# Optimal Uncertainty Quantification of a risk measurement from a thermal-hydraulic code using Canonical Moments

Jerome Stenger, Fabrice Gamboa, Merlin Keller, Bertrand Iooss

# Optimal Uncertainty Quantification of a risk measurement from a thermal-hydraulic code using Canonical Moments

By Jérôme Stenger[1,2], Fabrice Gamboa[1], Merlin Keller[2], and Bertrand Iooss[1,2]

[1] Institut Mathématiques de Toulouse, 118 route de Narbonne.
fabrice.gamboa@math.univ-toulouse.fr
[2] EDF R&D, 6 quai Watier, Chatou.
jerome.stenger@edf.fr ; merlin.kerller@edf.fr
bertrand.iooss@edf.fr

**Abstract.** We study an industrial computer code related to nuclear safety. A major topic of interest is to assess the uncertainties tainting the results of a computer simulation. In this work we gain robustness on the quantification of a risk measurement by accounting for all sources of uncertainties tainting the inputs of a computer code. To that extent, we evaluate the maximum quantile over a class of distributions defined only by constraints on their moments. Two options are available when dealing with such complex optimization problems: one can either optimize under constraints; or preferably, one should reformulate the objective function. We identify a well suited parameterization to compute the optimal quantile based on the theory of canonical moments. It allows an effective, free of constraints, optimization.

## 1 Introduction

Computer codes are increasingly used to measure safety margins in nuclear accident management analysis instead of conservative procedures Pourgol-Mohamad et al. (2009). In this context, it is essential to evaluate the accuracy of the numerical model results, whose uncertainties come mainly from the lack of knowledge of the underlying physic and the model input parameters. The Best Estimate Plus Uncertainty (BEPU) methods Iooss and Marrel (2018) were developed in safety analyses, especially for the large break loss of coolant accident (see Proek and Mavko (2007), Sanchez-Saez et al. (2018)). Its principles rely mainly on a probabilistic modeling of the model input uncertainties, on Monte Carlo sampling for running the thermal-hydraulic computer code on sets of input, and on the application of statistical tools to infer high quantiles of the scalar output variables of interest Wallis (2004) .

This takes place in a more general setting, known as Uncertainty Quantification (UQ) methods de Rocquigny et al. (2008). Quantitative assessment of the uncertainties tainting the results of computer simulations is a major topic of interest in both industrial and scientific communities. In the context of nuclear safety, the computer models are expensive to run. Uncertainty propagation, risk measurement such as high quantile inference, or system robustness analysis become a difficult task using such models. In order to circumvent this problem, a widely accepted method consists in replacing the cpu time expensive numerical simulations by inexpensive mathematical functions called metamodels. This metamodel is build from a set of computer code simulations that must be as representative as possible of the code in the variation domain of its uncertain inputs. Generally, space-filling designs of experiments are created with a given budget of $n$ code evaluations, that provide a full coverage of the input space Fang et al. (2005). In the presence of a high number of input parameters, screening strategies are then performed in order to identify the Primary Influential Inputs (PII) on the model output variability and rank them by decreasing influence. From the learning sample, a metamodel is therefore built to fit the simulator output, considering only the PII as the explanatory inputs, while the remaining inputs remain fixed to a default value. Among all the metamodel-based solutions (polynomials, splines, neural networks, etc.), Gaussian process metamodeling, also known as Kriging Rasmussen and Williams (2005), has been very attractive. It makes the assumption that the response is a realization of a Gaussian process, conditioned on code observations. This approach provides the basis for statistical inference. In that, we dispose of simple analytical formulations of the predictor and the mean squared error of the predictions. The metamodel is then validated before being used. Several works have shown how this technique can help estimate quantile or probability of failure in thermal-hydraulic calculations (see Cannamela et al. (2008), Lorenzo et al. (2011)).

Our use-case consists in thermal-hydraulic computer experiments, typically used in support of regulatory work and nuclear power plant design and operation. The test case under study is a simplified one, as regards both physical phenomena and dimensions of the system, with respect to a realistic modeling of a reactor. the numerical model is based on code CATHARE 2 (V2.5_3mod3.1) which simulates the time evolution of physical quantities during a thermal hydraulic transient. The simulated accidental transient is an Intermediate Break Loss Of Coolant Accident (IBLOCA) with a break on the cold leg and no safety injection on the broken leg. In this use-case, $d = 27$ scalar inputs variables of CATHARE code are uncertain and defined by their probability density function. They correspond to various physical parameters, for instance: interfacial friction, critical flow rates, heat transfer coefficients, etc. They are all considered mutually independent. The output variable $Y$ is a single scalar which is the maximal peak cladding temperature (PCT) during the accident transient, see an example in Figure 1.

The number $n$ of simulations chosen for the design of experiments is a com-

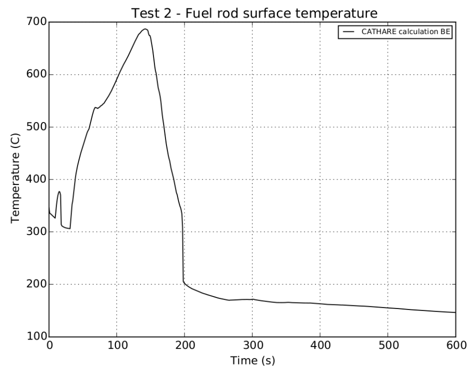


Figure 1: CATHARE temperature output for nominal parameters, the maximal temperature value is 687°C after 140 secondes.

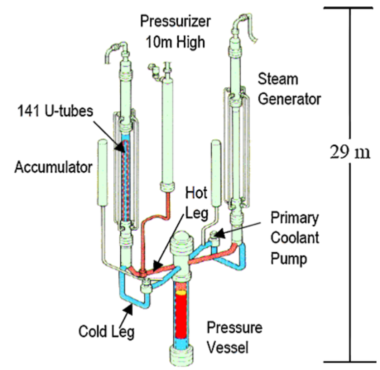


Figure 2: The replica of a water pressured reactor, with the hot and cold leg.

promise between the CPU time required for each simulation and the number of input parameters. For uncertainty propagation and metamodel building purpose, it is a common rule to chose $n$ at least 10times the dimension $d$ of the input vector. Here $n = 1000$ simulations were performed using a space filling Latin Hypercube Sample (LHS) in dimension 27, thus providing a nice coverage of the high-dimensional input space Fang et al. (2005).

A screening based on the Hilbert-Schmidt Independence Criterion (HSIC) dependence measure Da Veiga (2013), was performed on the $n = 1000$ learning simulations. The hypothesis "$\mathcal{H}_0^{(k)}$: the input $X_k$ and the output $Y$ are independent" was rejected for 11 inputs with a significance level $\alpha = 0.1$. In this work, we selected the 9 most important by ranking influence. Those inputs, designated as PII, are given by Table 1.

| Variable | Bounds | Initial distribution (truncated) | Mean | Second order moment |
|---|---|---|---|---|
| $n°10$ | $[0.1, 10]$ | $LogNormal(0, 0.76)$ | 1.33 | 3.02 |
| $n°22$ | $[0, 12.8]$ | $Normal(6.4, 4.27)$ | 6.4 | 45.39 |
| $n°25$ | $[11.1, 16.57]$ | $Normal(13.79$ | 13.83 | 192.22 |
| $n°2$ | $[-44.9, 63.5]$ | $Uniform(-44.9, 63.5)$ | 9.3 | 1065 |
| $n°12$ | $[0.1, 10]$ | $LogNormal(0, 0.76)$ | 1.33 | 3.02 |
| $n°9$ | $[0.1, 10]$ | $LogNormal(0, 0.76)$ | 1.33 | 3.02 |
| $n°14$ | $[0.235, 3.45]$ | $LogNormal(-0.1, 0.45)$ | 0.99 | 1.19 |
| $n°15$ | $[0.1, 3]$ | $LogNormal(-0.6, 0.57)$ | 0.64 | 0.55 |
| $n°13$ | $[0.1, 10]$ | $LogNormal(0, 0.76)$ | 1.33 | 3.02 |

Table 1: Corresponding moment constraints of the 9 most influential inputs of the CATHARE model.

The Gaussian process (Gp) is then build on the PII reduced space, conditioned from the available $n = 1000$ learning simulations. We usually consider in computer experiments an anisotropic (stationary) covariance, and the covariance kernel is here chosen as a Matrn 5/2. The metamodel accuracy is evaluated using the predictivity coefficient $Q^2$ Gratiet et al. (2017):

$$Q^2 = 1 - \frac{\sum_{i=1}^{n_{test}} (y^{(i)} - \widehat{y}^{(i)})^2}{\sum_{i=1}^{n_{test}} (y^{(i)} - \frac{1}{n_{test}} \sum_{i=1}^{n_{test}} y^{(i)})^2}$$

where $(x^{(i)})_{1 \leq i \leq n_{test}}$ is a test sample, $(y^{(i)})_{1 \leq i \leq n_{test}}$ are the corresponding observed outputs and $(\widehat{y}^{(i)})_{1 \leq i \leq n_{test}}$ are the metamodel predictions. We use a Leave one out strategy in order to perform the validation on the learning sample. The result of $Q^2 = 0.92$ confirms the efficiency of the screening strategy to increase the quality of the prediction.

As already discussed, once the predictive metamodel has been built, it can be used to perform uncertainty propagation and in particular, estimate probabilities or, as here, quantiles. This measure of risk will be designated from now on as our quantity of interest. The most trivial approach to estimate a quantile with a Gp metamodel, called *plug-in* approach, is to apply the quantile definition to the predictor of the metamodel. As the expectation of the Gp mean is a deterministic function of the input, this provides a deterministic expression of the quantile but no confidence intervals. Moreover for high quantiles, this methods tends to substantially underestimate the true quantile Cannamela et al. (2008). To assess this problem, Oakley (2004) has proposed to take into account the covariance structure of the Gp metamodel. The quantile definition is therefore applied to the global Gp metamodel and yields a random variable, whose expectation can be considered as the quantile estimator and its variance an indicator of the accuracy of its prediction. This *full-Gp* approach leads to confidence intervals. In practice, the estimation of a quantile with the full Gp approach is based on stochastic simulations (conditional on the learning sample) of the Gp metamodel.

This overall methodology yields the estimation of the 95%-quantile of the model output temperature. In nuclear safety, as in other engineering domains, methods of conservative computation of quantiles have been largely studied. Though the above construction work largely increases the robustness of the metamodel, the evaluation of the quantile remains tainted by the uncertainty of the input distributions. The inputs probability densities are usually chosen in parametric families (uniform, normal, log-normal, etc. See for instance Table 1), and their parameters are estimated using available data and/or an expert opinion. However, they may differ from reality. This uncertainty on the input probability densities is propagated to the quantile, hence, different choices of distributions will lead to different quantile values, thus different safety margins.

In this work, we propose to gain robustness on the quantification of this measure of risk. We aim to account for the uncertainty on the input distributions by evaluating the maximum quantile over a class of probability measures $\mathcal{A}$. In this optimization problem, the set $\mathcal{A}$ must be large enough to effectively represent our uncertainty on the inputs, but not too large in order to keep the estimation of the quantile representative of the physical phenomena. For example, the maximum quantile over the very large class $\mathcal{A} = \{\text{all distributions}\}$, proposed in Huber (1973), will certainly be too conservative to remain physically meaningful. Several articles which discuss possible choices of classes of distributions can be found in the literature of Bayesian robustness (see Ruggeri et al. (2005)). DeRoberts and Hartigan (1981), consider a class of measures specified by a type of upper and lower envelope on their density. Sivaganesan and Berger (1989) study the class of unimodal distributions. In more recent work, Owhadi et al. (2013) propose to optimize the measure of risk over a class of distributions specified by constraints on their *generalized* moments. They call their work Optimal Uncertainty Quantification (OUQ). However, in practical engineering cases, the available information on an input distribution is often reduced to the knowledge of its mean and/or variance. This is why in this paper, we are interested in a specific case of the framework introduced by Owhadi et al. (2013). We consider the class of measures known by some of their *classical* moments, chich we refer to as the moment class:

$$\begin{aligned} \mathcal{A} \quad = \quad & \Bigg\{ \mu = \otimes \mu_i \in \bigotimes_{i=1}^{d} \mathcal{M}_1([l_i, u_i]) \mid \mathbb{E}_{\mu_i}[x^j] = c_j^{(i)} \ , \\ & \qquad c_j^{(i)} \in \mathbb{R}, \ \text{for } 1 \leq j \leq N_i \text{ and } 1 \leq i \leq d \Bigg\} \ , \end{aligned} \tag{1}$$

where $\mathcal{M}_1([l_i, u_i])$ denotes the set of scalar probability measure on the interval $[l_i, u_i]$. The tensorial product of measure sets traduces the mutual independence of the $d$-components of the input vector $\mu$.

The solution to our optimization problem is numerically computed thanks to the OUQ reduction theorem (Owhadi et al. (2013), Winkler (1988)). This theorem states that the measure corresponding to the extremal CDF (can be extended to quantile through Proposition 2.1), is located on the extreme points of the distribution set. In the context of the moment class, the extreme distributions are located on the $d$-fold product of finite convex combinations of Dirac masses:

$$\mathcal{A}_\Delta \quad = \quad \left\{ \mu \in \mathcal{A} \mid \mu_i = \sum_{k=1}^{N_i+1} w_k^{(i)} \delta_{x_k^{(i)}} \quad \text{for} \ \ 1 \leq i \leq d \right\} \ , \tag{2}$$

To be more specific it holds that when $N$ pieces of information are available on the moments of a scalar measure $\mu$, it is enough to pretend that the measure is supported on at most $N + 1$ points. This powerful theorem gives the basis for practical optimization of our optimal quantity of interest. In this matter, Semi-Definite-Programming Henrion et al. (2009) has been already already explored by Betrò (2000) and Lasserre (2010), but the deterministic solver used rapidly reaches its limitation as the dimension of the problem increases. One can also find in the literature a Python toolbox developed by McKerns et al. (2012) called Mystic framework that fully integrates the OUQ framework. However, it was built as a generic tool for generalized moment problems and the enforcement of the moment constraints is not optimal. By restricting the work to our moment class, we propose an original and practical approach based on the theory of canonical moments Dette and Studden (1997). Canonical moments of a measure can be seen as the relative position of its moment sequence in the moment space. It is inherent to the measure and therefore presents many interesting properties. It allows to explore very efficiently the optimization space $\mathcal{A}_\Delta$, were the maximum quantile is to be found. Hence, we rewrite the optimization problem on the highly constrained set $\mathcal{A}_\Delta$ into a simplified and constraints free optimization problem.

The paper proceeds as follows. Section 2 describes the OUQ framework and the OUQ reduction theorem. In Section 3, we then describe step by step the algorithm calculating our quantity of interest with the canonical moments parameterization. We present in Section 4, an extended algorithm to deal with inequality constraints on the moments. Section 5 and 6 are dedicated to the application of our algorithm to a toy example, and to the peak cladding temperature for the IBLOCA application presented in the introduction. Section 7 gives some conclusions and perspectives.

# 2 OUQ principles

## 2.1 Duality transformation

In this work, we consider the quantile of the output of a computer code $G : \mathbb{R}^d \to \mathbb{R}$, seen as a black box function. As we said, in order to gain robustness on the risk measurement, our goal is to find the maximum quantile over the moment class $\mathcal{A}$ described in Equation (1). The objective value writes:

$$\begin{aligned} \overline{Q}_{\mathcal{A}}(p) &= \sup_{\mu\in\mathcal{A}} \left[ \inf \{ h \in \mathbb{R} \mid F_\mu(h) \geq p \} \right] , \\ &= \sup_{\mu\in\mathcal{A}} \left[ \inf \{ h \in \mathbb{R} \mid P_\mu(G(X) \leq h) \geq p \} \right] , \end{aligned} \tag{3}$$

The objective value written as a quantile (3) is not very convenient to work with. In order to applied the OUQ reduction Theorem 2.1 Owhadi et al. (2013), one must optimize an affine functional of the measure. In particular, it is necessary to optimize a probability instead of a quantile. The following result, illustrated in Figure 3, can be interpreted as a duality transformation of our optimization problem (3), into the optimization of a probability of failure (P.O.F). The proof is postponed to Appendix A.1.

**Proposition 2.1.** *The following duality result holds*

$$\overline{Q}_{\mathcal{A}}(p) = \inf \left\{ h \in \mathbb{R} \mid \inf_{\mu\in\mathcal{A}} F_\mu(h) \geq p \right\} .$$

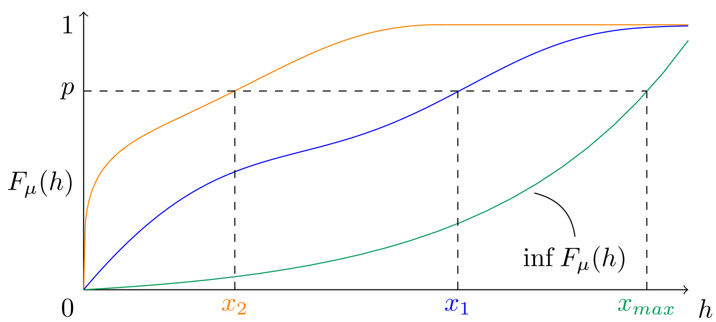


Figure 3: Illustration of the duality result 2.1. The green curve represents the CDF lower envelope; we can see that the maximum quantile $x_{max}$ is actually the quantile of the lowest CDF.

Proposition 2.1 reads like this: the optimal quantile over a class of distributions is equal to the quantile of the CDF lower envelope. Our problem is therefore to evaluate the lowest probability of failure $\inf_{\mu \in \mathcal{A}} P_\mu(G(X) \leq h)$ for a given threshold $h$.

## 2.2 Reduction Theorem

Under the form of Proposition 2.1, the OUQ reduction theorem applies (see Owhadi et al. (2013), Winkler (1988)). It states that the optimal solution of the P.O.F optimization is a product of discrete measures. A general form of the theorem reads as follows:

**Theorem 2.1** (OUQ reduction Owhadi et al. (2013, p.37))**.** *Suppose that $\mathcal{X} := \mathcal{X}_1 \times \cdots \times \mathcal{X}_d$ is a product of Radon spaces. Let*

$$\mathcal{A} := \left\{ (G, \mu) \;\middle|\; \begin{array}{l} G : \mathcal{X} \to \mathcal{Y}, \text{ is a real valued measurable function,} \\ \mu = \mu_1 \otimes \cdots \otimes \mu_p \in \bigotimes_{i=1}^{d} \mathcal{M}_1(\mathcal{X}_i) \ , \\ \text{for each } G, \text{ and for some measurable functions} \\ \varphi_l : \mathcal{X} \to \mathbb{R} \text{ and } \varphi_j^{(i)} : \mathcal{X}_i \to \mathbb{R} \ , \\ \qquad \bullet\ \mathbb{E}_\mu[\varphi_l] \leq 0 \text{ for } l = 1, \dots, N_0 \ , \\ \qquad \bullet\ \mathbb{E}_{\mu_i}[\varphi_j^{(i)}] \leq 0 \text{ for } j = 1, \dots, N_i \text{ and } i = 1, \dots, d \end{array} \right\}$$

*Let $\Delta_n(\mathcal{X})$ be the set of all discrete measure supported on at most $n+1$ points of $\mathcal{X}$, and*

$$\mathcal{A}_\Delta := \{(G, \mu) \in \mathcal{A} \mid \mu_i \in \Delta_{N_0+N_i}(\mathcal{X}_i)\} \ .$$

*Let $q$ be a measurable real function on $\mathcal{X} \times \mathcal{Y}$. Then*

$$\sup_{(G,\mu) \in \mathcal{A}} \mathbb{E}_\mu[q(X, G(X))] = \sup_{(G,\mu) \in \mathcal{A}_\Delta} \mathbb{E}_\mu[q(X, G(X))] \ .$$

This theorem derives from the work of Winkler (1988), who has shown that the extreme measures of a moment class $\{\mu \in \mathcal{M}_1(\mathcal{X}) \mid \mathbb{E}_\mu[\varphi_1] \leq 0, \ldots, \mathbb{E}_\mu[\varphi_n] \leq 0\}$ are the discrete measures that are supported on at most $n+1$ points. The strength of Theorem 2.1 is that it extends the result to a tensorial product of moment sets. The proof relies on a recursive argument using Winkler's classification on every set $\mathcal{X}_i$. A remarkable fact is that, as long as the quantity to be optimized is an affine function of the underlying measure $\mu$, this theorem remains true whatever the function $G$ and the quantity of interest $q$ are. Applying Theorem 2.1 to the optimization of the probability of failure, it is rewritten as:

$$\begin{aligned} \inf_{\mu \in \mathcal{A}} F_\mu(h) &= \inf_{\mu \in \mathcal{A}_\Delta} F_\mu(h) \ , \\ &= \inf_{\mu \in \mathcal{A}_\Delta} P_\mu(G(X) \leq h) \ , \\ &= \inf_{\mu \in \mathcal{A}_\Delta} \sum_{i_1=1}^{N_1+1} \cdots \sum_{i_d=1}^{N_d+1} \omega_{i_1}^{(1)} \ldots \omega_{i_d}^{(d)} \ \mathbb{1}_{\{G(x_{i_1}^{(1)}, \ldots, x_{i_d}^{(d)}) \leq h\}} \ , \end{aligned} \tag{4}$$

# 3 Parameterization using canonical moments

The optimization problem in Equation (4) shows that the weights and positions of the input distributions provide a natural parameterization for the computation of the P.O.F. However, in order to compute the lowest P.O.F, one must be able to explore the whole set of admissible measures $\mathcal{A}_\Delta$. Two ways to handle the problem appear. The first one consists in optimizing the objective value $F_\mu(h)$ under constraints, that is $\mu \in \mathcal{A}_\Delta$: this is the approach taken by McKerns et al. (2012) with the Mystic Framework. The second option, always favored when possible, consists in reformulating the objective function. This requires to identify a new parameterization adapted to the problem. Here, canonical moments Dette and Studden (1997) provide a surprisingly well tailored reparameterization.

The work on canonical moments was first introduced by Skibinsky (1967). His main contribution covered the original study of the geometric aspect of general moment space Skibinsky (1977), Skibinsky (1986). In a number of further papers, Skibinsky proves numerous other interesting properties of the canonical moments. Dette and Studden (1997) have shown the intrinsic relation between a measure $\mu$ and its canonical moments. They highlight the interest of canonical moments in many areas of statistics, probability and analysis such as problem of design of experiments, or the Hausdorff moment problem Hausdorff (1923). In the following, we describe step by step the algorithm used to transform the optimization problem of Equation 4 under the canonical moments parameterization.

## 3.1 Step 1. From classical moments to canonical moments

We enforce some moments on the input distributions of the code $G$. In this section, we present how to transform these *classical* moment constraints, into

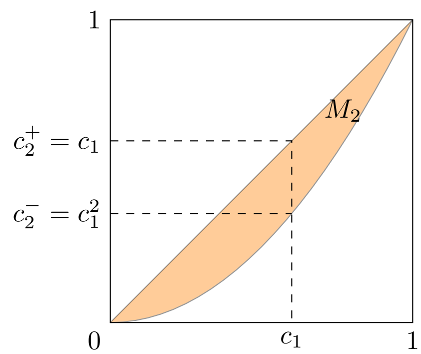


Figure 4: The moment set $M_2$ and definition of $c_2^+$ and $c_2^-$ for $(a,b) = (0,1)$.

canonical moments constraints.

We define the moment space $M := M(a,b) = \{\mathbf{c}(\mu) \mid \mu \in \mathcal{M}_1([a,b])\}$ where $\mathbf{c}(\mu)$ denote the sequence of all moments of some measure $\mu$. The $n$th moment space $M_n$ is defined by projecting $M$ onto its first $n$ coordinates, $M_n = \{\mathbf{c}_n(\mu) = (c_1, \ldots, c_n) \mid \mu \in \mathcal{M}_1([a,b])\}$. $M_2$ is depicted in Figure 4. We first define the extreme values,

$$\begin{aligned} c_{n+1}^+ &= \max\left\{c \in \mathbb{R} : (c_1, \ldots, c_n, c) \in M_{n+1}\right\} \ , \\ c_{n+1}^- &= \min\left\{c \in \mathbb{R} : (c_1, \ldots, c_n, c) \in M_{n+1}\right\} \ , \end{aligned}$$

which represent the maximum and minimum values of the $(n+1)$th moment that a measure can have, when its moments up to order $n$ equal to $\mathbf{c}_n$ . The $n$th canonical moment is then defined recursively as

$$p_n = p_n(\mathbf{c}) = \frac{c_n - c_n^-}{c_n^+ - c_n^-} \ . \tag{5}$$

Note that the canonical moments are defined up to the degree $N = N(\mathbf{c}) = \min\{\ n \in \mathbb{N} \mid \mathbf{c}_n \in \partial M_n\ \}$, and $p_N$ is either 0 or 1. Indeed, we know from (Dette and Studden, 1997, Theorem 1.2.5) that $\mathbf{c}_n \in \partial M_n$ implies that the underlying $\mu$ is uniquely determined, so that, $c_n^+ = c_n^-$. We also introduce the quantity $\zeta_n = (1 - p_{n-1})p_n$ that will be of some importance in the following. The very nice properties of canonical moments are that they belong to $[0,1]$ and are invariant by any affine transformation of the support of the underlying measures. Hence, we may restrict ourselves to the case $a = 0$, $b = 1$.

Therefore, for every $i = 1, \ldots, d$, the support of the measure $\mu_i$ is transformed into $[0,1]$ using the affine transformation $y = l_i + (u_i - l_i)x$. The sequences of moments of the corresponding measures are written $\mathbf{c}'_i = ({c'}_1^{(i)}, \ldots, {c'}_{N_i}^{(i)})$ where ${c'}_j^{(i)}$ reads

$${c'}_j^{(i)} = \frac{1}{(u_i - l_i)^j} \sum_{k=0}^{j} \binom{j}{k} (-l_i)^{j-k} c_k^{(i)} \ , \quad \text{for } j = 1, \ldots, N_i\,. \tag{6}$$

Given a sequence of moment constraints $(c_j^{(i)})_{1\leq j\leq N_i}$ enforced to the $i$th input, it is then possible to calculate the corresponding sequence of canonical moments $\mathbf{p}_i = (p_1^{(i)}, \dots, p_{N_i}^{(i)})$. Dette and Studden (1997, p. 29) propose a recursive algorithm named *Q-D algorithm* that allows this computation. It drastically fastens the computational time compared to the raw formula that consists of computing Hankel determinants (Dette and Studden, 1997, p. 32). In practical applications, we wish to enforce low order of moments, typically order 2 (see for instance Table 2). In this case we dispose of the simple analytical formulas

$$p_1 = c_1 \qquad , \qquad p_2 = \frac{c_2 - c_1^2}{c_1(1-c_1)} \ .$$

One can easily see that enforcing $N_i$ moments or $N_i$ canonical moments to the $i$th input is equivalent. Indeed, Equations (5) and (6) can be inverted.

### 3.2 Step 2. From canonical moments to support points

From a given sequence of canonical moments, one wishes to reconstruct the support of a discrete measure. We introduce the Stieltjes Transform, which connects canonical moments of a measure to its support. The Stieltjes transform Dette and Studden (1997) of $\mu$ is defined as

$$S(z) = S(z,\mu) = \int_a^b \frac{d\mu(x)}{z-x} \ , \quad (z \in \mathbb{C}\backslash\{\text{supp}(\mu)\}) \ .$$

The transform $S(z,\mu)$ is an analytic function of $z$ in $\mathbb{C}\backslash\text{supp}(\mu)$. If $\mu$ has a finite support then

$$S(z) = \int_a^b \frac{d\mu(x)}{z-x} = \sum_{i=1}^{n} \frac{\omega_i}{z-x_i} \ ,$$

where the support points of the measure $\mu$ are distinct and denoted by $x_1$, $\dots, x_n$, with corresponding weights $\omega_1, \dots, \omega_n$. Alternatively, the weights are given by $\omega_i = \lim_{z\to x_i}(z-x_i)S(z)$. We can rewrite the transform as a ratio of two polynomials with no common zeros. The zeros of the denominator being the support of $\mu$.

$$S(z) = \frac{Q_{n-1}(z)}{P_n^*(z)} \ , \tag{7}$$

where $P_n^*(z) = \prod_{i=1}^n (z-x_i)$ and

$$\omega_i = \frac{Q_{n-1}(x_i)}{\frac{d}{dx}P_n^*(x)|_{x=x_i}} \ .$$

The Stieljes transform can also be written as a continuous fraction, some basic definitions and properties of continuous fraction are postponed to Appendix A.2.

**Theorem 3.1** (Dette and Studden (1997, Theorem 3.3.1))**.** *Let $\mu$ be a probability measure on the interval $[a,b]$ and $z \in \mathbb{C}\backslash[a,b]$, then the Stieltjes transform of $\mu$ has the continued fraction expansion (see Appendix A.2 for notation):*

$$\begin{aligned} S(z) &= \frac{1 \rfloor}{\lceil z-a} - \frac{\zeta_1(b-a) \rfloor}{\lceil 1} - \frac{\zeta_2(b-a) \rfloor}{\lceil z-a} - \dots \\ &= \frac{1 \rfloor}{\lceil z-a-\zeta_1(b-a)} - \frac{\zeta_1\zeta_2(b-a)^2 \rfloor}{\lceil z-a-(\zeta_2+\zeta_3)(b-a)} \\ &\quad - \frac{\zeta_3\zeta_4(b-a)^2 \rfloor}{\lceil z-a-(\zeta_4+\zeta_5)(b-a)} - \dots \end{aligned}$$

*Where we recall that $\zeta_n := p_{n-1}(1-p_n)$.*

Theorem 3.1 states that the Stieltjes transform can be computed when one knows the canonicals moments. It immediatly follows from Equation (7), Theorem 3.1 and Lemma A.1 that we have the following recursive formula for $P_n^*$

$$P_{k+1}^*(x) = (x-a-(b-a)(\zeta_{2k}+\zeta_{2k+1}))P_k^*(x) - (b-a)^2\zeta_{2k-1}\zeta_{2k}P_{k-1}^*(x)\ , \quad (8)$$

where $P_{-1}^* = 0$, $P_0^* = 1$. The support of $\mu$ thus consists of the roots of $P_n^*$. This obviously leads to the following theorem.

**Theorem 3.2** (Dette and Studden (1997, Theorem 3.6.1))**.** *Let $\mu$ denote a measure on the interval $[a,b]$ supported on $n$ points with canonical moments $p_1, p_2, \dots$. Then, the support of $\mu$ is the set of $\{x : P_n^*(x) = 0\}$ defined by Equation (8).*

In the following we consider a fixed sequence of moments $\mathbf{c}_n = (c_1, \dots, c_n) \in M_n$, let $\mu$ be a measure supported on at most $n+1$ points, such that its moments up to order $n$ coincide to $\mathbf{p}_n = (p_1, \dots, p_n)$ the corresponding sequence of canonical moments related to $\mathbf{c}_n$, as described in Section 3.1. Corollary 3.3 is the moment version of Theorem 3.2. The only difficulty compared to Theorem 3.2 is that one tries to generate admissible measures supported on *at most* $n+1$ Dirac masses. Given a measure supported on strictly less than $n+1$ points, the question is therefore to know whether it makes sense to evaluate the $n+1$ roots of $P_{n+1}^*$. A limit argument is used for the proof.

**Corollary 3.3.** *Consider a sequence of moments $\mathbf{c}_n = (c_1, \dots, c_n) \in M_n$, and the set of measures*

$$\mathcal{A}_\Delta = \left\{ \mu = \sum_{i=1}^{n+1} \omega_i \delta_{x_i} \in \mathcal{M}_1([a,b]) \mid \mathbb{E}_\mu(x^j) = c_j,\ j = 1, \dots, n \right\}\ .$$

*We define*

$$\Gamma = \left\{ (p_{n+1}, \dots, p_{2n+1}) \in [0,1]^{n+1} \mid p_i \in \{0,1\} \Rightarrow p_k = 0,\ k > i \right\}\ .$$

*Then there exists a bijection between $\mathcal{A}_\Delta$ and $\Gamma$.*

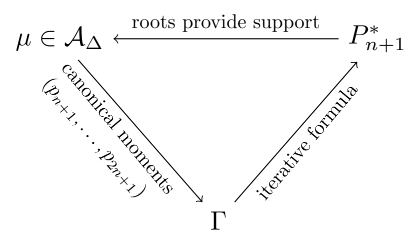


Figure 5: Relation between the set of admissible measures and the canonical moments.

*Proof.* Without loss of generality we can always assume $a = 0$ and $b = 1$ as the problem is invariant using affine transformation. We first consider the case where card(supp($\mu$)) is exactly $n+1$. From Theorem 3.2, the polynomial $P^*_{n+1}$ is well defined with $n+1$ distinct roots corresponding to the support of $\mu$. Notices that this implies that $(p_1, \ldots, p_{2n-1})$ belongs to $]0,1[^{2n-1}$ and that $p_{2n}, p_{2n+1}$ or $p_{2n+2}$ belong to $\{0,1\}$.

Now, the functions $g(x,z) = 1/(z-x)$ are equicontinuous for $z$ in any compact region which has a positive distance from $[0,1]$. The Stieljes transform is a finite sum of equicontinuous functions and therefore also equicontinuous. Thus if a measure $\mu$ converges weakly to $\mu^*$, the convergence must be uniform in any compact set with positive distance from $[0,1]$ (see Royden (1968)). It is then always possible to restrict ourselves to measures of cardinal $m < n+1$, by letting $p_k$ converge to 0 or 1 for $2m-2 \leq k \leq 2m$. Note that by doing so the polynomials $P^*_m$ and $P^*_{n+1}$ will have the same roots. But, $P^*_{n+1}$ and $Q_{n-1}$ will have some others roots of multiplicity strictly equal (see Equation (7) and (8)). The corresponding weights of these roots are vanishing, so that the measures extracted from $P^*_m$ and $P^*_{n+1}$ are the same. □

*Remark* 1. From a computational point of view, as the proof relies on a limit argument, we can always generate $p_k \in ]0,1[$ , for $n+1 \leq k \leq 2n+1$. This prevents the condition $p_k \in \{0,1\} \Rightarrow p_j = 0$ for $j > k$.

We use Section 3.1 to transform the $N_i$ constraints on the moments of the $i$th input into $N_i$ canonical moment constraints. However, the construction of the polynomial $P^{*(i)}_{N_i+1}$ requires the sequence $(p^{(i)}_1, \ldots, p^{(i)}_{2N_i+1})$. The $N_i$ first canonical moments of this sequence are known by the constraints, while the canonical moments $(p_k)_{N_i+1 \leq k \leq 2N_i+1} \in \Gamma$ constitute $N_i+1$ free parameters, in equal number to the cardinal of $\mu_i$. The computation of $\Gamma$ is very simple, it is basically done by random generation of $N_i+1$ numbers in $]0,1[$, yet it allows to generate the support of all the measures in $\mathcal{A}_\Delta$. This provides a very nice parameterization of the problem that takes naturally into account the constraints.

### 3.3 Step 3. From support points to weights

From the positions of a discrete measure $\mu$ in $\mathcal{A}_\Delta$ generated in Section 3.2. We easily recover the associated weights. Indeed, we enforce $N_i$ constraints on the moments of a scalar measure $\mu_i = \sum_{j=1}^{N_i+1} \omega_j^{(i)} \delta_{x_j^{(i)}}$, supported by at most $N_i+1$ points according to Theorem 2.1. A noticeable fact is that as soon as the $N_i+1$ support points of the distribution are set, the corresponding weights are uniquely determined. Indeed, the $N_i$ constraints lead to $N_i$ equations, and one last equation derives from the measure mass equals to 1. For each $1 \leq i \leq d$, the following $N_i+1$ linear equations holds

$$\begin{cases} \omega_1^{(i)} & + & \dots & + & \omega_{N_i+1}^{(i)} & = 1 \\ \omega_1^{(i)} x_1^{(i)} & + & \dots & + & \omega_{N_i+1}^{(i)} x_{N_i+1}^{(i)} & = c_1^{(i)} \\ \vdots & & & & \vdots & \vdots \\ \omega_1^{(i)} {x_1^{(i)}}^{N_i} & + & \dots & + & \omega_{N_i+1}^{(i)} {x_{N_i+1}^{(i)}}^{N_i} & = c_{N_i}^{(i)} \end{cases} \tag{9}$$

The determinant of the previous system is a Vandermonde matrix. Hence, the system is invertible as long as the $(x_j^{(i)})_j$ are distinct.

### 3.4 Step 4. Computation of the objective function

Thanks to Sections 3.1, 3.2, and 3.3, we can compute the positions $(x_j^{(i)})_{1 \leq j \leq N_i+1}$ and the weights $(\omega_j^{(i)})_{1 \leq j \leq N_i+1}$ of the $i$th input of some $\mu \in \mathcal{A}_\Delta$. We can therefore compute the following probability of failure (in Equation (4)):

$$P_\mu(G(X) \leq h) = \sum_{i_1=1}^{N_1+1} \cdots \sum_{i_d=1}^{N_d+1} \omega_{i_1}^{(1)} \dots \omega_{i_d}^{(d)} \, \mathbb{1}_{\{G(x_{i_1}^{(1)}, \dots, x_{i_d}^{(d)}) \leq h\}} \, ,$$

We recall that the positions and consequently the weights, were determined using Corrolary 3.3 from a sequence of canonical moments $(p_k)_{N_i+1 \leq k \leq 2N_i+1}$ $\in \Gamma$. So that, the exploration of $\mathcal{A}_\Delta$ is parameterized with canonical moments in $\Gamma$. No constraints need to be enforced, as a discrete measure generated from the canonical moments naturally satisfies the moment constraints. The P.O.F is then optimized globally using a differential evolution solver Price et al. (2005). Algorithm 1 summarizes step 1 to step 4 in order to compute the lowest probability of failure (4). The main cost of the algorithm arises from the high number of metamodel calls for $G$, evaluated on a $d$-dimensional grid of size $\prod_{i=1}^{d}(N_i+1)$.

**Algorithm 1** Calculation of the P.O.F

**Inputs:**

- lower bounds, $\mathbf{l} = (l_1, \dots, l_d)$
- upper bounds, $\mathbf{u} = (u_1, \dots, u_d)$
- constraints sequences of moments, $\mathbf{c}_i = (c_1^{(i)}, \dots, c_{N_i}^{(i)})$ and its corresponding sequences of canonical moments, $\mathbf{p}_i = (p_1^{(i)}, \dots, p_{N_i}^{(i)})$ for $1 \le i \le d$.

**Ensure:** $p_j^{(i)} \in ]0, 1[$ for $1 \le j \le N_i$ and $1 \le i \le d$

**function** P.O.F$(p_{N_1+1}^{(1)}, \dots, p_{2N_1+1}^{(1)}, \dots, p_{N_d+1}^{(d)}, \dots, p_{2N_d+1}^{(d)})$
  **for** $i = 1, \dots, d$ **do**
    **for** $k = 1, \dots N_i$ **do**
      $P_{k+1}^{*(i)} = (X - l_i - (u_i - l_i)(\zeta_{2k}^{(i)} + \zeta_{2k+1}^{(i)}))P_k^{*(i)} - (u_i - l_i)^2 \zeta_{2k-1}^{(i)} \zeta_{2k}^{(i)} P_{k-1}^{*(i)}$
    **end for**
    $x_1^{(i)}, \dots, x_{N_i+1}^{(i)} = \text{roots}(P_{N_i+1}^{*(i)})$
    $\omega_1^{(i)}, \dots, \omega_{N_i+1}^{(i)} = \text{weight}(x_1^{(i)}, \dots, x_{N_i+1}^{(i)}, \mathbf{c}_i)$
  **end for**
  **return** $\sum_{i_1=1}^{N_1+1} \cdots \sum_{i_d=1}^{N_d+1} \omega_{i_1}^{(1)} \dots \omega_{i_d}^{(d)} \, \mathbb{1}_{\{G(x_{i_1}^{(1)}, \dots, x_{i_d}^{(d)}) \le h\}}$
**end function**

# 4 Modified algorithm for inequality constraints

In the following, we consider inequality constraints for the moments. The optimization set reads

$$\mathcal{A} = \left\{ \mu = \otimes \mu_i \in \bigotimes_{i=1}^{d} \mathcal{M}_i([l_i, u_i]) \mid \alpha_j^{(i)} \le \mathbb{E}_{\mu_i}[x^j] \le \beta_j^{(i)} \ , \ 1 \le j \le N_i \right\} .$$

One can notice that $\alpha_j^{(i)} \le \mathbb{E}_{\mu_i}[x^j] \le \beta_j^{(i)}$ is equivalent to enforcing two constraints, thus drastically increasing the dimension of the problem. However, it is possible to restrict ourselves to one constraint. Considering the convex function $\varphi_j^{(i)} : x \mapsto (x^j - \alpha_j^{(i)})(x^j - \beta_j^{(i)})$, Jensen's inequality states that $\varphi_j^{(i)}(\mathbb{E}_{\mu_i}(x)) \le \mathbb{E}_{\mu_i}(\varphi_j^{(i)}(x))$. Therefore, the sole constraint $\mathbb{E}(\varphi_j^{(i)}(x)) \le 0$ ensures $\alpha_j^{(i)} \le \mathbb{E}_{\mu_i}[x^j] \le \beta_j^{(i)}$. Without loss of generality we still consider measures $\mu_i$ that are convex combinations of $N_i + 1$ Dirac masses, for $i = 1, \dots, d$.

We now propose a modified version of Algorithm 1 to solve the problem with inequality constraints. For $i = 1, \dots, d$, we denote the moments lower bounds $\boldsymbol{\alpha}_i = (\alpha_1^{(i)}, \dots, \alpha_{N_i}^{(i)})$ and the moments upper bounds $\boldsymbol{\beta}_i = (\beta_1^{(i)}, \dots, \beta_{N_i}^{(i)})$. We use Equation (6) to calculate the corresponding moment sequence $\boldsymbol{\alpha}_i'$ and $\boldsymbol{\beta}_i'$ after affine transformation to $[0, 1]$.

The P.O.F of algorithm 2 has $d + 2 \times \sum_{i=1}^{d} N_i$ arguments. The new parameters are actually the first $(N_i)_{|i=1,\dots,d}$th moments of the inputs that were previously fixed. A new step in the algorithm is needed to calculate the canonical moments

**Algorithm 2** Calculation of the P.O.F with inequality constraints

**Inputs:**

- lower bounds, $\mathbf{l} = (l_1, \ldots, l_d)$
- upper bounds, $\mathbf{u} = (u_1, \ldots, u_d)$
- moments lower bounds, $\boldsymbol{\alpha}'_i = ({\alpha'_1}^{(i)}, \ldots, {\alpha'_{N_i}}^{(i)})$ for $i = 1, \ldots, d$
- moments upper bounds, $\boldsymbol{\beta}'_i = ({\beta'_1}^{(i)}, \ldots, {\beta'_{N_i}}^{(i)})$ for $i = 1, \ldots, d$

**Ensure:**
$p_j^{(i)} \in [0,1]$ and ${c'_j}^{(i)} \in [{\alpha'_j}^{(i)}, {\beta'_j}^{(i)}]$ for $1 \leq j \leq N_i$ and $1 \leq i \leq d$.

**function** P.O.F$({c'_1}^{(1)}, \ldots, {c'_{N_i}}^{(1)}, p_{N_1+1}^{(1)}, \ldots, p_{2N_1+1}^{(1)}, \ldots, {c'_1}^{(d)}, \ldots, {c'_{N_d}}^{(d)}, p_{N_d+1}^{(d)}, \ldots, p_{2N_d+1}^{(d)})$
    **for** $i = 1, \ldots, d$ **do**
        **for** $k = 1, \ldots N_i$ **do**
            $p_k^{(i)} = f({c'_1}^{(i)}, \ldots {c'_k}^{(i)})$     *% f transform moments to canonical moments*
        **end for**
        **for** $k = 1, \ldots N_i$ **do**
            $P_{k+1}^{*(i)} = (X - l_i - (u_i - l_i)(\zeta_{2k}^{(i)} + \zeta_{2k+1}^{(i)}))P_k^{*(i)} - (u_i - l_i)^2 \zeta_{2k-1}^{(i)} \zeta_{2k}^{(i)} P_{k-1}^{*(i)}$
        **end for**
        $x_1^{(i)}, \ldots, x_{N_i+1}^{(i)} = \text{roots}(P_{N_i+1}^{*(i)})$
        $\omega_1^{(i)}, \ldots, \omega_{N_i+1}^{(i)} = \text{weight}(x_1^{(i)}, \ldots, x_{N_i+1}^{(i)}, \mathbf{c}'_i)$
    **end for**
    **return** $\sum_{i_1=1}^{N_1+1} \cdots \sum_{i_d=1}^{N_d+1} \omega_{i_1}^{(1)} \ldots \omega_{i_d}^{(d)} \ \mathbb{1}_{\{G(x_{i_1}^{(1)}, \ldots, x_{i_d}^{(d)}) \leq h\}}$
**end function**

up to degree $N_i$ for $i = 1, \ldots, d$. This ensures that the constraints are satisfied while the canonical moments from degree $N_i + 1$ up to degree $2N_i + 1$ can vary between $]0,1[$ in order to generate all possible measures. The increase of the dimension does not affect the computational times neither the complexity. Indeed, the main cost still arises from the large number of evaluation of the code $G$, that remains equal to $\prod_{i=1}^{d}(N_i + 1)$. Once again this new P.O.F function can be optimized using any global solver.

# 5 Numerical tests on a toy example

## 5.1 Presentation of the hydraulic model

In the following, we address a simplified hydraulic model Pasanisi et al. (2012). This code calculates the water height $H$ of a river subject to a flood event. It takes four inputs whose initial joint distribution is detailed in Table 2. It is always possible to calculate the *plug-in* quantiles for those particular distributions. The result is given in Figure 6, which present the initial CDF. However, as we desire to evaluate the robust quantile over a class of measures, we present in Table 3 the corresponding moment constraints that the variables must satisfy. The constraints are calculated based on the initial distributions, while the bounds are chosen in order to match the initial distributions most representative values.

| Variable | Distribution |
|---|---|
| $Q$: annual maximum flow rate | $Gumbel(mode = 1013, scale = 558)$ |
| $K_s$: Manning-Strickler coefficient | $\mathcal{N}(\overline{x} = 30, \sigma = 7.5)$ |
| $Z_v$: Depth measure of the river downstream | $\mathcal{U}(49, 51)$ |
| $Z_m$: Depth measure of the river upstream | $\mathcal{U}(54, 55)$ |

Table 2: Initial distribution of the 4 inputs of the hydraulic model.

| Variable | Bounds | Mean | Second order moment | Third order moment |
|---|---|---|---|---|
| $Q$ | $[160, 3580]$ | 1320.42 | $2.1632 \times 10^6$ | $4.18 \times 10^9$ |
| $K_s$ | $[12.55, 47.45]$ | 30 | 949 | 31422 |
| $Z_v$ | $[49, 51]$ | 50 | 2500 | 125050 |
| $Z_m$ | $[54, 55]$ | 54.5 | 2970 | 161892 |

Table 3: Corresponding moment constraints of the 4 inputs of the hydraulic model.

The height of the river $H$ is calculated through the analytical model

$$H = \left( \frac{Q}{300 K_s \sqrt{\frac{Z_m - Z_v}{5000}}} \right)^{3/5} . \tag{10}$$

We are interested in the flood probability $\sup_{\mu \in \mathcal{A}} P(H \geq h)$.

## 5.2 Maximum constraints order influence

We will compare the influence of the constraint order on the optimum. The initial distributions and the constraints enforced are available in Table 3. The value of the constraints correspond to the moments of the initial distributions. Figure 6 shows how the size of the optimization space $\mathcal{A}$ decreases by adding new constraints. A differential evolution solver was used to perform the optimization. The initial CDF was computed with a Monte Carlo algorithm. One can observe that enforcing only one constraint on the mean will give a robust quantile significantly larger than the one of the initial distribution. On the other hand, adding three constraints on every inputs reduces quite drastically the space so that the optimal quantile found are close to the one of the initial CDF.

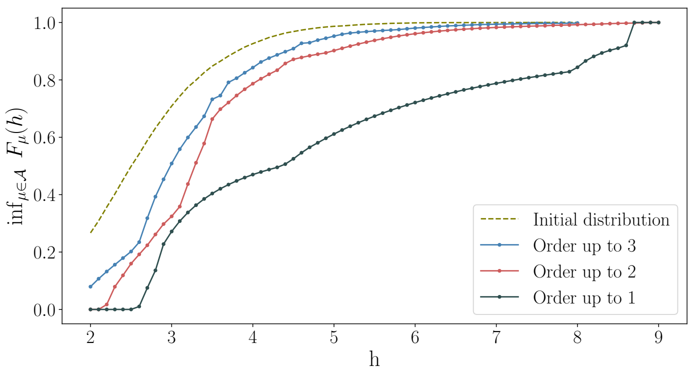


Figure 6: Influence of the number of constraints enforced on the minimal CDF.

## 5.3 Comparison with the Mystic framework

We highlight the interest of the canonical moments parameterization by comparing its performances with the Mystic framework McKerns et al. (2012). Mystic is a Python toolbox suitable for OUQ. In Figure 7 one can see the comparison beween Mystic and our algorithm. Both computations were realized with an identical solver, and computational times were similar (≈30 min). We enforced one constraint on the mean of each input (see Table 3). The performance of the Mystic framework is outperformed by our algorithm. Indeed, the generation of the weights and support points of the input distributions is not optimized in the Mystic framework. Hence, an intermediary transformation of the measure is needed in order to respect the constraints. During this transformation, the support points can be send out of bounds so that the measure is no more admissible. Many population vectors are rejected, which reduces the overall performance of the algorithm. Meanwhile, our algorithm warrants the exploration of the whole admissible set of measure without any vector rejection.

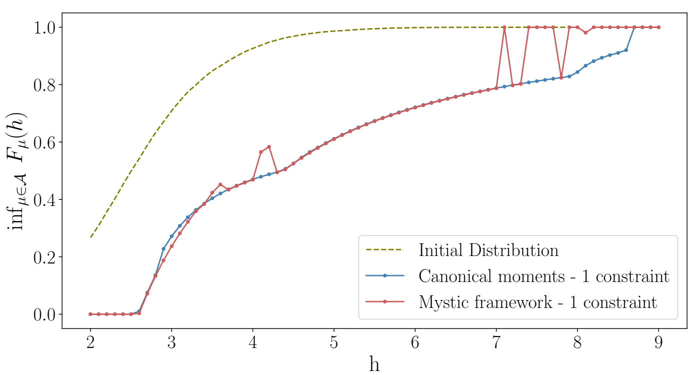


Figure 7: Comparison of the performance of the Mystic framework and our algorithm on the hydraulic code.

# 6 Application to the use-case

Two constraints were enforced on the first two moments of each inputs as displayed in Table 1. We successfully applied the methodology on the 9 dimensional restricted Gp metamodel of the CATHARE code. However, the computation was one day long for each threshold. We restricted the computation of the CDF to a small specific area of interest (high quantile 0.5-0.99) and we parallelized the task so that the computation did not exceed one week. One can compare, in Figure 8, the results of the computation realized with the Mystic framework and our algorithm. It confirms the difficulty for the Mystic framework to explore the whole space of admissible measures. On the other hand, this proves the efficiency of canonical moments to solve this optimization problem.

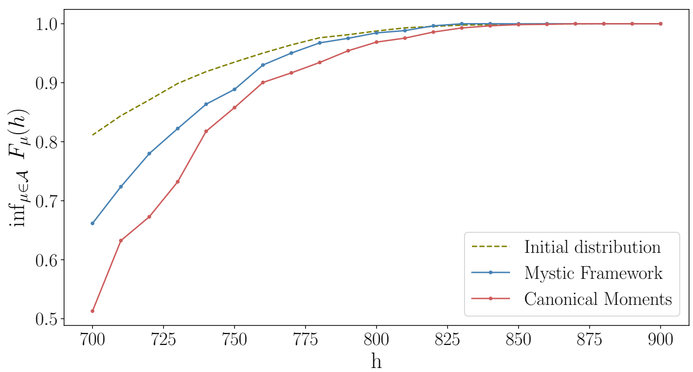


Figure 8: Comparison of the performance and our algorithm on the 9 dimensional restricted CATHARE Gp metamodel.

One can compare the estimation of the 95%-quantile of the peak cladding temperature for the IBLOCA application in Table 4. A 90%-confidence interval for the empirical quantile estimator was constructed with a bootstrap method. The *plug-in* and *full-Gp* approaches were defined in the introduction and correspond to a classical estimation of the quantile using respectively the predictor of the metamodel and the full Gaussian process Iooss and Marrel (2018). The robust method corresponds to the optimal quantile, when the input distributions are only defined by two of their moments (see Table 1). With this information, industrials are able to quantify the worst impact of the inputs uncertainty on the measure of risk, and adapt their choice of safety margins.

| | Empirical | Plug-in | Full-Gp | Robust |
|---|---|---|---|---|
| Mean | 746.80 | 735.83 | 741.46 | 788 |
| 90%-CI | $[736.7, 747.41]$ | | $[738.76, 744.17]$ | |

Table 4: Results for the 95%-quantile estimates.

# 7 Summary

Metamodels are widely used in industry to perform uncertainty propagation, in particular to evaluate measures of risk such as high quantiles. In this work, we successfully increased the robustness of the quantile evaluation by removing the main sources of uncertainties tainting the inputs of the computer code. We evaluated the maximum measure of risk over a class of distribution. We focus on set of measures only known by some of their moments, and adapted the theory of canonical moments into an improved methodology for solving OUQ problems. Our objective function has been parameterized with the canonical moments, which allows the natural integration of the constraints. The optimization can therefore be performed free of constraints, thus drastically increasing its efficiency. The restriction to moment constraints suits most of practical engineering cases. We also provide an algorithm to deal with inequality constraints on the moments, if an uncertainty lies in their values. Our algorithm shows very good performances and great adaptability to any constraints order. However, the optimization is subject to the curse of dimension and should be kept under 10 input parameters.

The joint distribution of the optimum is a discrete measure. One can criticize that it hardly corresponds to a physical, real world, interpretation. In order to address this issue, we will search for new optimization sets whose extreme points are not discrete measure. The unimodal class found in the literature of robust Bayesian analysis or the $\varepsilon$-contamination class, might be of some interest in this situation. New measures of risk will also be explored, for instance, superquantiles Rockafellar and Royset (2014), and Bayesian estimates associated to a given utility or loss function Berger (1985), which are of particular industrial interest.

# A Appendix

## A.1 Proof of duality proposition 2.1

*Proof.* we denote by $a = \sup_{\mu\in\mathcal{A}} \left[ \inf \{ h \in \mathbb{R} \; ; \; F_\mu(h) \geq p \} \right]$ and $b = \inf \left\{ h \in \mathbb{R} \mid \inf_{\mu\in\mathcal{A}} F_\mu(h) \geq p \right\}$. In order to prove $a = b$, we proceed in two step. First step, we have

$$
\begin{aligned}
& \text{for all } h \geq b \; ; \; \inf_{\mu\in\mathcal{A}} F_\mu(h) \geq p \; , \\
\Leftrightarrow \quad & \text{for all } h \geq b \; \text{ and for all } \mu \in \mathcal{A} \; ; \; F_\mu(h) \geq p \; , \\
\Leftrightarrow \quad & \text{for all } \mu \in \mathcal{A} \; \text{ and for all } h \geq b \; ; \; F_\mu(h) \geq p \; , \\
\Rightarrow \quad & \text{for all } \mu \in \mathcal{A} \; ; \; \inf\{ h \in \mathbb{R} \,|\, F_\mu(h) \geq p \} \leq b \; ,
\end{aligned}
$$

so that $b \geq a$. Second step, because $a$ is the sup of the quantiles,

$$
\begin{aligned}
& \text{for all } h \geq a \; ; \; \text{ for all } \mu \in \mathcal{A} \; ; \; F_\mu(h) \geq p \; , \\
\Rightarrow \quad & \text{for all } h \geq a \; ; \; \inf_{\mu\in\mathcal{A}} F_\mu(h) \geq p \; ,
\end{aligned}
$$

so that

$$\inf \left[ h \in \mathbb{R} \mid \inf_{\mu \in \mathcal{A}} F_\mu(h) \geq p \right] \leq a \ ,$$

and $b \leq a$. □

## A.2 Basic properties of continuous fraction

**Lemma A.1.** *A finite continued fraction is an expression of the form*

$$b_0 + \frac{a_1}{b_1 + \frac{a_2}{b_2 + \dots}} = b_0 + \frac{a_1 \rfloor}{\lceil b_1} + \frac{a_2 \rfloor}{\lceil b_2} + \dots + \frac{a_n \rfloor}{\lceil b_n} = \frac{A_n}{B_n} \ .$$

*The quantities $A_n$ and $B_n$ are called the $n$th partial numerator and denominator. There are basic recursive relations for the quantities $A_n$ and $B_n$ given by*

$$\begin{aligned} A_n &= b_n A_{n-1} + a_n A_{n-2} \ , \\ B_n &= b_n B_{n-1} + a_n B_{n-2} \ , \end{aligned}$$

*for $n \geq 1$ with initial conditions*

$$\begin{aligned} A_{-1} = 1 \quad &, \quad A_0 = b_0 \ , \\ B_{-1} = 0 \quad &, \quad B_0 = 1 \ . \end{aligned}$$

# References


Berger, J. O. 1985. *Statistical Decision Theory and Bayesian Analysis*, 2 ed. Springer Series in Statistics. Springer-Verlag, New York.

Betrò, B. 2000. *Robust Bayesian Analysis*. Lecture Notes in Statistics. Springer-Verlag, New York, Chapter 15, Methods for Global Prior Robustness under Generalized Moment Conditions.

Cannamela, C., Garnier, J., and Iooss, B. 2008. Controlled Stratification for Quantile Estimation. *The Annals of Applied Statistics 2,* 4, 1554–1580.

Da Veiga, S. 2013. Global Sensitivity Analysis with Dependence Measures. *Journal of Statistical Computation and Simulation 85*.

de Rocquigny, E., Devictor, N., Tarantola, S., Lefebvre, Y., Prot, N., Castaings, W., Mangeant, F., Schwob, C., Lavin, R., Masse, J.-R., Limbourg, P., Kanning, W., and Gelder, P. 2008. Uncertainty in industrial practice: A guide to quantitative uncertainty management.

DeRoberts, L. and Hartigan, J. A. 1981. Bayesian Inference Using Intervals of Measures. *The Annals of Statistics 9,* 2 (Mar.), 235–244.

Dette, H. and Studden, W. J. 1997. *The Theory of Canonical Moments with Applications in Statistics, Probability, and Analysis.* Wiley-Blackwell, New York.

Fang, K.-T., Li, R., and Sudjianto, A. 2005. *Design and Modeling for Computer Experiments (Computer Science & Data Analysis).* Chapman & Hall/CRC.

Gratiet, L. L., Marelli, S., and Sudret, B. 2017. Metamodel-Based Sensitivity Analysis: Polynomial Chaos Expansions and Gaussian Processes. *Handbook of Uncertainty Quantification*, 1289–1325.

Hausdorff, F. 1923. Momentprobleme für ein endliches Intervall. *Mathematische Zeitschrift 16*, 220–248.

Henrion, D., Lasserre, J.-B., and Lfberg, J. 2009. GloptiPoly 3: moments, optimization and semidefinite programming. *Optimization Methods and Software 24,* 4-5 (Oct.), 761–779.

Huber, P. J. 1973. The use of Choquet capacities in statistics. *Bulletin of the International Statistical Institute 45,* 4.

Iooss, B. and Marrel, A. 2018. Advanced methodology for uncertainty propagation in computer experiments with large number of inputs. *hal: 01907198*.

Lasserre, J.-B. 2010. *Moments, positive polynomials and their applications.* Number v. 1 in Imperial College Press optimization series. Imperial College Press ; Distributed by World Scientific Publishing Co, London : Signapore ; Hackensack, NJ. OCLC: ocn503631126.

Lorenzo, G., Zanocco, P., Gimnez, M., Marqus, M., Iooss, B., Bolado Lavn, R., Pierro, F., Galassi, G., D'Auria, F., and Burgazzi, L. 2011. Assessment of an Isolation Condenser of an Integral Reactor in View of Uncertainties in Engineering Parameters.

McKerns, M., Owhadi, H., Scovel, C., Sullivan, T. J., and Ortiz, M. 2012. The optimal uncertainty algorithm in the mystic framework. *CoRR abs/1202.1055*.

Oakley, J. 2004. Estimating Percentiles of Uncertain Computer Code Outputs. *Journal of the Royal Statistical Society. Series C (Applied Statistics) 53,* 1, 83–93.

Owhadi, H., Scovel, C., Sullivan, T. J., McKerns, M., and Ortiz, M. 2013. Optimal Uncertainty Quantification. *SIAM Review 55,* 2 (Jan.), 271–345. arXiv: 1009.0679.

Pasanisi, A., Keller, M., and Parent, E. 2012. Estimation of a quantity of interest in uncertainty analysis: Some help from Bayesian decision theory. *Reliability Engineering & System Safety 100*, 93–101.

Pourgol-Mohamad, M., Modarres, M., and Mosleh, A. 2009. Integrated Methodology for Thermal-Hydraulic Code Uncertainty Analysis with Application. *Nuclear Technology 165,* 3 (Mar.), 333–359.

Price, K., Storn, R. M., and Lampinen, J. A. 2005. *Differential Evolution: A Practical Approach to Global Optimization.* Natural Computing Series. Springer-Verlag, Berlin Heidelberg.

Proek, A. and Mavko, B. 2007. The State-of-the-Art Theory and Applications of Best-Estimate Plus Uncertainty Methods. *Nuclear Technology 158,* 1 (Apr.), 69–79.

Rasmussen, C. E. and Williams, C. K. I. 2005. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning).* The MIT

Press.

ROCKAFELLAR, R. T. AND ROYSET, J. O. 2014. Random Variables, Monotone Relations, and Convex Analysis. *Math. Program. 148,* 1-2 (Dec.), 297–331.

ROYDEN, H. 1968. *Real analysis.* Macmillan.

RUGGERI, F., RIOS INSUA, D., AND MARTIN, J. 2005. Robust Bayesian Analysis. In *Handbook of Statistics*, D. K. Dey and C. R. Rao, Eds. Bayesian Thinking, vol. 25. Elsevier, 623–667.

SANCHEZ-SAEZ, F., SNCHEZ, A. I., VILLANUEVA, J. F., CARLOS, S., AND MARTORELL, S. 2018. Uncertainty analysis of a large break loss of coolant accident in a pressurized water reactor using non-parametric methods. *Reliability Engineering and System Safety 174,* C, 19–28.

SIVAGANESAN, S. AND BERGER, J. O. 1989. Ranges of Posterior Measures for Priors with Unimodal Contaminations. *The Annals of Statistics 17,* 2 (June), 868–889.

SKIBINSKY, M. 1967. The range of the $(n + 1)$th moment for distributions on [0,1]. *Journal of Applied Probability 4,* 3 (Nov.), 543–552.

SKIBINSKY, M. 1977. The Maximum Probability on an Interval When the Mean and Variance Are Known. *Sankhy: The Indian Journal of Statistics, Series A (1961-2002) 39,* 2, 144–159.

SKIBINSKY, M. 1986. Principal representations and canonical moment sequences for distributions on an interval. *Journal of Mathematical Analysis and Applications 120,* 1 (Nov.), 95–118.

WALLIS, G. 2004. Uncertainties and probabilities in nuclear reactor regulation. *Nuclear Engineering and Design 237*, 1586–1592.

WINKLER, G. 1988. Extreme Points of Moment Sets. *Math. Oper. Res. 13,* 4 (Nov.), 581–587.