi-1D configuration to further assess the reliability of the strategy followed in this study; besides, this strategy will also be applied to the optimal design of a 2D BZT turbine including thermophysical, operating and geometric uncertainties.

Acknowledgments

This work was partially supported by the French National Research Agency (ANR) in the context of the OPUS (Open source Platform for Uncertainty treatment in Simulation) project (Ref. ANR-07-CIS7–010 and ANR-07-TLOG-015).

References

[1] P.A. Thompson, A fundamental derivative in gas dynamics, Phys. Fluids 14 (1971) 1843–1849.

[2] P.M. Congedo, C. Corre, P. Cinnella, Airfoil shape optimization for transonic flows of Bethe–Zel'dovich–Thompson fluids, AIAA J. 45 (6) (2007).
[3] M.S. Cramer, L.M. Best, Steady, isentropic flows of dense gases, Phys. Fluids A 3 (1991) 219–226.
[4] M.S. Cramer, G.M. Tarkenton, Transonic flows of Bethe–Zel'dovich–Thompson fluids, J. Fluid Mech. 240 (1992) 197–228.
[5] J.F. Monaco, M.S. Cramer, L.T. Watson, Supersonic flows of dense gases in cascade configurations, J. Fluid Mech. 330 (1997) 31–59.
[6] B.P. Brown, B.M. Argrow, Application of Bethe–Zel'dovich–Thompson fluids in organic Rankine cycles, J. Propul. Power 16 (2000) 1118–1124.
[7] C.W. Wang, Z. Rusak, Numerical studies of transonic BZT gas flows around thin airfoils, J. Fluid Mech. 396 (1999) 109–141.
[8] M.S. Cramer, A. Kluwick, On the propagation of waves exhibiting both positive and negative nonlinearity, J. Fluid Mech. 142 (1984) 9–37.
[9] P. Cinnella, P.M. Congedo, Inviscid and viscous aerodynamics of dense gases, J. Fluid Mech. 580 (2007) 179–217.
[10] C. Zamfirescu, A. Guardone, P. Colonna, Admissibility region for rarefaction shock waves in dense gases, J. Fluid Mech. 599 (2008) 363–381.
[11] <http://www.turboden.eu/en/rankine>.
[12] A. Guardone, B.M. Argrow, Non classical gasdynamic region of selected fluorocarbons, Phys. Fluids 17 (2005). pp. 116102–116111, 17.
[13] P. Colonna, N.R. Nannan, A. Guardone, E.W. Lemmon, Multiparameter equations of state for selected siloxanes, Fluid Phase Equilibr. 244 (2006) 193–211.
[14] P. Cinnella, P.M. Congedo, V. Pediroda, L. Parussini, Quantification of thermodynamic uncertainties in real gas flows, Int. J. Engrg. Syst. Model. Simul. 2 (1/2) (2010) 12–24.
[15] J.K. Guest, T. Igusa, Structural optimization under uncertain loads and nodal locations, Comput. Methods Appl. Mech. Engrg. 198 (2008) 116–124.
[16] X. Guo, W. Bai, W. Zhang, X. Gao, Confidence structural robust design and optimization under stiffness and load uncertainties, Comput. Methods Appl. Mech. Engrg. 198 (2009) 3378–3399.
[17] C. Jiang, X. Han, G.R. Liu, Optimization of structures with uncertain constraints based on convex model and satisfaction degree of interval, Comput. Methods Appl. Mech. Engrg. 196 (2007) 4791–4800.
[18] C. Jiang, X. Han, G.P. Liu, A sequential nonlinear interval number programming method for uncertain structures, Comput. Methods Appl. Mech. Engrg. 197 (2008) 4250–4265.
[19] I. Doltsinis, Z. Kang, G. Cheng, Robust design of non-linear structures using optimization methods, Comput. Methods Appl. Mech. Engrg. 194 (2005) 1779–1795.
[20] L. Huyse, S.L. Padula, R. Michael Lewis, Wu Li, Probabilistic approach to free-form airfoil shape optimization under uncertainty, AIAA J. 40 (9) (2002).
[21] Y. Deremaux, N. Pietremont, J. Négrier, E. Herbin, M. Ravachol, Environmental MDO and Uncertainty Hybrid Approach Applied to a Supersonic Business Jet, AIAA 2008-5832.
[22] E.N. Tinoco, Validation and minimizing CFD uncertainty for commercial aircraft applications, AIAA 2008-6902, in: 26th AIAA Applied Aerodynamics Conference, 18–21 August 2008, Honolulu, Hawaii.
[23] R. Hassan, W. Crossley, Approach to discrete optimization under uncertainty: The population-based sampling genetic algorithm, AIAA J. 45 (11) (2007).
[24] D. Lucor, C. Enaux, H. Jourdren, P. Sagaut, Stochastic design optimization: Application to reacting flows, Comput. Methods Appl. Mech. Engrg. 196 (2007) 5047–5062.
[25] Sethuraman Sankaran, Stochastic optimization using a sparse grid collocation scheme, Probab. Engrg. Mech. 24 (2009) 382–396.
[26] P. Cinnella, P.M. Congedo, C. Corre, Shape optimization for dense gas flows through turbine cascades, in: H. Deconinck, E. Dick (Eds.), Proceedings of the Fourth International Conference on CFD, ICCFD4, Ghent, Belgium, 10–14 July, 2006, Springer, 2009.
[27] P. Cinnella, P.M. Congedo, Optimal airfoil shapes for viscous transonic flows of BZT fluids, Comput. Fluids 37 (2008) 250–264.
[28] A. Harten, P.D. Lax, B. van Leer, On upstream differencing and Godunov-type schemes for hyperbolic conservation laws, SIAM Rev. 25 (1) (1983) 35–61.
[29] E.F. Toro, M. Spruce, W. Speares, Restoration of the contact surface in the HLL Riemann solver, Shock Waves 4 (25) (1994) 25–34.
[30] H. Luo, J.D. Baum, R. Löhner, Extension of Harten–Lax–van Leer scheme for flows at all speeds, AIAA J. 43 (6) (2005) 1160–1166.
[31] P.M. Congedo, P. Cinnella, C. Corre, Efficient numerical simulation of dense gas flows past airfoils and wings, in: H. Choi, H.G. Choi, J.Y. Yoo (Eds.), Proceedings of the Fifth International Conference on CFD, ICCFD5, Seoul, Republic of Korea, 7–11 July 2008, Springer, 2009.
[32] B. Van Leer, Toward the ultimate conservative difference scheme, 5: A second-order sequel to Godunov's method, J. Comput. Phys. 32 (2) (1979) 101–136.
[33] T.J. Barth, D.C. Jespersen, The Design and Application of Upwind Schemes on Unstructured Meshes, AIAA Paper 89-0366, 1989.
[34] D. Sridar, N. Balakrishnan, An upwind difference scheme for meshless solvers, J. Comput. Phys. 189 (1) (2003) 1–29.
[35] C. Corre, X. Du, A residual-based scheme for computing compressible flows on unstructured grids, Comput. Fluids 38 (2009) 1338–1347.
[36] V. Venkatakrishnan, Convergence to steady-state solutions of the Euler equations on unstructured grids with limiters, J. Comput. Phys. 118 (1) (1995) 120–130.
[37] H. Luo, J. Baum, R. Löhner, A fast matrix-free implicit method for compressible flows on unstructured grids, J. Comput. Phys. 146 (2) (1998) 664–690.
[38] P. Glaister, An approximate linearised Riemann solver for the Euler equations for real gases, J. Comput. Phys. 74 (1988) 382–408.
[39] J. Hoffren, T. Talonpoika, J. Larjola, T. Siikonen, Numerical simulation of real-gas flow in a supersonic turbine nozzle ring, J. Engrg. Gas Turbine Power 124 (2002) 395–403.
[40] R.S. Miller, K.G. Harstad, J. Bellan, Direct numerical simulations of supercritical fluid mixing layers applied to heptane–nitrogen, J. Fluid Mech. 436 (2001) 1–39.
[41] A. Guardone, L. Vigevano, Roe linearization for the van der Waals gas, J. Comput. Phys. 175 (2002) 50–78.
[42] P. Boncinelli, F. Rubechini, A. Arnone, M. Cecconi, C. Cortese, Real gas effects in turbomachinery flows: A computational fluid dynamics model for fast computations, ASME J. Turbomachinery 126 (2004) 268–276.
[43] P. Cinnella, P.M. Congedo, Numerical solver for dense gas flows, AIAA J. 43 (11) (2005) 2457–2461.
[44] F. Rubechini, M. Marconcini, A. Arnone, M. Maritano, S. Cecchi, The impact of gas modeling in the numerical analysis of a multistage gas turbine, in: Proceedings of ASME Turbo Expo 2006, Barcelona, Spain, 8–11 May, ASME Paper GT2006-90129.
[45] R.C. Reid, J.M. Prausnitz, B.E. Poling, The Properties of Gases and Liquids, fourth ed., McGraw-Hill, London, UK, 1987.
[46] R. Stryjek, J.H. Vera, An improved Peng–Robinson equation of state for pure compounds and mixtures, Can. J. Chem. Engrg. 64 (1986) 323–333.
[47] P. Colonna, N.R. Nannan, A. Guardone, Multiparameter equations of state for siloxanes: [(CH$_3$)3-Si–O1/2]2-[O–Si–(CH3)2]i = 1,...,3, and [O–Si–(CH$_3$)$_2$]6, Fluid Phase Equilibr. 263 (2008) 115–130.
[48] P. Colonna, N.R. Nannan, A. Guardone, T.P. van der Stelt, On the computation of the fundamental derivative of gas dynamics using equations of state, Fluid Phase Equilibr. 286 (2009) 43–44.
[49] Z. Rusak, C.W. Wang, Low-drag airfoils for transonic flow of dense gases, Z. Angew. Math. Phys. (ZAMP) 51 (3) (2000) 467–480.
[50] R.G. Ghanem, S.D. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer-Verlag, 1991.
[51] D. Xiu, G.E. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, J. Sci. Comput. 26 (2002).
[52] R.H. Cameron, W.T. Martin, The orthogonal development of nonlinear functionals in series of Fourier–Hermite functionals, Ann. Math. 48 (1947) 385.
[53] Th. Crestaux, O. Le Maître, J.M. Martinez, Polynomial chaos expansion for sensitivity analysis, Reliab. Engrg. Syst. Safe. 94 (2009) 1161–1172.
[54] T. Gerstner, M. Griebel, Numerical Integration using sparse grids, Numer. Algorithms (18) (1998) 209–223.
[55] <http://opus-project.fr>.
[56] K. Petras, SmolPack – A software for Smolyaks quadrature with Clenshaw Curtis basic sequences, <http://www.public.tu-bs.de:8000/~petras/software.html>.
[57] A. Saltelli, S. Tarantola, F. Campolongo, Sensitivity Analysis in Practice, Wiley, 2004.
[58] I.M. Sobol, Sensitivity estimates for non linear mathematical models, Math. Model. Comput. Exp. 1 (4) (1993) 407–414.
[59] N. Srinivas, K. Deb, Multiobjective function optimization using nondominated sorting genetic algorithms, Evol. Comput. 2 (3) (1995) 221–248.

Pietro Marco Congedo is a Postdoctoral Research Fellow at LEGI Grenoble, France. He graduated with honours in Materials Engineering at University of Lecce (Italy). After his Master in Fluid Mechanics at Arts et Métiers (Paris, France), he received his Ph.D. in Energy Systems at University of Lecce in 2007. His research interests are in the numerical simulation and optimisation of real gas flows, thermodynamics of complex flows, multiphase flows, and uncertain optimisation.

Christophe Corre is a Professor at "Institut National Polytechnique" in Grenoble, France. He got a Master of Science in Aerospace Engineering from Michigan University in 1991. He received his Ph.D. (1995) in Fluid Mechanics at Arts et Métiers (Paris-France), where he worked as Assistant Professor until 2006. His research interests are in computational fluid dynamics, aerodynamics, implicit schemes, real gas dynamics and shape optimisation.

Jean-Marc Martinez received the Doctorat ès Sciences in physics from University Paris XI, France, in 1983. Currently, he is with the Systems and Structures Modelisation Department, Commissariat à l'Energie Atomique, Saclay, Gif-sur-Yvette, France. His activities are devoted to machine learning and uncertainties modelling in the field of numerical simulation to design and analyse computer experiments.