NON-INTRUSIVE STOCHASTIC STUDIES
FOR EXTERNAL AND INTERNAL FLOWS
USING THE elsA SOFTWARE

Marc Lazareff*, Jacques Peter†, Antoine Fourmaux‡

*†ONERA/DSNA
29, Avenue de la Division Leclerc F-92320 Châtillon
e-mail: *marc.lazareff@onera.fr †jacques.peter@onera.fr

‡ONERA/DAAP
8 rue des Vertugadins F-92190 Meudon
e-mail: antoine.fourmaux@onera.fr

Key words: Fluid Dynamics, elsA, stochastic, NIPCM, Python

Abstract. While the accuracy of deterministic CFD computations is increasing, physical
and geometrical inherent variations cannot be neglected. In the context of the NODESIM-
CFD European projecta, the influence of numerical options (e.g. numerical dissipation
coefficients, degree of convergence towards steady state) and modelisation (e.g. choice
of RANS turbulence model) on the outputs of CFD computations were first investigated.
Uncertainty quantification (UQ) using non-intrusive methods coupled with the elsA soft-
wareb were then carried out. Examples are presented for external and internal flows, using
Monte-Carlo (MC), complemented with a surrogate model, and non-intrusive polynomial
chaos method (NIPCM).

---

ahttp://www.nodesim.eu/
bhttp://elsa.onera.fr/
1 INTRODUCTION

The NODESIM-CFD European project has been an important incentive for ONERA/DSNA to develop NIPCM tools as a complement for its elsA CFD software, and also an efficient platform for comparison with other approaches. Here we are trying to build on this experience, trying to provide better confidence on standard results as well as an indication on perturbation propagation for industrial cases.

Stochastic studies involving polynomial chaos (an instance of homogenous chaos\(^1\)) in CFD may be broken down into two main approaches: the intrusive and non-intrusive approaches\(^{2,3,4}\). The former (intrusive) requires extensive modifications to be performed on the CFD software, so as to introduce the chaos polynomials inside the algebra, but once these costly modifications (in terms of software development) are completed it is computationally efficient\(^2\); as it depends on a polynomial representation of fluid phenomena, it is more suited to flow solutions not involving shocks or other quasi-discontinuities. The latter (non-intrusive) involves no modifications of the software (only additions to it), but is CPU-intensive; it provides the same capabilities as the original software regarding the flow-field regularity.

Here, being mainly interested in flows with shocks, we chose the Non-Intrusive Polynomial Chaos Method (NIPCM), so that stochastic processing is reduced to pre- and post-processing stages, with in-between deterministic CFD computations performed here by the (un-modified) elsA software\(^5\).

2 STOCHASTIC PROCESSING

The pre-processing stage involves the choice of the location of quadrature (Gauss-Jacobi, Gauss-Hermite . . . ) or collocation points, and the post-processing stage the computation of the elements of the stochastic output. In the case of Gauss quadrature, this requires the knowledge of the input stochastic distribution, which determines the associated orthogonal polynomials; the processed data is then possibly very sensitive to the exactness of the assumption on the input distribution type; the identification process may also be delicate, see 3.4.1.

Depending on the choice of observable quantities and on the regularity of the CFD process (linking the input and output distributions), it may be necessary to iterate on the global (pre- and post-processing) process, using several sizes of stochastic sets. The use of response surfaces, see 2.1, to interpolate between the CFD points allows to re-use all the existing sets.

All the processing has been performed using purposely developed Python modules, with use of mathematical libraries and of external software.

We have tried to evaluate both the adequacy of each chosen stochastic set for the representation of existing data and the individual quality of each CFD computation, see 3.4.2, so as to reduce the dependence of the conclusions on the CFD procedure.

A first impression on the set of computed results is generally obtained through the
M. Lazareff, J. Peter, A. Fourmaux

computation of a response surface (yielding smoothed results). As this technique is also used for generating interpolated sets in stochastic pre-processing, it will be introduced first.

2.1 TPS response surface technique

The Thin Plate Splines (TPS) technique\(^6\) has been selected for the definition of response surfaces, owing to its good mathematical properties and also to its lack of dependence on specific hypotheses\(^6\). We are using the regularised TPS implementation provided by the \texttt{fields} package\(^d\) of the GNU R software\(^e\). This package provides 2D response surfaces; for applications requiring 1D interpolation, we have used extrusion to simulate 2D data, so that a common interpolation technique is used throughout this work.

Thanks to its internal evaluation of the regularity of the input data (smoothing is determined using generalized cross-validation), the TPS technique may be applied without specifying any parameters. On the other hand, when respecting the data is more important than regularity, it is possible to obtain the internal parameters determined by the TPS algorithm and to re-apply the interpolation with different values. For example, the global scalar smoothing coefficient may be reduced by a factor of 10 from the computed value if precise representation is essential, while still yielding stable interpolation.

2.2 Stochastic pre-processing

We have used both Gauss and Monte-Carlo sets, the former being mainly useful for regular responses and the latter – especially when using over-sampling through TPS interpolation, see 2.4.2 – for the more complex behaviour newly observed on the RAE2822 profile, see 3.3.2.

It must be kept in mind in this respect that the discretisation in parameter space must be adequate not only for the representation of the input stochastic distribution, but also for the output distribution, when the response of the system includes steep regions, or plateau regions leading to peaks in the output pdf. Also, for distributions with non-limited support space like the normal distributions, it may occur that “events” on the low-probability parts of the distribution (ex. beyond 2\(\sigma\)) may have a noticeable contribution to the output pdf; thus, it may be important to check that the computed domain is sufficiently wide to include all the significant parts of the pdf.

2.3 Parametric variations with optimised chaining

On the software side, once the stochastic set is determined (as a list of parameter values), it is prepared for processing by the \texttt{elsA} software through the \texttt{variator} class of the Python-\texttt{elsA} interface. This class provides a generic parametric variations capability,

\(^a\)Such as for example the Gaussian distribution hypothesis for the Kriging technique.
\(^b\)By R. Furrer, D. Nychka and S. Sain.
\(^c\)\texttt{http://www.r-project.org/}
based on the definition of a list of perturbations to a base CFD case, itself defined for example by the Python-\texttt{elsA} script object corresponding to the “central” conditions. Each perturbation is defined (in the Python language) by a dictionary of (parameter, value) pairs; the \texttt{variate()} method of the \texttt{variator} class is then invoked to build “perturbed” copies of the base script object, which are all stored in a database keyed by the variation parameters. A single invocation of the \texttt{elsA} software with the \texttt{--span} command-line option then starts the actual CFD runs on the stochastic set; parallel computations sharing a common perturbations database are made possible, while avoiding duplicate runs, by the simple inclusion of a “run status” field in the database.

This object-oriented interface to parametric variations also provides improved operational efficiency through a programmable restart capability. A user-defined (or provided default) distance function defined in the parameter space may be used to automatically choose the next computation as the one involving the smallest “jump” in parameter space according to the specified distance. This allows to easily add collocation points to refine the DoE, while reducing the computational cost for added points thanks to improved initial conditions on restart. The mesh refinement level is treated the same as other parameters, which allows the “Full MultiGrid” (FMG) technique to be integrated into this optimised chaining scheme, providing it with a degree of tolerance to non-convergence because it allows neighbours on the same mesh level to be automatically used in lieu of the coarse-grid solution – for the same other parameters – if needed. This DoE strategy is also fully programmable, allowing for example to impose intermediate computations (linearisation) for long jumps involving float-type parameters such as the Mach number or angle-of-attack. Also, we may introduce metric skewing in parameter space to take account of already detected (or known) “steep” regions, see 3.4.2.

2.4 Stochastic post-processing

Stochastic post-processing has been performed using both direct summation (over Monte-Carlo sets) and Gauss quadrature (or Gauss collocation) over Gauss-Jacobi sets.

The direct summation technique does not need any explanation, as the expressions are quite straightforward. The quadrature/collocation methods are briefly described below.

2.4.1 Stochastic moments computation by Gauss quadrature/collocation

The Gauss quadrature method involves the representation of the output stochastic data on an orthogonal basis, selected to yield simplified integration expressions for the specific input distribution. It is thus a fast method, the more so if the polynomial coefficients and associated integration weights have been previously tabulated, but it may be strongly dependent on the exactness of the identification of this input distribution, whereas the cruder “direct summation” method is only independent of it to first order.

A Python module has been developed for general manipulations of stochastic data involving orthogonal polynomials, with recurrence relations providing theoretical arbitrary
order. Adequate precision for computations of integer-based rational fractions is preserved (when going through the recurrence iterations\textsuperscript{1} by using the \texttt{clnum} Python package, interface to the \texttt{CLN} ("Class Library for Numbers") software. This precision may be checked by comparison between the quadrature and collocation methods, the former being more dependent on the precision of the polynomial coefficients obtained through recurrence, while the latter could (at large orders) become dependent on the precision of the solution of the involved linear system.

This post-processing module allows a number of options, including: use of quadrature or collocation; use of maximum polynomial order (set count minus one), or lower; using sets created through \texttt{TPS} interpolation, from existing CFD data.

All the available data may be inspected using queries to a database-like structure (a Python dictionary); possible queries also include convenient comparison of results obtained using different options.

2.4.2 Set oversampling using a response surface

To provide additional precision in the evaluation of the propagated perturbations, which may involve steep variations of the CFD observables, we use the interpolation provided by a TPS response surface, see 2.1, to refine the discretisation of the parameter space. This interpolator is used as a surrogate function for the CFD solver, allowing to obtain large ($10^5$ to $10^6$) Monte-Carlo sets respecting the specified input stochastic distribution.

2.4.3 Computation of output probability density functions

The direct (pdf) and cumulated (cdf) probability density functions are computed here using set oversampling, see 2.4.2; these (large) interpolated sets provide enough samples to apply the “bins” method, which does not depend on any hypothesis on the output probability distribution. This output pdf, which is the composite of the input pdf and of the CFD response, may contain sharp peaks corresponding to extrema of the (real or approximated) response surface. These peaks, if any, prohibit (as they are significant and must be kept) the use of a smoothing function. When no peaks are found, a high-order polynomial smoothing is applied to aid in the comparison between variants of the computation.

3 EXTERNAL FLOWS: RAE2822 TRANSONIC PROFILE

A perturbation propagation analysis was performed on the \texttt{RAE2822} transonic profile, near Mach-induced stall conditions arising from shock-boundary layer interaction ($M = 0.734$, $\alpha = 2.79$ degree, $Re = 6.510^6$).

\textsuperscript{1}For example for the Jacobi polynomials

\textsuperscript{2}http://www.ginac.de/CLN/
3.1 Solution procedure for one point in parameter space

As the flow solution is noticeably dependent on turbulence modeling, we have selected the \( k - \omega \) model with the Menter SST correction, as previous studies have shown\(^7,8\) that without the SST correction the stall tendency is widely under-estimated. Computations are performed using only the finest \( (737 \times 177) \) available mesh of the FMG suite, as this correction has been found to require a fine mesh for efficiency; use of the restart technique provides reduced CPU costs for a neighborhood of points in the parameter space, so that the global cost of using this fine mesh is reduced.

Convergence is accelerated by the v-cycle variant of the multi-grid technique. Also, after some initial tests, a non-monotonic (simulated annealing) CFL law has been used here, which has removed some convergence difficulties previously encountered on the finest used mesh. Specifically, some “high \( \alpha \), medium \( M \)” combinations would lead to large-amplitude limit cycles, where the final value would depend essentially on the precise iteration count, which defines the phase in the cycle. These high-amplitude limit cycles, illustrated by above-surface points in Figure 3 (and also some hidden under-surface ones), have been eliminated here by this “CFL annealing” technique.

3.2 Specification of the stochastic perturbations

A normal distribution was chosen with the following parameters (taken from the NODESIM-CFD project specifications) for the mean value \( \mu \) and the standard deviation \( \sigma \) for \( \alpha \) and \( M \):

- \( \mu = 2.79 \) and \( \sigma = 0.1 \) (in degree) for the angle-of-attack \( \alpha \);
- \( \mu = 0.734 \) and \( \sigma = 0.005\mu \) for the Mach number \( M \).

The resultant parameter domain is defined by the \([0.704, 0.764] \times [2.09, 3.49] \) rectangle in \((M, \alpha)\) space.

3.3 Design of Experiment and results

The Design of Experiment (DoE) for the RAE2822 case has eventually been performed with user-driven increments, because of the steep variations which have found. Thus, the perturbation propagation analysis is here dependent on visual analysis of the CFD results.

3.3.1 1D analysis with stochastic moments computation

Two approaches have been combined to represent the normal distribution: a 9-point Gauss-Hermite set (reaching to \( \pm 2.65\sigma \)) and a specific Monte-Carlo set with 25 realisations (reaching to \([-2.1\sigma, +2.4\sigma]\)); the latter has been selected (by filtering out inadequate candidates) to nearly match the specified mean and standard deviation, and also the skewness (0.) and kurtosis (3.) of a normal distribution.
Results for this 1D study are presented below in 3.4.1, together with a tentative inspection of possible problems in identifying the input distribution type.

![Figure 1: rae2822 profile pdf and cdf responses: left: M perturbations, right: \( \alpha \) perturbations — for each sub-figure: left: pdf, right: cdf – top to bottom: input, inviscid lift, drag, moment outputs](image)

For the Mach number variations, these approaches have been for initial evaluation only, as the resultant discretisation of the parameter space has been found to be insufficient; thus, additional points were added “manually” to improve the set’s resolution, and the Gauss set was not used as such but only to complement the Monte-Carlo set, see below.

Stochastic moments were evaluated (both for the \( M \) and \( \alpha \) variations) by defining Monte-Carlo sets (using the original normal distributions) with respectively \( 10^5 \) and \( 10^6 \) points. For these high-density distributions, the CFD solution has been approximated by using TPS interpolation, see 2.4.2, on the existing CFD set (comprising both the Gauss and Monte-Carlo sets). This large number of points is not necessary for the computation of the first stochastic moments, but allows the output pdf (and cdf), see Figure 1, to be approximated (using “bins”) without requiring any additional hypothesis on the output probability distribution.

Please notice that the unusual aspect of the output pdf and cdf of the lift for Mach perturbations is due to the existence of a plateau for the high-\( C_L \) part of the \( C_L(M) \) response; this plateau leads both to the existence of a peak in the pdf and (through graphical scaling) of a reduced size of apparent size of the main part of the curve. The notable dissymmetry of this main part (peak excluded) is due to the near-stall conditions which gives smaller weights to the low-\( C_L \) segments where \( C_L \) is rapidly decreasing with increasing \( M \).
3.3.2 2D analysis with response surface computation

![Diagram of 2D analysis with response surface computation]

The above 1D explorations for $M = 0.734$ and $\alpha = 2.79$ degree have been extended with “off-axis” computations, to build a 2D response surface for combined $M$ and $\alpha$ variations.

For previous studies of the $C_L(M, \alpha)$ response surface on the rae2822, where the stall behaviour was clearly present but the surface appeared rather regular, we have used a simple 9-point stencil, or a more elaborate one including two levels of 9-point stencils complemented by boundary points, see the left part of Figure 2. As an unusually complex structure was detected in the response surface, we used a different strategy to account for this new behaviour, as explained below.

The incremental definition of the DoE starts with the 1D distribution at $\alpha = 2.79$, see Figure 2, including a 9-point Hermite-Gauss set and a 25-point Monte-Carlo set. As a steep feature was detected in the resultant $C_L(M)$ aerodynamic lift profile, the Gauss-Jacobi set for $M$ used at $\alpha = 2.79$ degree has been repeated for $\alpha = 2.49, 2.59, 2.69, 2.89, 2.99$ degree, yielding a total of 6 profiles. An unusual “canyon”-like feature then appeared diagonally, coherently across all 6 profiles, and additional CFD points were computed along the diagonal feature to characterise its structure. This rather large number of collocation points was needed to ascertain the existence of the “canyon”, which we had never observed before. The analysis is limited here to the appearance of the response surface. This aspect is nonetheless sufficiently complex, see Figure 3, to warrant a noticeable deviation from a Gaussian response in the output pdf for almost any direction of analysis in the $(M, \alpha)$ plane. Also, a detailed discretisation has been performed only along the $\alpha = 2.79$ degree profile; more structures could be revealed by a more detailed exploration.
3.4 Additional (non-standard) analysis of results

3.4.1 Influence of the input distribution type identification

In order to obtain some information on the influence of the specified type of input distribution on the output characteristics, we have re-processed the 1D data for the $C_L(M)$ profile – discussed in 3.3.1 – by assuming a symmetrical $\beta$-distribution instead of a normal one. The CFD computations used for this analysis have been obtained with the “classical” solution procedure, leading to a regular response surface (see Figure 3, left part).

On each sub-figure of Figures 5 and 4, we represent each computation by a point in the $(\mu, \sigma)$ plane, so that mean values are represented horizontally and standard deviations vertically. We first compare the clouds of points for 30 realisations of respectively 50-point (blue circles) and 10000-point (red circle) normal distribution sets, which clearly indicate that the 10000-point sets are much less scattered than the 50-point ones.

We then consider the points (brown diamonds) labeled “G-J 3” to “G-J 6”, corresponding to the (improper) post-processing of the data using a $\beta$-distribution (instead of a normal-distribution) hypothesis, with 3 to 6 (interpolated) Gauss-Jacobi points. It is observed on this case that, except for the moment coefficient (at bottom left on each figure), the “G-J 4” point is quite close to the 10000-point set “cluster” (red circles).

Thus, a representation of the data by third-order Jacobi polynomials would appear quite plausible as regards $\mu$ and $\sigma$, and might falsely lead to the identification of the $h$ moment coefficient combines lift and application point, which leads to a more complex profile, and may explain the different behaviour; also, convergence is not so good between “G-J 5” and “G-J 6” for this observable for Mach variations (Figure 5).

The moment coefficient combines lift and application point, which leads to a more complex profile, and may explain the different behaviour; also, convergence is not so good between “G-J 5” and “G-J 6” for this observable for Mach variations (Figure 5).
input distribution as a $\beta$-distribution and not a normal one. This conclusion must be tempered by the fact that this study was performed only on the first two moments (represented by the mean value $\mu$ and the standard deviation $\sigma$); using more moments would certainly provide greater discrimination between the two distribution types, and to different conclusions.

### 3.4.2 “Spiral” convergence study

A non-standard convergence analysis was performed by plotting the convergence history in the plane of the $(F_x, F_y)$ integrated convective fluxes, respectively equivalent (to a rotation and a normalisation) to the $(C_D, C_L)$ aerodynamic drag and lift pair. It was found that, for the tested cases of external 2D or 3D flows, using either the Euler, laminar or turbulent flow model, good convergence on the residuals was accompanied by a spiral aspect of the convergence figure in the $(F_x, F_y)$ plane. Other observed figures in this plane, which respectively seem to denote a lack of convergence and slow convergence, include the limit cycle (the curve representing the current flux values adheres to a closed curve,
as in an attractor) and the “pig-tail” (a helix-like figure with a decreasing radius).

The CFD computations used for this analysis have been obtained with the “CFL annealing” solution procedure, leading to an irregular response surface with a “canyon” (see 3.3.2).

Various quasi-spiral shapes are apparent in the left side of Figure 6, with the two cases of the limit cycle and of a “pig-tail” ending as a spiral on the right side. It may be that more convergence would have transformed the limit cycle into a converging spiral, so that further analysis is required (possibly using Poincaré sections to characterise limit cycles).

The link from the \((F_x, F_y)\) values, Figure 6 to the response surface, Figure 3, is not easy, but it has been noticed that the “limit cycle” case is prevalent in the “low \(M\), low \(\alpha\)” region of the current parameter space.

This geometrical study has been furthered in a tentative computation of a numerical value characterising the geometrical rate of shrinking of the spiral radius, Figure 7; this preliminary analysis shows a noticeable correlation between the “canyon” feature (left part of the Figure) and the values of this “spiral” criterion (right part), with the left half
of the canyon represented by negative values and the right half by positive values.

This is not yet fully investigated, but seems to show promise in this “spiral” criterion as in indication for DoE refinement.

4 INTERNAL FLOWS: NASA ROTOR 37 COMPRESSOR

4.1 Perturbations of the outlet static pressure

Computations have been performed using symmetric $\beta$-distributions around six specified values ($p = 110000, 114074, 119035, 121033, 123008, 124027 \text{ Pa}$) of the outlet pressure $p$ along the characteristic, using 6-point Gauss-Jacobi distributions in each case.

Results for the mass flow rate, stagnation pressure, stagnation temperature and isentropic efficiency have been computed by the \textit{elsA} software. Statistical moments to order 2 (mean, standard deviation) are computed here using the Gauss-Jacobi quadrature method.

There is a clear evolution in the ratio of the output perturbations, shifting from the mass flow at high pressure ratios to the pressure ratio at high mass flow, see Figure 4.1, left.

4.2 Perturbations of the tip clearance

Stochastic propagation computations have been performed for variations of the tip clearance at 4 points ($p = 105000, 110000, 114000, 119035 \text{ Pa}$) along the characteristic, using 5-point Gauss-Jacobi distributions in each case.

The same post-processing as for the outlet static pressure has been performed here (albeit on a different number of Gauss points).
It can be observed – Figure 4.1, right – that the output perturbations are here of almost constant amplitude on both the mass flow and the pressure ratio, contrary to the case of outlet static pressure perturbations.

It can be observed, Figure 4.2, that order-3 polynomials are more than adequate to represent the mean value and standard deviation for the chosen observables (mass flow, total pressure ratio, total temperature ratio and adiabatic efficiency), so that 4-point distributions would have been sufficient here.
5 CONCLUSION

A Python tool-chain including the ellsA CFD software has been developed for non-intrusive (Monte-Carlo and NIPC) stochastic studies. It has been applied to both external (rae2822 case) and internal (NASA Rotor 37 case) flows. A new “spiral” convergence criterion is being developed, based on the geometrical analysis of the convergence history in the plane of the $(F_x, F_y)$ integrated convective fluxes, which provides an alternate quality indicator for CFD results. A methodology for semi-automatic DoE refinement, based on this “spiral” convergence criterion, is being developed. It is hoped that, coupled with the optimised chaining technique of the variator class, it may become efficient even for non-regular response surfaces such as the “canyon” feature newly observed for the RAE2822 profile.

ACKNOWLEDGEMENT

This work was supported by the project NODESIM-CFD ”Non-Deterministic Simulation for CFD-based Design Methodologies” funded by the European Community represented by the CEC, Research Directorate-General, in the 6th Framework Programme, under Contract No. AST5-CT-2006-030959.

References


Contents

1 INTRODUCTION 2

2 STOCHASTIC PROCESSING 2
   2.1 TPS response surface technique 3
   2.2 Stochastic pre-processing 3
   2.3 Parametric variations with optimised chaining 3
   2.4 Stochastic post-processing 4
      2.4.1 Stochastic moments computation by Gauss quadrature/collocation 4
      2.4.2 Set oversampling using a response surface 5
      2.4.3 Computation of output probability density functions 5

3 EXTERNAL FLOWS : RAE2822 TRANSONIC PROFILE 5
   3.1 Solution procedure for one point in parameter space 6
   3.2 Specification of the stochastic perturbations 6
   3.3 Design of Experiment and results 6
      3.3.1 1D analysis with stochastic moments computation 6
      3.3.2 2D analysis with response surface computation 8
   3.4 Additional (non-standard) analysis of results 9
      3.4.1 Influence of the input distribution type identification 9
      3.4.2 “Spiral” convergence study 10

4 INTERNAL FLOWS : NASA ROTOR 37 COMPRESSOR 12
   4.1 Perturbations of the outlet static pressure 12
   4.2 Perturbations of the tip clearance 12

5 CONCLUSION 14