



HAL
open science

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

Jerome Stenger, Fabrice Gamboa, Merlin M. Keller, Bertrand Iooss

► **To cite this version:**

Jerome Stenger, Fabrice Gamboa, Merlin M. Keller, Bertrand Iooss. Optimal Uncertainty Quantification of a risk measurement from a thermal-hydraulic code using Canonical Moments. *International Journal of Uncertainty Quantification*, 2020, 10, pp.35-53. hal-01987449v2

HAL Id: hal-01987449

<https://hal.science/hal-01987449v2>

Submitted on 20 Aug 2019

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

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

OPTIMAL UNCERTAINTY QUANTIFICATION OF A RISK MEASUREMENT FROM A THERMAL-HYDRAULIC CODE USING CANONICAL MOMENTS

J. Stenger,^{1,2,*} F. Gamboa,² M. Keller,¹ & B. Iooss^{1,2}

¹EDF R&D, 6 quai Watier, 78401 Chatou, France

²Université Paul Sabatier, 118 route de Narbonne, 31400 Toulouse, France

*Address all correspondence to: J. Stenger, EDF R&D, 6 quai Watier, 78401 Chatou, France, E-mail: jerome.stenger@edf.fr

Original Manuscript Submitted: 12/04/2019; Final Draft Received: 12/04/2019

In uncertainty quantification studies, 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 bounded distributions satisfying 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 maximal quantile based on the theory of canonical moments. It allows an effective, free of constraints, optimization. This methodology is applied on an industrial computer code related to nuclear safety.

KEY WORDS: *uncertainty quantification, canonical moments, computational statistics, stochastic optimization*

1. INTRODUCTION

Computer codes are increasingly used to measure safety margins in nuclear accident management analysis instead of conservative procedures [1]. 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 [2] were developed in safety analyses, especially for the large break loss of coolant accident (see [3], [4]). 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 [5].

This takes place in a more general setting, known as Uncertainty Quantification (UQ) methods [6]. 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 [7]. 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 [8], 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.

Once the predictive metamodel has been built, it can be used to perform uncertainty propagation and in particular, estimate probabilities or, as here, quantiles (see [9], [10]). 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 [9]. To assess this problem, [11] 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 p -quantile of the model output. In nuclear safety, as in other engineering domains, methods of conservative computation of quantiles have been largely studied [12]. 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. For simplicity, the inputs probability densities are usually chosen in parametric families (uniform, normal, log-normal, etc. See for instance Table 3). The distribution's parameters are itself set with the available information coming from data and/or an expert opinion. This information is often reduced to an inputs mean value or a variance. Nevertheless, the distribution choice differs inevitably 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 [13], 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 [14]). [15], consider a class of measures specified by a type of upper and lower envelope on their density. [16] study the class of unimodal distributions. In more recent work, [17] propose to optimize the measure of risk over a class of distributions specified by constraints on their *generalized* moments. Their work (and ours) is called Optimal Uncertainty Quantification (OUQ), because given a set of assumptions and information, there exist optimal bounds on uncertainties. 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 [17]. We consider the class of measures known by some of their *classical* moments, which we refer to as the moment class:

$$\mathcal{A} = \left\{ \mu = \otimes_{i=1}^d \mu_i \in \bigotimes_{i=1}^d \mathcal{P}([l_i, u_i]) \mid \mathbb{E}_{\mu_i}[x^j] = c_i^{(j)}, \right. \\ \left. c_i^{(j)} \in \mathbb{R}, \text{ for } 1 \leq j \leq N_i \text{ and } 1 \leq i \leq d \right\}, \quad (1)$$

where $\mathcal{P}([l_i, u_i])$ denotes the set of scalar probability measure on the interval $[l_i, u_i]$. This set traduces simply that each random input is restricted to a given range and has some of its moments fixed. The tensorial product of measure sets

traduces the mutual independence of the d inputs. The choice of range enforcement is not a very strong hypothesis, as input variables represent physical parameters which are rarely unbounded.

The solution to our optimization problem is numerically computed thanks to the OUQ reduction theorem ([17], [18]). 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:

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

To be more specific it holds that when N pieces of information are available on the moments of a scalar measure μ , 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 [19] has been already explored by [20] and [21], 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 [22] 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 [23]. 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}_Δ , where the maximum quantile is to be found. Hence, we rewrite the optimization problem on the highly constrained set \mathcal{A}_Δ 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 $G : \mathbb{R}^d \rightarrow \mathbb{R}$, 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:

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

The objective value written as a quantile (3) is not very convenient to work with. In order to applied the OUQ reduction Theorem 1 [17], 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 1, 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.

Proposition 2.1.

The following duality result holds

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

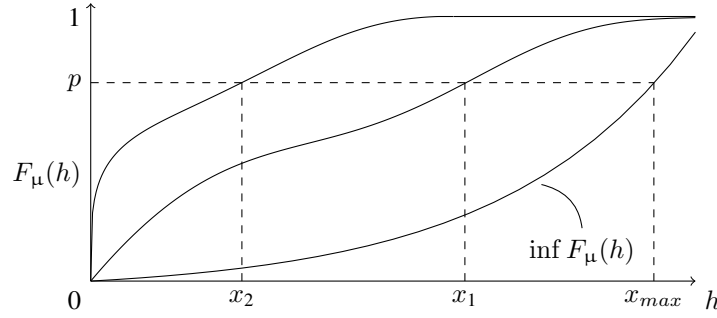


FIG. 1: Illustration of the duality result 2.1. The lower curve represents the CDF lower envelope; we can see that the maximum quantile x_{max} is actually the quantile of the lowest CDF.

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 $\inf_{\mu \in \mathcal{A}} \mathbb{P}_{\mu}(G(X) \leq h)$ for a given threshold h .

2.2 Reduction Theorem

Under the form of Proposition 2.1, the OUQ reduction theorem applies (see [17], [18]). 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 1 (OUQ reduction [17, p.37]). *Suppose that $\mathcal{X} := \mathcal{X}_1 \times \dots \times \mathcal{X}_d$ is a product of Radon spaces. Let*

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

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 [18], who has shown that the extreme measures of a moment class $\{\mu \in \mathcal{P}(\mathcal{X}) \mid \mathbb{E}_{\mu}[\varphi_1] \leq 0, \dots, \mathbb{E}_{\mu}[\varphi_n] \leq 0\}$ are the discrete measures that are supported on at most $n + 1$ points. The strength of Theorem 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 μ , this theorem remains true whatever the function G and the quantity of interest q are. Applying Theorem 1 to the optimization of the probability of failure, it is rewritten

as:

$$\begin{aligned}
\inf_{\mu \in \mathcal{A}} F_{\mu}(h) &= \inf_{\mu \in \mathcal{A}_{\Delta}} F_{\mu}(h), \\
&= \inf_{\mu \in \mathcal{A}_{\Delta}} \mathbb{P}_{\mu}(G(X) \leq h), \\
&= \inf_{\mu \in \mathcal{A}_{\Delta}} \sum_{i_1=1}^{N_1+1} \cdots \sum_{i_d=1}^{N_d+1} \omega_1^{(i_1)} \cdots \omega_d^{(i_d)} \mathbf{1}_{\{G(x_1^{(i_1)}, \dots, x_d^{(i_d)}) \leq h\}}, \tag{4}
\end{aligned}$$

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}_{Δ} . 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 [22] 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 [23] provide a surprisingly well tailored reparameterization.

The work on canonical moments was first introduced by [24]. His main contribution covered the original study of the geometric aspect of general moment space [25], [26]. In a number of further papers, Skibinsky proves numerous other interesting properties of the canonical moments. [23] have shown the intrinsic relation between a measure μ 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 [27]. 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 canonical moments constraints.

We define the moment space $M := M(a, b) = \{c(\mu) \mid \mu \in \mathcal{P}([a, b])\}$ where $c(\mu)$ denote the sequence of all moments of some measure μ . The n th moment space M_n is defined by projecting M onto its first n coordinates, $M_n = \{c_n(\mu) = (c_1, \dots, c_n) \mid \mu \in \mathcal{P}([a, b])\}$. M_2 is depicted in Figure 2. We first define the extreme values,

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

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 c_n . The n th canonical moment is then defined recursively as

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

Note that the canonical moments are defined up to the degree $N = N(c) = \min \{n \in \mathbb{N} \mid c_n \in \partial M_n\}$, and p_N is either 0 or 1. Indeed, we know from [23, Theorem 1.2.5] that $c_n \in \partial M_n$ implies that the underlying μ 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 $i = 1, \dots, d$, the support of the measure μ_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 $c'_i =$

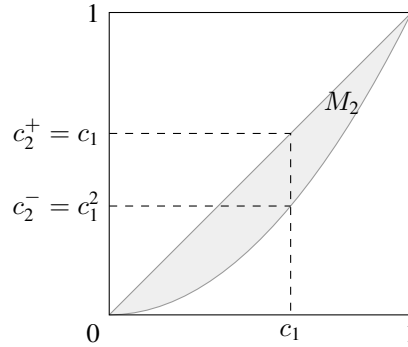


FIG. 2: The moment set M_2 and definition of c_2^+ and c_2^- for $(a, b) = (0, 1)$.

$(c_i^{(1)}, \dots, c_i^{(N_i)})$ where $c_i^{(j)}$ reads

$$c_i^{(j)} = \frac{1}{(u_i - l_i)^j} \sum_{k=0}^j \binom{j}{k} (-l_i)^{j-k} c_i^{(k)}, \quad \text{for } j = 1, \dots, N_i. \quad (6)$$

Given a sequence of moment constraints $(c_i^{(j)})_{1 \leq j \leq N_i}$ enforced to the i th input, it is then possible to calculate the corresponding sequence of canonical moments $\mathbf{p}_i = (p_i^{(1)}, \dots, p_i^{(N_i)})$. [23, 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 [23, p. 32]. In practical applications, we wish to enforce low order of moments, typically order 2 (see for instance Table 1). In this case we dispose of the simple analytical formulas

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

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 [23] of a scalar measure μ is defined as

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

The transform $S(z, \mu)$ is an analytic function of z in $\mathbb{C} \setminus \text{supp}(\mu)$. If μ 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 scalar measure μ are distinct and denoted by x_1, \dots, x_n , with corresponding weights $\omega_1, \dots, \omega_n$. Alternatively, the weights are given by $\omega_i = \lim_{z \rightarrow x_i} (z - x_i)S(z)$. We can rewrite the transform as a ratio of two polynomials with no common zeros. The zeros of the denominator being the support of μ .

$$S(z) = \frac{Q^{(n-1)}(z)}{P_*^{(n)}(z)}, \quad (7)$$

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 Stieltjes transform can also be written as a continuous fraction, some basic definitions and properties of continuous fraction are postponed to APPENDIX B.

Theorem 2 ([23, Theorem 3.3.1]). *Let μ be a probability measure on the interval $[a, b]$ and $z \in \mathbb{C} \setminus [a, b]$, then the Stieltjes transform of μ has the continued fraction expansion (see Appendix APPENDIX B for notation):*

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

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

Theorem 2 states that the Stieltjes transform can be computed when one knows the canonicals moments. It immediately follows from Equation (7), Theorem 2 and Lemma 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 μ thus consists of the roots of $P_*^{(n)}$. This obviously leads to the following theorem.

Theorem 3 ([23, Theorem 3.6.1]). *Let μ denote a measure on the interval $[a, b]$ supported on n points with canonical moments p_1, p_2, \dots . Then, the support of μ 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 μ 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 1 is the moment version of Theorem 3. The only difficulty compared to Theorem 3 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 1. *Consider a sequence of moments $\mathbf{c}_n = (c_1, \dots, c_n) \in M_n$, and the set of measures*

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

We define

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

Then there exists a bijection between \mathcal{A}_Δ and Γ .

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 $\text{card}(\text{supp}(\mu))$ is exactly $n + 1$. From Theorem 3, the polynomial $P_*^{(n+1)}$ is well defined with $n + 1$ distinct roots corresponding to the support of μ . Notices that this implies that (p_1, \dots, p_{2n-1}) belongs to $]0, 1[^{2n-1}$ and that p_{2n}, p_{2n+1} or p_{2n+2} belong to $\{0, 1\}$.

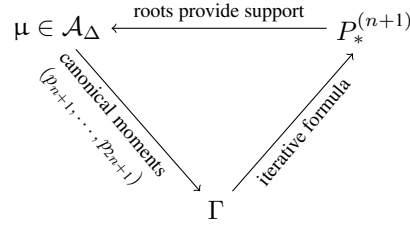


FIG. 3: Relation between the set of admissible measures and the canonical moments.

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 μ converges weakly to μ^* , the convergence must be uniform in any compact set with positive distance from $[0, 1]$ (see [28]). 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. \square

Remark 4. 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 $(p_1^{(i)}, \dots, p_{2N_i+1}^{(i)})$. 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 μ_i . The computation of Γ 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}_Δ . This provides a very nice parameterization of the problem that takes naturally into account the constraints.

We also provide a geometrical interpretation of this results. Consider the support of point (x_1, \dots, x_{n+1}) of a discrete measure in \mathcal{A}_Δ . The set of support point for all measure in \mathcal{A}_Δ is a manifold in \mathbb{R}^{n+1} . Corollary 1 described the structure of the manifold and states it is an algebraic variety. Indeed, it is the zero locus of the set of polynomial $P_*^{(n)}$ built from a sequence of canonical moments in $(p_1, \dots, p_n) + \Gamma$.

3.3 Step 3. From support points to weights

From the positions of a discrete measure μ in \mathcal{A}_Δ 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 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

$$\begin{cases} \omega_i^{(1)} & + \dots + \omega_i^{(N_i+1)} & = 1 \\ \omega_i^{(1)} x_i^{(1)} & + \dots + \omega_i^{(N_i+1)} x_i^{(N_i+1)} & = c_i^{(1)} \\ \vdots & & \vdots \\ \omega_i^{(1)} x_i^{(1)N_i} & + \dots + \omega_i^{(N_i+1)} x_i^{(N_i+1)N_i} & = c_i^{(N_i)} \end{cases} \quad (9)$$

The determinant of the previous system is a Vandermonde matrix. Hence, the system is invertible as long as the $(x_i^{(j)})_j$ 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 $(x_i^{(j)})_{1 \leq j \leq N_i+1}$ and the weights $(\omega_i^{(j)})_{1 \leq j \leq N_i+1}$ of the i th input of some $\mu \in \mathcal{A}_\Delta$. We can therefore compute the following probability of failure (in Equation (4)):

$$\mathbb{P}_\mu(G(X) \leq h) = \sum_{i_1=1}^{N_1+1} \cdots \sum_{i_d=1}^{N_d+1} \omega_1^{(i_1)} \cdots \omega_d^{(i_d)} \mathbf{1}_{\{G(x_1^{(i_1)}, \dots, x_d^{(i_d)}) \leq h\}},$$

We recall that the positions and consequently the weights, were determined using Corollary 1 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}_Δ is parameterized with canonical moments in Γ . 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 [29]. 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

Input: - lower bounds, $\mathbf{l} = (l_1, \dots, l_d)$
 - upper bounds, $\mathbf{u} = (u_1, \dots, u_d)$
 - constraints sequences of moments, $\mathbf{c}_i = (c_i^{(1)}, \dots, c_i^{(N_i)})$ and its corresponding sequences of canonical moments, $\mathbf{p}_i = (p_i^{(1)}, \dots, p_i^{(N_i)})$ for $1 \leq i \leq d$.

```

function P.O.F( $p_1^{(N_1+1)}, \dots, p_1^{(2N_1+1)}, \dots, p_d^{(N_d+1)}, \dots, p_d^{(2N_d+1)}$ )
  for  $i = 1, \dots, d$  do
    for  $k = 1, \dots, N_i$  do
       $P_{i*}^{(k+1)} = (X - l_i - (u_i - l_i)(\zeta_i^{2k} + \zeta_i^{(2k+1)}))P_{i*}^{(k)} - (u_i - l_i)^2 \zeta_i^{(2k-1)} \zeta_i^{(2k)} P_{i*}^{(k-1)}$ ;
    end
     $x_i^{(1)}, \dots, x_i^{(N_i+1)} = \text{roots}(P_i^{*(N_i+1)})$ ;
     $\omega_i^{(1)}, \dots, \omega_i^{(N_i+1)} = \text{weight}(x_i^{(1)}, \dots, x_i^{(N_i+1)}, \mathbf{c}_i)$ ;
  end
  return  $\sum_{i_1=1}^{N_1+1} \cdots \sum_{i_d=1}^{N_d+1} \omega_1^{(i_1)} \cdots \omega_d^{(i_d)} \mathbf{1}_{\{G(x_1^{(i_1)}, \dots, x_d^{(i_d)}) \leq h\}}$ ;
end

```

4. MODIFIED ALGORITHM FOR INEQUALITY CONSTRAINTS

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

$$\mathcal{A} = \left\{ \mu = \otimes \mu_i \in \bigotimes_{i=1}^d \mathcal{P}([l_i, u_i]) \mid \alpha_i^{(j)} \leq \mathbb{E}_{\mu_i}[x^j] \leq \beta_i^{(j)}, 1 \leq j \leq N_i \right\}.$$

One can notice that $\alpha_i^{(j)} \leq \mathbb{E}_{\mu_i}[x^j] \leq \beta_i^{(j)}$ 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 $\varphi_i^{(j)} : x \mapsto (x^j - \alpha_i^{(j)})(x^j - \beta_i^{(j)})$, Jensen's inequality states that $\varphi_i^{(j)}(\mathbb{E}_{\mu_i}(x)) \leq \mathbb{E}_{\mu_i}(\varphi_i^{(j)}(x))$. Therefore, the sole constraint $\mathbb{E}(\varphi_i^{(j)}(x)) \leq 0$ ensures $\alpha_i^{(j)} \leq \mathbb{E}_{\mu_i}[x^j] \leq \beta_i^{(j)}$. Without loss of generality we still consider measures μ_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 $\alpha_i = (\alpha_i^{(1)}, \dots, \alpha_i^{(N_i)})$ and the moments upper bounds $\beta_i = (\beta_i^{(1)}, \dots, \beta_i^{(N_i)})$. We use Equation (6) to calculate the corresponding moment sequence α'_i and β'_i after affine transformation to $[0, 1]$.

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

Input: - lower bounds, $\mathbf{l} = (l_1, \dots, l_d)$
 - upper bounds, $\mathbf{u} = (u_1, \dots, u_d)$
 - moments lower bounds, $\alpha'_i = (\alpha_i'^{(1)}, \dots, \alpha_i'^{(N_i)})$ for $i = 1, \dots, d$
 - moments upper bounds, $\beta'_i = (\beta_i'^{(1)}, \dots, \beta_i'^{(N_i)})$ for $i = 1, \dots, d$

```

function P.O.F ( $c_1^{(1)}, \dots, c_1^{(N_1)}, p_1^{(N_1+1)}, \dots, p_1^{(2N_1+1)}, \dots, c_d^{(1)}, \dots, c_d^{(N_d)}, p_d^{(N_d+1)}, \dots, p_d^{(2N_d+1)}$ )
  for  $i = 1, \dots, d$  do
    for  $k = 1, \dots, N_i$  do
      |  $p_i^{(k)} = f(c_i^{(1)}, \dots, c_i^{(k)})$ ; /* f transform moments to canonical moments */
    end
    for  $k = 1, \dots, N_i$  do
      |  $P_{i*}^{(k+1)} = (X - l_i - (u_i - l_i)(\zeta_i^{(2k)} + \zeta_i^{(2k+1)}))P_{i*}^{(k)} - (u_i - l_i)^2 \zeta_i^{(2k-1)} \zeta_i^{(2k)} P_{i*}^{(k-1)}$ ;
    end
     $x_i^{(1)}, \dots, x_i^{(N_i+1)} = \text{roots}(P_{i*}^{(N_i+1)})$ ;
     $\omega_i^{(1)}, \dots, \omega_i^{(N_i+1)} = \text{weight}(x_i^{(1)}, \dots, x_i^{(N_i+1)}, \mathbf{c}_i)$ ;
  end
  return  $\sum_{i_1=1}^{N_1+1} \dots \sum_{i_d=1}^{N_d+1} \omega_1^{(i_1)} \dots \omega_d^{(i_d)} \mathbf{1}_{\{G(x_1^{(i_1)}, \dots, x_d^{(i_d)}) \leq h\}}$ ;
end

```

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 up to degree N_i for $i = 1, \dots, d$. 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 [30]. 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 1. It is always possible to calculate the *plug-in* quantiles for those particular distributions. The result is given in Figure 4, which present the initial CDF. However, as we desire to evaluate the robust quantile over a class of measures, we present in Table 2 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.

TABLE 1: Initial distribution of the 4 inputs of the hydraulic model.

Variable	Description	Distribution
Q	annual maximum flow rate	$Gumbel(mode = 1013, scale = 558)$
K_s	Manning-Strickler coefficient	$\mathcal{N}(\bar{x} = 30, \sigma = 7.5)$
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: Corresponding moment constraints of the 4 inputs of the hydraulic model.

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

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

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

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

5.2 Maximum constraints order influence

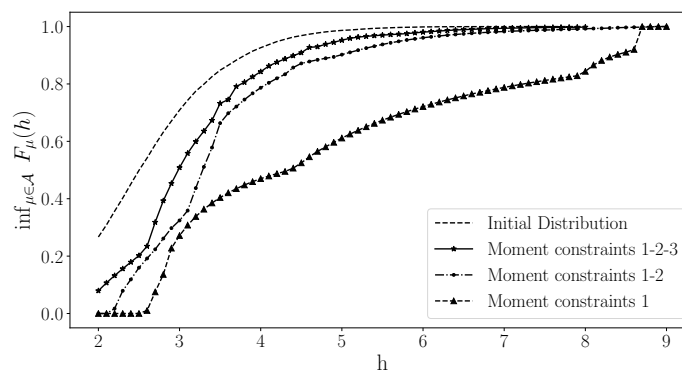


FIG. 4: h is the height of the river. We observe the influence of the number of moment constraints enforced on the CDF lower envelope.

We compare the influence of the constraint order on the optimum. The initial distributions and the constraints enforced are available in Table 2. The value of the constraints correspond to the moments of the initial distributions. Figure 4 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. A good compromise is to enforce two constraints on the first two moments. This is equivalent to enforcing the mean and the variance on the input distributions, which is a natural way of proceeding. Using this rule, the moment space is not too large so that the worst case quantile remains physically viable.

5.3 Comparison with the Mystic framework

We highlight the interest of the canonical moments parameterization by comparing its performances with the Mystic framework [22]. Mystic is a Python toolbox suitable for OUQ. In Figure 5 one can see the comparison between 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 2). 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 sent 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.

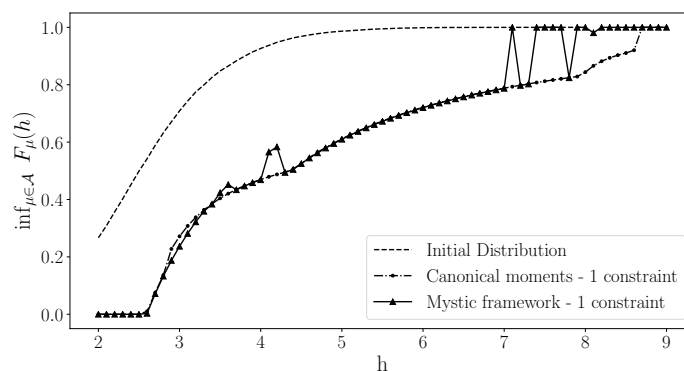


FIG. 5: Comparison of the performance of the Mystic framework and our algorithm on the hydraulic code.

6. APPLICATION TO THE USE-CASE

6.1 Presentation of the use-case

Our use-case consists in thermal-hydraulic computer experiments, typically used in support of regulatory work and nuclear power plant design and operation. 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. One run of the computer code takes around 20 minutes on an ordinary computer. Nevertheless, in regards of both physical phenomena and dimensions of the system, we could consider our case under study as a simplified model with respect to a realistic modeling of a reactor. 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 6.

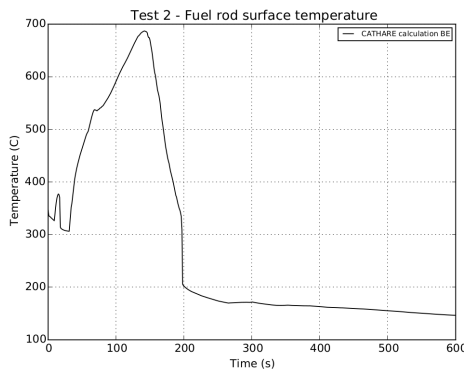


FIG. 6: CATHARE temperature output for nominal parameters, the maximal temperature value is 687°C after 140 secondes.

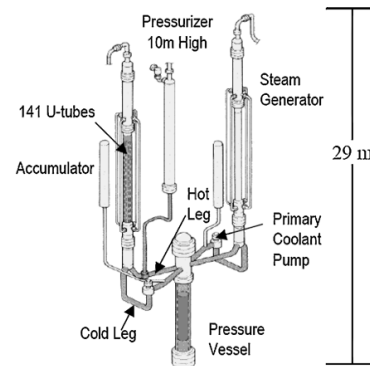


FIG. 7: The replica of a water pressured reactor, with the hot and cold leg.

The number n of simulations chosen for the design of experiments is a compromise 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 10 times the dimension d of the input vector [31]. 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 [7].

A screening based on the Hilbert-Schmidt Independence Criterion (HSIC) dependence measure [32], 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 9 inputs with a significance level $\alpha = 0.1$. Those inputs, designated as PII, are given by Table 3. The screening based on the HSIC takes into account the whole output variability. However, in some pathological case a low influential input for global variability may have an important impact on a p -quantile. A perspective to solve this issue, is to consider the conditional and target HSIC [33, p.23], this will be the aim of a future work.

TABLE 3: Corresponding moment constraints of the 9 most influential inputs of the CATHARE model [2].

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

The input distributions described in Table 3 derive from the CIRCE method [34,35], where in a Bayesian setting the posterior distributions are computed with respect to an experimental database. The moment presented in Table 3 correspond to the moments of the the calibrated distributions. Those are the moments enforced to the measures in the moment class. This way, we can compare the quantiles obtained through the *plug-in* and *full-Gp* approaches with the maximal quantile of the OUQ framework. Indeed, with these constraints the initial distribution is also an element of the moment class.

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 Matérn 5/2. The metamodel accuracy is evaluated using the predictivity coefficient Q^2 [36]:

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

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 $(\hat{y}_i)_{1 \leq i \leq n_{test}}$ are the metamodel predictions. We use a Leave one out strategy in order to perform the validation on the learning sample, we obtain $Q^2 = 0.92$. As already discussed in the introduction, once the predictive metamodel has been built, it can be used to perform uncertainty propagation and in particular, estimate quantiles.

6.2 Results on the use-case

Two constraints were enforced on the first two moments of each inputs as depicted in Table 3. Indeed, informations on the input distributions are often reduced to the knowledge of the means and variances. Moreover, in regards to the toy example, only one constraints on the mean defines a moment class too large, which leads to a worst case scenario not physically viable. At last, more constraints increases the number of support points of each discrete measure on the extreme points, thus increases the computational cost.

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. Once again, the computations were performed with identical solvers. It confirms the difficulty for the Mystic framework to explore the whole space of admissible measures. Indeed, the maximal quantile obtained with the Mystic framework is lower than the result of our optimization. This proves the efficiency of the canonical moments parameterization to solve this optimization problem.

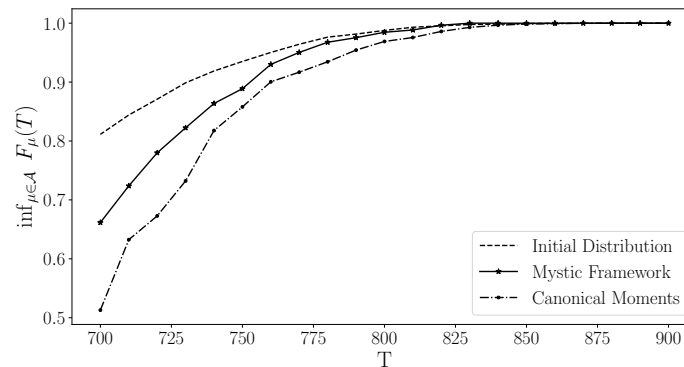


FIG. 8: Comparison of the performance and our algorithm on the 9 dimensional restricted CATHARE Gp metamodel. h denotes the maximal peak cladding temperature during the accident transient (see Figure 6).

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 [2]. The OUQ method corresponds to the maximal quantile, when the input distributions are only defined by two of their moments (see Table 3). It is optimal in a sense that it minimize the uncertainty on the input distribution considering the available information. 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. The overall optimization can also be realized via Algorithm

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

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

2, by relaxing the equality constraints into inequality constraints. This way, it is possible to quantify the sensitivity of the maximal quantile to the moment values.

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 ε -contamination class, might be of some interest in this situation. New measures of risk will also be explored, for instance, superquantiles [37], and Bayesian estimates associated to a given utility or loss function [38], which are of particular industrial interest.

APPENDIX A. 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

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

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

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

so that

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

and $b \leq a$. □

APPENDIX B. BASIC PROPERTIES OF CONTINUOUS FRACTION

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

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

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

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

REFERENCES

1. Pourgol-Mohamad, M., Modarres, M., and Mosleh, A., Integrated Methodology for Thermal-Hydraulic Code Uncertainty Analysis with Application, *Nuclear Technology*, 165(3):333–359, March 2009.
2. Iooss, B. and Marrel, A., Advanced methodology for uncertainty propagation in computer experiments with large number of inputs, *Nuclear Technology*, pp. 1–19, 2019.
3. Proek, A. and Mavko, B., The State-of-the-Art Theory and Applications of Best-Estimate Plus Uncertainty Methods, *Nuclear Technology*, 158(1):69–79, April 2007.
4. Sanchez-Saez, F., Snchez, A.I., Villanueva, J.F., Carlos, S., and Martorell, S., 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, 2018.
5. Wallis, G., Uncertainties and probabilities in nuclear reactor regulation, *Nuclear Engineering and Design*, 237:1586–1592, 2004.
6. 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., Uncertainty in industrial practice: A guide to quantitative uncertainty management, January 2008.
7. Fang, K.T., Li, R., and Sudjianto, A., *Design and Modeling for Computer Experiments (Computer Science & Data Analysis)*, Chapman & Hall/CRC, 2005.
8. Rasmussen, C.E. and Williams, C.K.I., *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*, The MIT Press, 2005.
9. Cannamela, C., Garnier, J., and Iooss, B., Controlled Stratification for Quantile Estimation, *The Annals of Applied Statistics*, 2(4):1554–1580, 2008.
10. Lorenzo, G., Zanooco, P., Gimnez, M., Marqus, M., Iooss, B., Bolado Lavn, R., Pierro, F., Galassi, G., D’Auria, F., and Burgazzi, L., Assessment of an Isolation Condenser of an Integral Reactor in View of Uncertainties in Engineering Parameters, *Science and Technology of Nuclear Installations*, 2011.
11. Oakley, J., Estimating Percentiles of Uncertain Computer Code Outputs, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 53(1):83–93, 2004.
12. Hessling, J.P. and Uhlmann, J., ROBUSTNESS OF WILKS’ CONSERVATIVE ESTIMATE OF CONFIDENCE INTERVALS, *International Journal for Uncertainty Quantification*, 5(6):569–583, 2015.
13. Huber, P.J., The use of Choquet capacities in statistics, *Bulletin of the International Statistical Institute*, 45(4), 1973.

14. Ruggeri, F., Rios Insua, D., and Martin, J. Robust Bayesian Analysis. In: Dey, D.K. and Rao, C.R. (Eds.), *Handbook of Statistics*, Vol. 25 of Bayesian Thinking, pp. 623–667. Elsevier, January 2005.
15. DeRoberts, L. and Hartigan, J.A., Bayesian Inference Using Intervals of Measures, *The Annals of Statistics*, 9(2):235–244, March 1981.
16. Sivaganesan, S. and Berger, J.O., Ranges of Posterior Measures for Priors with Unimodal Contaminations, *The Annals of Statistics*, 17(2):868–889, June 1989.
17. Owhadi, H., Scovel, C., Sullivan, T.J., McKerns, M., and Ortiz, M., Optimal Uncertainty Quantification, *SIAM Review*, 55(2):271–345, January 2013, arXiv: 1009.0679.
18. Winkler, G., Extreme Points of Moment Sets, *Math. Oper. Res.*, 13(4):581–587, November 1988.
19. Henrion, D., Lasserre, J.B., and Lfberg, J., GloptiPoly 3: moments, optimization and semidefinite programming, *Optimization Methods and Software*, 24(4-5):761–779, October 2009.
20. Betrò, B. *Robust Bayesian Analysis*, chapter 15, Methods for Global Prior Robustness under Generalized Moment Conditions. Lecture Notes in Statistics. Springer-Verlag, New York, 2000.
21. Lasserre, J.B., *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, 2010. OCLC: ocn503631126.
22. McKerns, M., Owhadi, H., Scovel, C., Sullivan, T.J., and Ortiz, M., The optimal uncertainty algorithm in the mystic framework, *CoRR*, abs/1202.1055, 2012.
23. Dette, H. and Studden, W.J., *The Theory of Canonical Moments with Applications in Statistics, Probability, and Analysis*, Wiley-Blackwell, New York, September 1997.
24. Skibinsky, M., The range of the $(n + 1)$ th moment for distributions on $[0,1]$, *Journal of Applied Probability*, 4(3):543–552, November 1967.
25. Skibinsky, M., 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, 1977.
26. Skibinsky, M., Principal representations and canonical moment sequences for distributions on an interval, *Journal of Mathematical Analysis and Applications*, 120(1):95–118, November 1986.
27. Hausdorff, F., Momentprobleme für ein endliches Intervall, *Mathematische Zeitschrift*, 16:220–248, 1923.
28. Royden, H., *Real analysis*, Macmillan, 1968.
29. Price, K., Storn, R.M., and Lampinen, J.A., *Differential Evolution: A Practical Approach to Global Optimization*, Natural Computing Series, Springer-Verlag, Berlin Heidelberg, 2005.
30. Pasanisi, A., Keller, M., and Parent, E., Estimation of a quantity of interest in uncertainty analysis: Some help from Bayesian decision theory, *Reliability Engineering & System Safety*, 100:93–101, April 2012.
31. Loepky, J.L., Sacks, J., and Welch, W.J., Choosing the Sample Size of a Computer Experiment: A Practical Guide, *Technometrics*, 51(4).
32. Da Veiga, S., Global Sensitivity Analysis with Dependence Measures, *Journal of Statistical Computation and Simulation*, 85, November 2013.
33. Raguét, H. and Marrel, A., Target and conditional sensitivity analysis with emphasis on dependence measures, *arXiv preprint arXiv:1801.10047*, 2018.
34. Damblin, G. and Gaillard, P., Bayesian inference and non-linear extensions of the circe method for quantifying the uncertainty of closure relationships integrated into thermal-hydraulic system codes, *arXiv preprint arXiv:1902.04931*, 2019.
35. De Crécy, A., Determination of the uncertainties of the constitutive relationships of the cathare 2 code, *M&C*, 3, 10 2001.
36. Gratiot, L.L., Marelli, S., and Sudret, B., Metamodel-Based Sensitivity Analysis: Polynomial Chaos Expansions and Gaussian Processes, *Handbook of Uncertainty Quantification*, pp. 1289–1325, 2017.
37. Rockafellar, R.T. and Royset, J.O., Random Variables, Monotone Relations, and Convex Analysis, *Math. Program.*, 148(1-2):297–331, December 2014.
38. Berger, J.O., *Statistical Decision Theory and Bayesian Analysis*, Springer Series in Statistics, Springer-Verlag, New York, 2 edition, 1985.