



**HAL**  
open science

# An information geometry approach for robustness analysis in uncertainty quantification of computer codes

Clement Gauchy, Jerome Stenger, Roman Sueur, Bertrand Iooss

## ► To cite this version:

Clement Gauchy, Jerome Stenger, Roman Sueur, Bertrand Iooss. An information geometry approach for robustness analysis in uncertainty quantification of computer codes. *Technometrics*, 2022, 64, pp.80-91. hal-02425477v3

**HAL Id: hal-02425477**

**<https://hal.science/hal-02425477v3>**

Submitted on 12 Dec 2020

**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.

# An Information Geometry Approach to Robustness Analysis for the Uncertainty Quantification of Computer Codes

Clement GAUCHY<sup>1\*</sup>

Jerome STENGER<sup>12</sup>

Roman SUEUR<sup>1</sup>

Bertrand IOOSS<sup>12†</sup>

<sup>1</sup>EDF R&D, Département PRISME, 6 Quai Watier, 78401, Chatou, France

<sup>2</sup>Institut de Mathématiques de Toulouse, 31062 Toulouse, France

December 12, 2020

## Abstract

Robustness analysis is an emerging field in the uncertainty quantification domain. It involves analyzing the response of a computer model—which has inputs whose exact values are unknown—to the perturbation of one or several of its input distributions. Practical robustness analysis methods therefore require a coherent methodology for perturbing distributions; we present here one such rigorous method, based on the Fisher distance on manifolds of probability distributions. Further, we provide a numerical method to calculate perturbed densities in practice which comes from Lagrangian mechanics and involves solving a system of ordinary differential equations. The method introduced for perturbations is then used to compute quantile-related robustness indices. We illustrate these “perturbed-law based” indices on several numerical models. We also apply our methods to an industrial setting: the simulation of a loss of coolant accident in a nuclear reactor, where several dozen of the model’s physical parameters are not known exactly, and where limited knowledge on their distributions is available.

*Keywords:* Computer experiments, Density perturbation, Fisher metric, Importance sampling, Quantile, Sensitivity analysis

---

\*Present affiliation: DES/ISAS - Service d’études mécaniques et thermiques (SEMT), CEA Saclay, 91191 Gif-sur-Yvette, France

†Corresponding author: [bertrand.iooss@edf.fr](mailto:bertrand.iooss@edf.fr)

# 1 Introduction

Over the last few decades, two major trends in industrial and research practices have led to a rise in the importance of uncertainty quantification (UQ) methods (de Rocquigny et al., 2008; Smith, 2014; Ghanem et al., 2017). The first is the replacement of full-scale physical experiments, considered costly and difficult to implement, by numerical models. This choice raises the issue of a potential mismatch between computer codes and the physical reality they aim to simulate. The second trend involves accounting for risk in an increasing number of industrial activities, and in particular, doing it quantitatively.

In both situations, UQ can be performed by considering the imperfectly known inputs of a computer code (given as a function  $G(\cdot)$ ) as a vector of random variables  $\mathbf{X} = (X_1, \dots, X_d)$ . The most widespread approach consists of running  $G(\cdot)$  with different combinations of inputs in accordance with their range of plausible values, in order to study the associated uncertainty found in the output  $Y = G(X_1, \dots, X_d)$ , or, in other settings, to estimate a specific quantity of interest (QoI). The latter means some statistical quantity derived from  $Y$ , e.g., the performance, via the mean of  $Y$ , or some risk criterion in terms of a high-level quantile.

For instance, the nuclear industry faces major issues as facilities age and regulatory authorities' requirements toughen (Bucalossi et al., 2010; Mousseau and Williams, 2017). One precise case is where operators have to study *loss of coolant accidents* (LOCA), which result in a break in the primary loop of pressurized water nuclear reactors. This scenario can be simulated using system thermal-hydraulic computer codes which involve dozens of physical parameters including such things as condensation and heat transfer coefficients (Mazgaj

et al., 2016; Sanchez-Saez et al., 2018). However, the values of many such parameters are known with only limited precision (Larget, 2019) as they are typically calculated by way of other quantities measured via small-scale physical experiments. Certain other variables can only be measured during periodic inspections, e.g., the characteristics of pumps in hydraulic systems.

Various methods from the UQ domain can be useful when dealing with such uncertainties in system safety analyses. First, certain methods aim to improve the exploration of the input domain  $\mathcal{X}$  by using design of experiments, e.g., space filling designs (Fang et al., 2006). These make it possible to cover an input domain as evenly as possible for a fixed number of code runs, and to limit unexplored areas. For the estimation of certain specific QoI such as the probability of threshold exceedance or an  $\alpha$ -order quantile of the output, Monte Carlo-type methods are often preferred. In particular, accelerated Monte Carlo methods (e.g., importance sampling or subset simulation) target the most informative areas of  $\mathcal{X}$  in the sampling algorithm in order to estimate the QoI while controlling its estimation error (Morio and Balesdent, 2016). As a preliminary or concomitant stage, global sensitivity analysis is also essential in order to eliminate non-influential parameters and rank influential ones according to their impact on the QoI (Iooss and Lemaître, 2015; Iooss and Marrel, 2019).

All of these approaches are useful for dealing with the existence of uncertainty in applied problems. However, industrial (e.g., nuclear facility) operators face the difficulty of justifying their risk assessment methods over and above providing simulation results. Risk assessment methods must demonstrate that they provide reliable safety margins. Doing so

means a probable overestimation of the likely risk after accounting for all known sources of uncertainty. This principle of conservatism, which can be easily implemented when dealing with very simple monotonic physical models, can be hard in practice to adapt to computer codes simulating complex and non-monotonic physical phenomena. Indeed, it is rarely simple to apply this principle when implementing UQ methods based on a set of computer experiments that provide a whole range of values for the output quantity  $Y$ .

To address this issue, a new UQ branch of robustness analysis has emerged in recent years in the field of sensitivity analysis. It consists of evaluating the impact of the choice of the inputs' distributions and, more precisely, involves an analysis of how values of the QoI vary with respect to these choices.

An initial strategy is to consider a whole set of input distributions and analyze the related output distributions. For global sensitivity analysis, Hart and Gremaud (2019) use “optimal perturbations” of probability density functions to analyze the robustness of variance-based sensitivity indices (called Sobol indices (Sobol, 1993)). Meynaoui et al. (2019) and Chabridon et al. (2018) propose approaches that deal with “second-level” uncertainty, i.e., uncertainty in the parameters of the input distributions. An alternative approach known as “optimal uncertainty quantification” avoids specifying the input probability distributions, transforming the problem into the definition of constraints on moments (Owhadi et al., 2013; Stenger et al., 2019). The latter is out of the scope of the present work; here we consider that the initial probability distributions input by the user remain important.

In real-world uncertainty quantification studies in engineering, input distributions are

typically truncated, as we are dealing with physical parameters that in reality belong to finite-dimensional compact sets. It is therefore natural to assume that there is no uncertainty in the support of input random variables. In this paper, we also assume that the inputs are mutually independent. Keeping in mind that our goal is to directly deal with input distributions (and not to consider second-level uncertainty), one particularly interesting solution has been proposed in the context of reliability-oriented sensitivity analysis by Lemaître (2014) (see also Lemaître et al. (2015); Sueur et al. (2016)): perturbed-law based indices (PLI). A density perturbation consists of replacing the density  $f_i$  of one input  $X_i$  by a perturbed one  $f_{i\delta}$ , where  $\delta \in \mathbb{R}$  represents a shift in the value of a specific moment (e.g., the mean or variance). Among all densities which have an equivalent shift in the chosen moment,  $f_{i\delta}$  is defined as that which minimizes the Kullback-Leibler divergence from  $f_i$ . This method has been applied to compute a QoI that corresponds to a failure probability (Iooss and Le Gratiet, 2019; Perrin and Defaux, 2019), a quantile (Sueur et al., 2017; Larget, 2019) and a superquantile (Iooss et al., 2020; Larget and Gautier, 2020).

However, this method is not fully satisfactory. Indeed, the interest of perturbing moments is arguable: perturbing the mean or variance of input distributions is essentially an arbitrary decision. Moreover, it has been observed that applying the same perturbation to two different parametric distributions can yield significant differences in the value of the Kullback-Leibler divergence (between the initial and perturbed density).

Another possibility described in Perrin and Defaux (2019) is to use an iso-probabilistic operator to transform all input random variables into standard Gaussian ones. This makes perturbations comparable when applied in this new space, but it remains difficult to trans-

late this interpretation back into the initial physical space, that which is of interest to users. We note in passing another type of robustness analysis that has been proposed in quantitative finance by Cont et al. (2010). These authors investigate whether the estimated QoI is sensitive to small perturbations of the empirical distribution function. To this end, they define the robustness of a QoI in terms of its continuity with respect to the Prokhorov distance on the set of integrable random variables.

The goal of the present paper is to propose a novel approach to perturbing probability distributions. It relies on density perturbations based on the Fisher distance (Costa et al., 2012) as a measure of dissimilarity between an initial density  $f_i$  and a perturbed one  $f_{i\delta}$ . This distance defines a geometry on spaces of probability measures known as *information geometry* (Nielsen, 2013). The statistical interpretation of the Fisher distance provides an equivalence between perturbations of non-homogeneous quantities, and consequently a coherent framework for robustness analysis. Before presenting this approach, we first review existing density perturbation methods in Section 2. Section 3 is dedicated to the description of our method and a discussion of numerical methods. Section 4 illustrates our density perturbation method for the practical robustness index (PLI). We apply this to a toy example function and an industrial case study in Section 5. This is followed by conclusions and research options going forward. The online supplementary materials contain four further sections describing: 1) some numerical results on Fisher spheres computations; 2) the *reverse importance sampling* method; 3) the theoretical properties of the PLI-quantile estimator; and 4) an application of the PLI to another toy model.

## 2 Previous approaches to density perturbation for uncertainty quantification robustness analysis

The method of Lemaître et al. (2015), later known as PLI by Sueur et al. (2016), is based on the idea of perturbing input densities. It aims to provide a practical counterpart to the general idea of analyzing the output QoI of a model in a UQ framework when one or several parameters of the input probabilistic model (considered as the reference) is/are modified. This strategy can be seen as a way to take into account an “error term” that could be added to an imperfectly-known input distribution.

### 2.1 Kullback-Leibler divergence minimization

To construct a perturbed distribution  $f_{i\delta}$  from a distribution  $f_i$ , the approach of Lemaître et al. (2015) is nonparametric. This method essentially aims to analyze perturbations in the most common characteristics of input distributions such as the mean or variance. To illustrate this for the mean, let us assume the random variable  $X_i \sim f_i$  has mean  $\mathbb{E}[X_i] = \mu$ . By definition, the perturbed density will have a mean of  $\mu + \delta$ . However, this is clearly insufficient to fully determine the perturbed distribution and in particular explicitly access the value of  $f_{i\delta}$  over the whole domain of  $X_i$ . Amongst all densities with a mean of  $\mu + \delta$ ,  $f_{i\delta}$  is defined as the solution of the minimization problem:

$$f_{i\delta} = \arg \min_{\pi \in \mathcal{P}, s.t. \mathbb{E}_\pi[\mathbf{X}_i] = \mathbb{E}_{f_i}[X_i] + \delta} KL(\pi || f_i), \quad (1)$$

where  $\mathcal{P}$  is the set of all probability measures which are absolutely continuous with respect to  $f_i$ . This approach basically consists of perturbing the chosen parameter while chang-



ing the initial model as little as possible. Under this definition, “changing” the model corresponds to increasing the entropy, where the Kullback-Leibler divergence between two densities  $f$  and  $\pi$  is given by:

$$KL(\pi||f) = \int \log \left( \frac{\pi(x)}{f(x)} \right) f(x) dx . \quad (2)$$

This method can be applied to higher-order moments (for instance moments of order 2—to define a variance perturbation) and, more generally, to constraints that can be expressed as a function of the perturbed density, such as quantiles (Lemaître, 2014). Notice that in the case of an initial Gaussian distribution, the perturbed distribution remains Gaussian with simply a change in the mean or variance.

In the general case, this method has several drawbacks. First, the likelihood ratio between  $f_{i\delta}$  and  $f_i$  (see section 4.1) may not have an analytic form, which leads to numerical difficulties. Second, this method requires that moments be defined for the initial density. Third, the main difficulty concerns the interpretation of results. Indeed, each input containing uncertainty in the UQ model is perturbed using a range of values of  $\delta$ . However, the choice of the moment to be perturbed remains arbitrary. Such problems are particularly worrisome when uncertainty in the input parameter is epistemic (Kiureghian and Ditlevsen, 2009) (meaning it comes from a lack of knowledge). Perturbing only the mean or variance of such an input distribution may be overly limiting when trying to satisfy regulatory authorities focused on being overly conservative.

We also recall that all input random variables are assumed to be mutually independent. Nonetheless, the effect of perturbations can be considered only for each variable individually, and in absolute terms (the same  $\delta$  shift may have a quite different impact on different

input densities). This method thus proves challenging when one wants to compare the relative impact of perturbations on different inputs.

## 2.2 Standard space transformations

To interpret a  $\delta$  shift on an input distribution and in particular to enable inputs to be compared with respect to the impact on the QoI of the same perturbation, an equivalence criterion between inputs is required. An idea developed by Perrin and Defaux (2019) consists of applying perturbations in the so-called *standard space* (instead of the initial physical space) in which all input distributions are identical, thus making all perturbations equivalent. Then the perturbed densities are obtained by applying a reverse transform.

In the case of independent inputs, the required distribution transform is a simple inverse probability one. Given a real-valued random vector  $\mathbf{X} = (X_1, \dots, X_d)^T$  with cumulative distribution function (cdf)  $F$ , the transform is the random vector  $\mathbf{S} = \Phi^{-1}(F(\mathbf{X}))$ , where  $\Phi$  is the cdf of the standard Gaussian distribution  $\mathcal{N}(\mathbf{0}, \mathbf{I}_d)$ . Consequently,  $\mathbf{S}$  follows a standard Gaussian distribution whatever the initial distribution  $F$ . In the Kullback-Leibler divergence minimization framework (see Section 2.1), a perturbation of the mean in the standard space simply consists of a mean shift without the standard deviation changing. For instance, for a mean shift of  $\delta \in \mathbb{R}^d$ , the perturbed distribution is defined by the inverse transformation  $F_\delta = F^{-1}(\Phi(\mathbf{S} + \delta))$  in order to get back from the standard space to the physical space. This therefore leads to an analytic expression for the perturbed density  $f_{i\delta}$  in the one-dimensional case via the change of variable formula (Stirzaker, 2003, p.318):

$$f_{i\delta}(x) = e^{\frac{-\delta^2 + 2\delta\Phi^{-1}(F_i(x))}{2}} f_i(x), \quad (3)$$

where  $F_i$  is the cdf of the variable  $X_i$  for  $i \in \{1, \dots, d\}$ . This simple formula makes the perturbed density and likelihood ratio easy to compute.

However, similar perturbations in the standard space can lead to quite different ones in the physical space, depending on the initial distributions. As an example, Figure 1 depicts the Kullback-Leibler divergence (approximated with Simpson's rule (Abramowitz and Stegun, 1974)) of two specific distributions (the Triangular<sup>1</sup>  $\mathcal{T}(-1, 0, 1)$  and the Uniform  $\mathcal{U}[-1, 1]$ ) with their respective distributions in the standard space as the value of  $\delta$  changes. We see that the Kullback-Leibler divergence can behave very differently for different distributions in the physical space, even though the exact same perturbation has been applied in the standard space. Furthermore, there is no general rule for estimating the mean of a perturbed physical input in its physical space for a given mean perturbation of the input in the standard space. Such difficulties are amplified when considering perturbations involving moments other than the mean. For instance, there is no general equivalence in the physical space between perturbations applied to the mean and standard deviation of the same input probability distribution in the standard space. Overall, it would seem generally difficult, if not impossible, to have a simple way to convert results provided by this method into a relationship between input and output physical quantities, making such results difficult to interpret.

---

<sup>1</sup>The triangular distribution  $\mathcal{T}(-1, 0, 1)$  is parametrized by its minimum  $a$ , mode  $b$ , and maximum  $c$ .

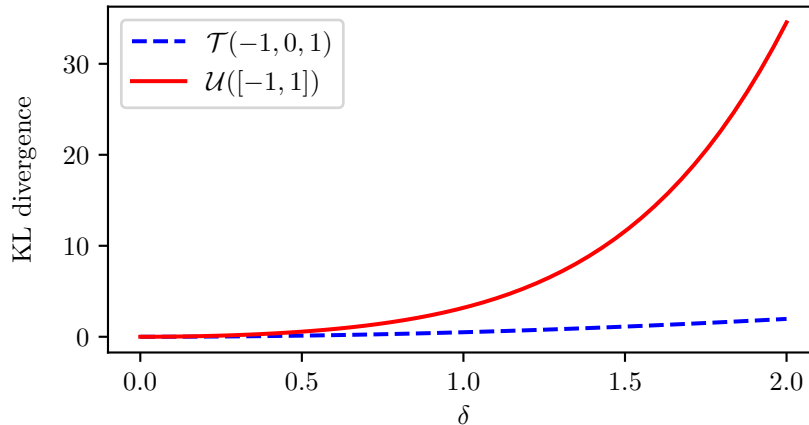


Figure 1: The Kullback-Leibler divergence between the initial distribution and the perturbed one for perturbations of the mean,  $\delta \in [0, 2]$ . Further details on perturbed distributions can be found in Lemaître (2014) and Lemaître et al. (2015).

### 3 A perturbation method based on information geometry

The Kullback-Leibler divergence can be interpreted as the power of a hypothesis test with null hypothesis: “ $X_i$  follows the distribution  $f_i$ ” and alternative hypothesis: “ $X_i$  follows the distribution  $f_{i\delta}$ ” (Eguchi and Copas, 2006). For this reason, it would appear to be an appropriate tool to measure how far a perturbed density is from the initial one and thus provides a formal counterpart to the vague idea of “uncertainty in the distribution”. However, in a practical robustness analysis context, its use often implies embedding inputs in a standard Gaussian space to begin with, in order to allow a comparison of perturbations of different input distributions, as well as to enable easy computation of perturbed densities. This transformation of the inputs leads to manipulating perturbations in a non-physical

input space, making it impossible to interpret them in terms of how they modify physical inputs' densities.

### 3.1 The Fisher distance

To try to form an intuitive understanding of the consequences of such perturbations on the distribution of the output variable  $Y$ , it is necessary to base our perturbation method on a metric which allows us to compare perturbations coming from different inputs distributions in the UQ model. For instance, if one input is Gaussian and another is log uniform, their respective perturbations (associated with the same level of perturbation) should be able to be interpreted in the same way. In particular, the perturbation method should result in identical perturbed densities when applied to two different parametric representations of the same input distribution. The Fisher distance (Rao, 1945) is consistent with these wishes. It is based on the local scalar product induced by the Fisher information matrix in a given parametric space, and defines a Riemannian geometry on the corresponding set of probability measures as on any Riemannian manifold with its associated metric.

Consider the family of parametric densities  $\mathcal{S} = \{f_\theta, \theta \in \Theta \subset \mathbb{R}^r\}$ . We recall that every input variable represents a physical parameter with known domain of validity, and therefore, for all  $\theta$  in  $\Theta$ , the support of  $f_\theta$  is assumed to be a compact set of  $\mathbb{R}$ . The metric associated with the coordinate function  $\theta$ , called the Fisher (or Fisher-Rao) metric, is defined as:

$$I(\theta_0) = \mathbb{E} \left[ \nabla_\theta \log f_\theta(X)|_{\theta=\theta_0} (\nabla_\theta \log f_\theta(X)|_{\theta=\theta_0})^T \right] ,$$

where  $I(\theta_0)$  is the Fisher information matrix evaluated at  $\theta_0$  for this statistical model.

The Fisher information, well-known in the fields of optimal design, Bayesian statistics, and machine learning, is a way of measuring the amount of information that an observable random variable  $X$  carries about an unknown parameter  $\theta$  of the distribution of  $X$ . The Fisher information matrix defines the following local inner product in  $\mathcal{S}$  for  $u \in \mathbb{R}^r$  and  $v \in \mathbb{R}^r$ :

$$\langle u, v \rangle_\theta = u^T I(\theta) v . \quad (4)$$

Given two distributions  $f_{\theta_0}$  and  $f_{\theta_1}$  in a manifold  $\mathcal{S}$ , a path from  $f_{\theta_0}$  to  $f_{\theta_1}$  is a piecewise smooth map  $q : [0, 1] \rightarrow \Theta$  satisfying  $q(0) = \theta_0$  and  $q(1) = \theta_1$ . Its length (Sternberg, 1999)  $l(q)$  satisfies the following equation:

$$l(q) = \int_0^1 \sqrt{\langle \dot{q}(t), \dot{q}(t) \rangle_{q(t)}} dt , \quad (5)$$

where  $\dot{q}$  is the derivative of  $q$ . Similarly, the energy (Sternberg, 1999)  $E(q)$  of a path is defined by the equation:

$$E(q) = \int_0^1 \frac{1}{2} \langle \dot{q}(t), \dot{q}(t) \rangle_{q(t)} dt . \quad (6)$$

The Fisher distance between  $f_{\theta_0}$  and  $f_{\theta_1}$  is defined as the minimal length over the set of paths  $\mathcal{P}(f_{\theta_0}, f_{\theta_1})$  from  $f_{\theta_0}$  to  $f_{\theta_1}$ :

$$d_F(f_{\theta_1}, f_{\theta_2}) = \inf_{q \in \mathcal{P}(f_{\theta_1}, f_{\theta_2})} l(q) . \quad (7)$$

The path  $\gamma$  minimizing this length—or equivalently minimizing the energy—is called a geodesic (Costa et al., 2012). The specific choice of the Fisher information matrix for a Riemannian metric matrix leads to a very interesting statistical interpretation, as shown in Amari (1985, p.27); it is directly related to the Cramer-Rao lower bound (Rao, 1945) which states that, for any unbiased estimator  $\hat{\theta}$  of  $\theta$ , the covariance matrix  $\text{Var}(\hat{\theta})$  is bounded by

$I(\theta)^{-1}$ . This means that the Fisher information is the maximum amount of information about the value of a parameter one can extract from a given sample. More formally, under some regularity conditions (given by (Newey and McFadden, 1994, Theorem 3.3)), if  $x_1, \dots, x_n$  are  $n$  independent observations distributed according to a density  $f_\theta$ , the maximum likelihood estimator  $\hat{\theta}_n$  of  $\theta$  converges weakly to a Gaussian distribution with mean  $\theta$  and covariance  $\frac{I(\theta)^{-1}}{n}$ . The density of  $\hat{\theta}_n$ , denoted  $p(\hat{\theta}_n, \theta)$  is written:

$$p(\hat{\theta}_n, \theta) = \frac{1}{\sqrt{\left(\frac{2\pi}{n}\right)^r \det(I(\theta)^{-1})}} \exp\left(-\frac{n(\hat{\theta}_n - \theta)^T I(\theta)(\hat{\theta}_n - \theta)}{2}\right). \quad (8)$$

When  $n \rightarrow +\infty$ , this probability density is proportional to  $(\hat{\theta}_n - \theta)^T I(\theta)(\hat{\theta}_n - \theta)$  (Amari, 1985), which is the local inner product defined in equation (4). This result can be interpreted as follows: the Fisher distance between two distributions  $f_\theta$  and  $f_{\theta'}$  represents the separability of the two distributions by a finite sample of independent observations sampled from the distribution  $f_\theta$  (Amari, 1985).

Let us illustrate the Fisher distance in a simple example. Consider the statistical manifold of univariate normal distributions  $\mathcal{S} = \{\mathcal{N}(\mu, \sigma^2), (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*\}$ . The Fisher information matrix has the analytical form (Costa et al., 2012):

$$I(\mu, \sigma) = \begin{pmatrix} 1/\sigma^2 & 0 \\ 0 & 2/\sigma^2 \end{pmatrix}. \quad (9)$$

If we apply the change of coordinates  $\phi(\mu, \sigma) \rightarrow \left(\frac{\mu}{\sqrt{2}}, \sigma\right)$ , the related geometry is the hyperbolic geometry in the Poincaré half-plane (Stillwell, 1997) in which the geodesic and distance between two normal distributions are known analytically (Costa et al., 2012). Geometrically, geodesics are the vertical lines and half-circle centered on the line  $\sigma = 0$ .

Further details on interpretations of information geometry can be found in Costa et al. (2012).

Figure 2 shows the position of four Gaussian distributions in the  $(\frac{\mu}{\sqrt{2}}, \sigma)$  half-plane. It is clear that the distributions  $C$  and  $D$  are more difficult to distinguish than  $A$  and  $B$ , though in both cases the 2-Wassertein distance  $W_2$  (Villani, 2009) is the same. The hyperbolic geometry induced by the Fisher information provides a representation in accordance with this intuition. Indeed, the two dashed curves are the geodesics between points  $A$  and  $B$ , and  $C$  and  $D$ . We observe that the Fisher distance between  $A$  and  $B$  is greater than that between  $C$  and  $D$ . This illustrates how information geometry provides a well-grounded framework to measure statistical dissimilarities in probability measure spaces.

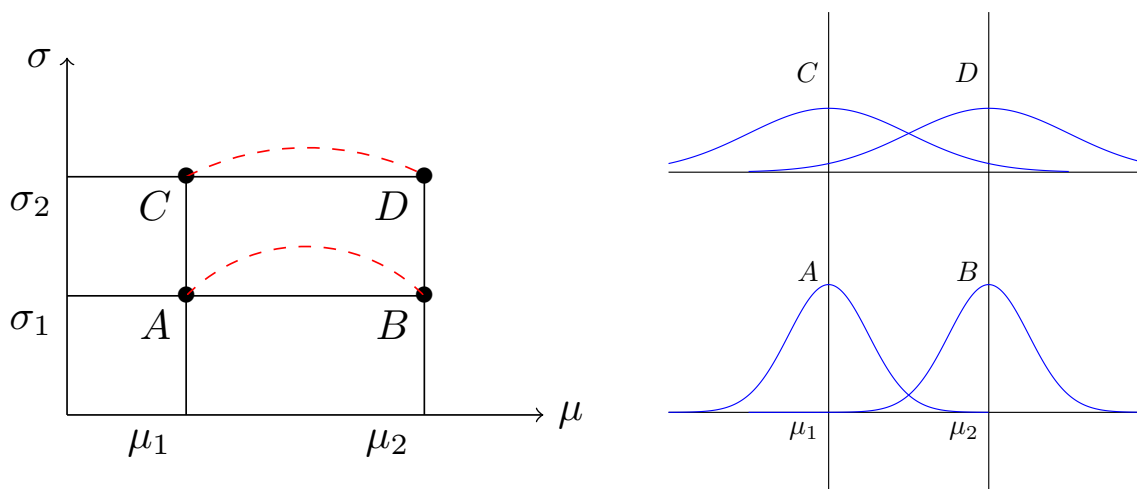


Figure 2: Four Gaussian distributions represented in the parameter space (left) and their respective distributions (right). Although  $W_2(A, B) = W_2(C, D)$ , it is easier to distinguish  $A$  from  $B$  than  $C$  from  $D$ . The dashed curved lines are geodesics of different length in the  $(\frac{\mu}{\sqrt{2}}, \sigma)$  plane.

In summary, the Fisher distance provides a satisfactory and coherent grounding for our



notion of density perturbation. Before continuing, let us define a perturbation of a density  $f$  to be of level  $\delta$  if the Fisher distance between  $f$  and the perturbed density  $f_\delta$  is equal to  $\delta$ . The set of all perturbations of  $f$  at level  $\delta$  is then the Fisher sphere of radius  $\delta$  centered at  $f$ , whenever this perturbation is applied to one or another of the parameters. This implies that, in this framework, we do not consider one specific perturbed distribution, but rather, a non-finite set of probability densities. The following section is dedicated to the development of a numerical method to compute Fisher spheres of radius  $\delta$  centered at  $f$ .

## 3.2 Computing Fisher spheres

As detailed in Section 3.1, geodesics are defined as the solution of a minimization problem. Specifically, a geodesic is a path with minimal length or energy (denoted  $E$ ). Given a smooth map  $q : [0, 1] \rightarrow \mathcal{S}$ , we have

$$E(q) = \int_0^1 \frac{1}{2} \langle \dot{q}(t), \dot{q}(t) \rangle_{q(t)} dt . \quad (10)$$

In the following we denote  $L(t, q, \dot{q}) = \frac{1}{2} \langle \dot{q}(t), \dot{q}(t) \rangle_{q(t)}$  the Lagrangian of the system. The energy of a path can be rewritten as

$$E(q) = \int_0^1 L(t, q, \dot{q}) dt . \quad (11)$$

A necessary condition for the path  $q$  to minimize the energy  $E$  is to satisfy the Euler-Lagrange equation (see Gelfand and Fomin (2012) for details):

$$\frac{\partial L}{\partial q} = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right) . \quad (12)$$

We denote  $p = \frac{\partial L}{\partial \dot{q}}$  and obtain by differentiation of the quadratic form  $L(t, q, \dot{q}) = \frac{1}{2} \dot{q}^T I(q) \dot{q}$  that  $p = I(q) \dot{q}$  and  $\dot{q} = I^{-1}(q)p$ . Then, inspired by Lagrangian mechanics theory (Arnold,

1997, p.65), the Hamiltonian  $H(p, q)$  defined by

$$\begin{aligned} H(p, q) &= p^T \dot{q} - L(t, q, \dot{q}) = p^T I^{-1}(q)p - \frac{1}{2} \dot{q}^T I(q) \dot{q} \\ &= \frac{1}{2} p^T I^{-1}(q)p \end{aligned} \tag{13}$$

is constant whenever  $q$  is a geodesic. Eq. (13) is derived from the Euler-Lagrange equation, and implies that  $(p, q)$  follows a system of ordinary differential equations (ODE) known as Hamilton's equations:

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p} = I^{-1}(q)p, \\ \dot{p} = -\frac{\partial H}{\partial q} = \frac{\partial L(t, q, I^{-1}(q)p)}{\partial q}. \end{cases} \tag{14}$$

The goal is to find any geodesics  $q$  that satisfy  $q(0) = \theta_0$  and  $d_F(f, q(1)) = \delta$ . This corresponds to computing the Fisher sphere centered at  $f_{\theta_0}$  with radius  $\delta$ . The only degree of freedom remaining to solve the ODE system (14) entirely is the initial velocity  $p(0)$ . Notice that the Hamiltonian is equal to the kinetic energy, since  $p = I(q)\dot{q}$ . As the Hamiltonian is constant on a geodesic, we have for all  $t$ :

$$\frac{1}{2} \langle \dot{q}(t), \dot{q}(t) \rangle_{q(t)} = k, \tag{15}$$

where  $k$  is non-negative. The length of  $q$  is therefore equal to

$$\int_0^1 \sqrt{\langle \dot{q}(t), \dot{q}(t) \rangle_{q(t)}} dt = \sqrt{2k}, \tag{16}$$

whereby  $\delta = \sqrt{2k}$ . Therefore, Eq. (13) can be rewritten:

$$\delta = \sqrt{2k} \iff p^T I^{-1}(q)p = \delta^2. \tag{17}$$

Taking equation (17) at the initial state  $t = 0$ , we can determine all of the initial velocities for which  $d_F(q(0), q(1)) = \delta$ . These are needed to solve the system of ODEs (14) and compute the geodesics.

Generally, computing the geodesic between two distributions is challenging. Methods relying on shooting algorithms have been developed for this problem (Le Brigant, 2019). However, in our framework, we don't want to compute the distance between two given distributions, which implies to find the particular geodesic joining these, but we aim at screening a Fisher sphere as well as possible for a fixed distance (as detailed in section 4.2). In the Section 1 of the online supplementary material, we focus on numerical methods for computing geodesics by solving the system of ODEs (14). We then illustrate these methods by computing Fisher spheres on the Gaussian manifold  $\mathcal{S} = \{\mathcal{N}(\mu, \sigma^2), (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*\}$ .

## 4 Applying this to perturbed-law based indices

The UQ robustness analysis detailed in Section 1 and Section 2 aims to quantify the impact of a lack of knowledge about an input distribution on the UQ of model outputs. In Section 3, a coherent formal definition of density perturbation was proposed. We now illustrate how to use this solution to define a practical robustness analysis method. First however, analyzing the effect of perturbing an input density requires the definition of some criterion or “index” to summarize the effects on a QoI.

### 4.1 Definition of perturbed-law based indices

A PLI aims to measure the impact of the modification of an input density on a QoI, such as a quantile or a threshold exceedance probability of the model output (Lemaître et al., 2015; Sueur et al., 2016). In the following, we focus on a quantile of order  $\alpha$ , often used as a risk measure in real-world applications (Mousseau and Williams, 2017; Delage et al.,

2018; Larget, 2019).

If we consider  $\mathcal{X}$  to be a compact set of  $\mathbb{R}^d$ , then given the random vector  $\mathbf{X} = (X_1, \dots, X_d) \in \mathcal{X}$  of our  $d$  independent imperfectly known input variables,  $G(\cdot)$  our numerical model, and  $Y = G(\mathbf{X}) \in \mathbb{R}$  the model output, the quantile of order  $\alpha$  of  $Y$  is given by:

$$q^\alpha = \inf\{t \in \mathbb{R}, F_Y(t) \geq \alpha\} , \quad (18)$$

where  $F_Y$  is the cdf of the random variable  $Y$ . In order to compute the  $i$ -th PLI (see below), we first change the density  $f_i$  of  $X_i$  into a density  $f_{i\delta}$ , where  $\delta \in \mathbb{R}^+$  represents the level of perturbation. The perturbed quantile is then given by:

$$q_{i\delta}^\alpha = \inf\{t \in \mathbb{R}, F_{Y,i\delta}(t) \geq \alpha\} , \quad (19)$$

where  $F_{Y,i\delta}$  is the cdf corresponding to the input variable  $X_i$  sampled from  $f_{i\delta}$ . The  $i$ -th PLI  $S_i$  is then simply defined as the relative change in the output quantile generated by the perturbation, i.e.,

$$S_i(f_{i\delta}) = \frac{q_{i\delta}^\alpha - q^\alpha}{q^\alpha} . \quad (20)$$

This definition slightly differs from that proposed in previous studies (Lemaître et al., 2015; Sueur et al., 2017), but is preferred here as it has been found to be more intuitive in a number of engineering studies that used PLI. Indeed, it simply consists of the relative variation of the quantile when the  $i$ -th input is submitted to a density perturbation and, as such, allows for a clearer interpretation.

In many applications including nuclear safety exercises, computer models are costly in terms of CPU time and memory. This means that only a limited number  $N$  of code runs

are available for estimating all of the PLIs. Thus, we have a sample  $\mathcal{Y}_N = \{y^{(n)}\}_{1 \leq n \leq N}$  of  $N$  outputs of the model from a sample  $\mathcal{X}_N = \{\mathbf{X}^{(n)} = (x_1^{(n)}, \dots, x_d^{(n)})\}_{1 \leq n \leq N}$  of  $N$  independent realizations of  $\mathbf{X}$ . The estimation of the quantile is then based on the empirical quantile estimator  $\widehat{q}_N^\alpha = \inf\{t \in \mathbb{R}, \widehat{F}_Y^N(t) \leq \alpha\}$ , where  $\widehat{F}_Y^N(t) = \frac{1}{N} \sum_{n=1}^N \mathbb{1}_{(y^{(n)} \leq t)}$  is the empirical estimator of the cdf of  $Y$ . In order to estimate the perturbed quantile  $\widehat{q}_{N,i\delta}^\alpha$  from the same sample  $\mathcal{X}_N$ , we use reverse importance sampling Hesterberg (1996) (see Section 2 of the online supplementary materials) to compute  $\widehat{F}_{Y,i\delta}^N$  (Delage et al., 2018):

$$\widehat{F}_{Y,i\delta}^N(t) = \frac{\sum_{n=1}^N L_i^{(n)} \mathbb{1}_{(y^{(n)} \leq t)}}{\sum_{n=1}^N L_i^{(n)}} , \quad (21)$$

with  $L_i^{(n)}$  the likelihood ratio  $\frac{f_{i\delta}(x_i^{(n)})}{f_i(x_i^{(n)})}$ . The estimator of the PLI is then

$$\widehat{S}_{N,i\delta} = \frac{\widehat{q}_{N,i\delta}^\alpha - \widehat{q}_N^\alpha}{\widehat{q}_N^\alpha} . \quad (22)$$

The theoretical properties of  $\widehat{S}_{N,i\delta}$  are studied in Section 3 of the online supplementary materials.

As presented in Section 3, the Fisher sphere of radius  $\delta$  centered on the initial input distribution  $f_i$  and written  $\partial\mathcal{B}_F(f_i, \delta) = \{g, d_F(f_i, g) = \delta\}$ , provides a good base for perturbing distributions. We do not consider one specific perturbation at level  $\delta$ , but instead a whole set of them:  $\partial\mathcal{B}_F(f_i, \delta)$ . Over this set, we compute the maximum  $S_{i\delta}^+$  and the minimum  $S_{i\delta}^-$  of the PLI:

$$S_{i\delta}^+ = \max_{g \in \partial\mathcal{B}_F(f_i, \delta)} S_i(g) , \quad (23)$$

$$S_{i\delta}^- = \min_{g \in \partial\mathcal{B}_F(f_i, \delta)} S_i(g) , \quad (24)$$

where  $S_i(g)$  is the PLI with  $g$  the perturbed density for the variable  $X_i$ .

Among all perturbed distributions at level  $\delta$ , we look for those which most deviate the quantile from its original value. These two quantities  $S_{i\delta}^+$  and  $S_{i\delta}^-$ , called OF-PLI (for ‘‘Optimal Fisher-based PLI’’), are measures of the robustness of the numerical code while taking into account uncertainty in the input distribution.

## 4.2 Practical implementation

The main problem when estimating the OF-PLI arises from the available sample size being finite, and possibly not as large as it needs to be. Indeed, due to inherent computational budget constraints in practice, there might not be enough sample points to correctly compute the perturbed quantile (and its confidence interval) at certain levels of perturbation. Thus, the key issue is to determine the extent to which an input distribution should be perturbed. To address this, we propose adapting the empirical criterion from Iooss et al. (2020) to help establish a maximal perturbed level  $\delta_{max}$ . For a proper estimation of OF-PLI indices, the number of points  $N_{\mathcal{Y}}$  in the output sample  $\mathcal{Y}_N$  exceeding the  $\delta$ -perturbed quantile needs to be sufficiently high. As the sample size is finite, the  $\delta$ -perturbed quantile can often not be computed accurately enough for high values of  $\delta$ . In practice, a value of  $N_{\mathcal{Y}} = 10$  has been decided on (after numerous numerical tests) as the smallest allowed for computing a PLI-quantile. As soon as a distribution on the Fisher sphere exceeds the criterion in (Iooss et al., 2020), the corresponding radius is taken as  $\delta_{max}$ .

The estimation of the quantities of interest  $S_{i\delta}^+$  and  $S_{i\delta}^-$  is described in Algorithm 1. We emphasize that this approach only applies to expensive computer models. Indeed, the bootstrap variance of the estimated quantile with reverse importance sampling tends to be

very large, as illustrated in Iooss et al. (2020). This is due to the likelihood ratio, which explodes locally. Thus, when dealing with a cheap code, one can directly resample over the perturbed distribution in order to estimate the output quantile. In this situation, there is no limiting level of perturbation  $\delta_{max}$ .

---

**Algorithm 1** Estimation of  $S_{i\delta}^+$  and  $S_{i\delta}^-$

---

1. Initialisation:  $\delta$  (level of perturbation),  $i \in \{1, \dots, d\}$  (input number), a sample of  $K$  points on the Fisher sphere of radius  $\delta$  centered at  $f_i$  (using the numerical method of Section 2 of the online supplementary material).
  2. For  $k = 1, \dots, K$ :
    - (a) estimate the  $\alpha$ -quantile  $q_{i\delta}^{\alpha, (k)}$  for the  $k$ -th perturbed density  $\{f_{i\delta}^{(k)}\}$  using reverse importance sampling based on the sample  $\mathcal{X}_N$ ,
    - (b) check whether the number of points in the output sample below or above the perturbed quantile  $q_{i\delta}^{\alpha, (k)}$  satisfies the stopping criterion  $N_y \geq 10$ ,
    - (c) compute the PLI estimator  $\widehat{S}_{N, i\delta}^{(k)}$ .
  3. the estimators  $\widehat{S}_{N, i\delta}^+$  and  $\widehat{S}_{N, i\delta}^-$  of the quantities of interest  $S_{i\delta}^+$  and  $S_{i\delta}^-$  are taken as the maximal and minimal values of the PLI sampled on the Fisher sphere  $\{\widehat{S}_{N, i\delta}^{(k)}\}$ .
- 

In future work, OF-PLI confidence intervals (computed via bootstrap) will provide valuable additional information like for instance confidence intervals, but further work on these is required and they are not included in the industrial application (section 5.2). The code for the OF-PLI is available at <https://github.com/JeromeStenger/PLI-Technometrics>.

The code for computing the older version of the PLI, called E-PLI (entropy-based PLI) in the following, is available in the `sensitivity` package of the software R.

## 5 Perturbed-law based indices in engineering studies

The PLI as defined here allows us to assess to what extent an output quantile can be affected by an error of magnitude  $\delta$  in the characterization of an input distribution. In the following section, we compare, using a toy example, our new method (OF-PLI) to the earlier one (E-PLI). Following that, we illustrate the use of OF-PLI in a nuclear safety analysis of a pressurized water nuclear reactor. Moreover, as PLIs are based on changes in input distributions, they differ from global sensitivity measures (Iooss and Lemaître, 2015), which evaluate the effect of input variability for fixed probabilistic models. To study potential coherence and/or divergence between the two approaches, we compare Sobol indices and OF-PLI on an analytic model in Section 4 of the online supplementary material.

### 5.1 A toy example: the Ishigami function

The Ishigami function (Ishigami and Homma, 1990) is often used as an example for uncertainty and sensitivity analysis methods, in particular because it exhibits strong non-linearity and non-monotonicity. In this section, we apply the methodology introduced in Section 4.2 to estimate the OF-PLI and compare our results to E-PLI. The Ishigami function, which takes three input random variables  $(X_1, X_2, X_3)$ , each distributed normally as  $\mathcal{N}(0, 1)$ , is defined with the following analytic formula:

$$G(x_1, x_2, x_3) = \sin(x_1) + 7 \sin(x_2)^2 + 0.1x_3^4 \sin(x_1) . \quad (25)$$



We intend to evaluate the impact of a perturbed input distribution on the value of the 95%-quantile. In this simple example where the function is cheap to evaluate, we do not use the reverse importance sampling estimator of the quantile as proposed in Section 4.2. Since the computational burden of running new code simulations is negligible, we instead draw new samples of size  $N = 2000$  directly from the perturbed input distributions in order to compute the output perturbed quantile. Hence, this toy example does not follow the reverse important sampling estimation procedure detailed in Section 4.2. The OF-PLI are computed for perturbation levels  $\delta$  from the interval  $[0, 0.9]$ . Notice that the choice  $\delta_{max} = 0.9$  is arbitrary here since the reverse importance sampling estimator has been put aside. Indeed, here there is no actual limit for the maximal perturbation level as the OF-PLI are computed by resampling from the perturbed distribution. We chose a value of  $K = 100$  trajectories over each Fisher sphere for computing the minimum and maximum of the OF-PLI. We also compute the 95%-confidence intervals calculated from 50 values of  $\widehat{S}_{N,i\delta}^+$  and  $\widehat{S}_{N,i\delta}^-$ .

The OF-PLI results are depicted in Figure 3. It appears that the third input has the most impact in shifting the quantile to the right. On the other hand, the second input has more impact in shifting the quantile to the left. These results are consistent with the well-known behavior of the Ishigami function in terms both of non-linearity of the model and primary influence of the third input.

Because the maximum and minimum of the OF-PLI are taken over the Fisher sphere, we depict in Figure 4 the distribution of the OF-PLI over the Fisher sphere with radius  $\delta = 0.9$  for the third input. One can see that in this case, the maximum and minimum are

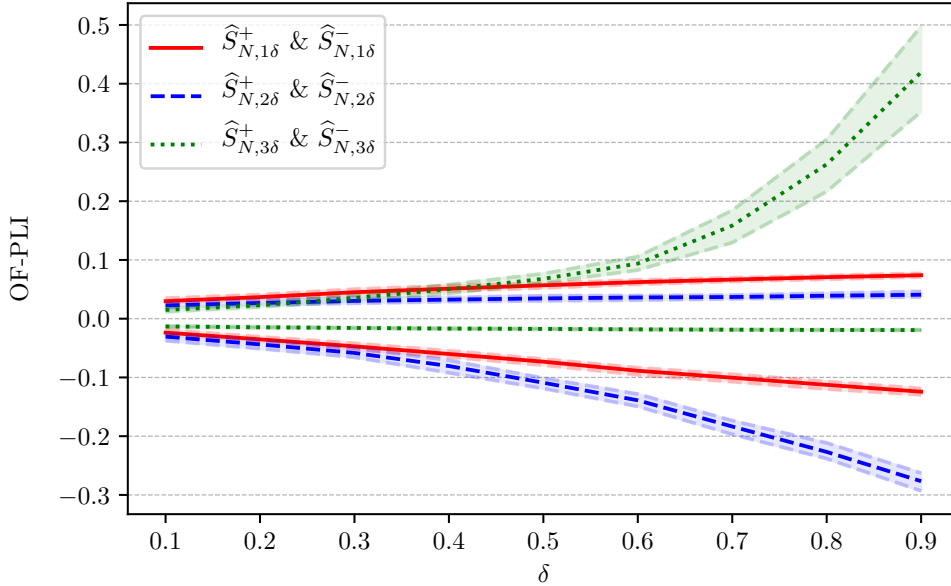


Figure 3: Minimum and maximum of the OF-PLI over the Fisher sphere with  $K = 100$  trajectories for  $\delta$  varying in  $[0, 0.9]$ , and their 95%-confidence intervals. For the Ishigami model, we directly resample from the perturbed distribution to compute the perturbed quantile instead of using the reverse important sampling estimator.

found respectively for high and low variance, with seemingly no change with respect to the mean.

These results can be compared to the E-PLI ones (see Section 2). As in this case, inputs follow normal distributions, applying an inverse cdf transformation is not necessary. Therefore, perturbing the mean (resp. the variance) of the input variable is equivalent to drawing straight horizontal (resp. vertical) trajectories in the parameter space (see Figure 1 of the online supplementary material). Results are depicted in Figure 5; the mean of the Gaussian is perturbed in  $[-1, 1]$  and its variance in  $[0, 4]$ . This corresponds to the range of variation of these parameters for the Fisher sphere radius varying in  $[0, 0.9]$ . We then

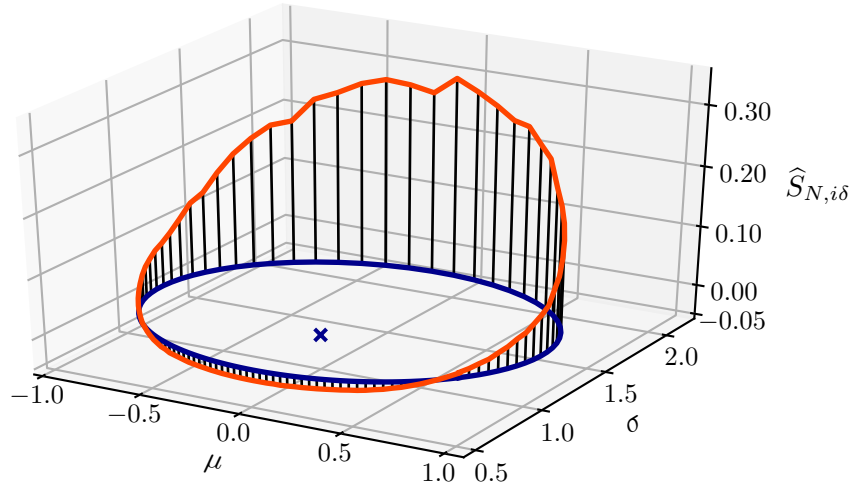


Figure 4: Value of the OF-PLI  $\hat{S}_{N,i\delta}$  (red line) for the third input of the Ishigami model ( $N = 100$ ,  $i = 3$ ) over a Fisher sphere of radius  $\delta = 0.9$  (blue line).

compare the third input for the two methods. We saw in Figure 4 that the maximal OF-PLI was obtained with high variance and no mean perturbation, which is coherent with the results in Figure 5. However, we miss the true impact a perturbed density can induce in situations where the maximal and minimal OF-PLI are not located in these two axes, such as for instance is the case for the first variable. Hence, the E-PLI, restricted to two directions in the Fisher sphere, has limited interpretation.

## 5.2 A nuclear safety application

This industrial application concerns the study of the peak cladding temperature (PCT) of fuel rods in the case of a loss of coolant accident caused by an intermediate-sized break in the primary loop (IB-LOCA) in a nuclear pressurized water reactor. According to operating rules, this temperature must remain below a threshold to prevent any deterioration of

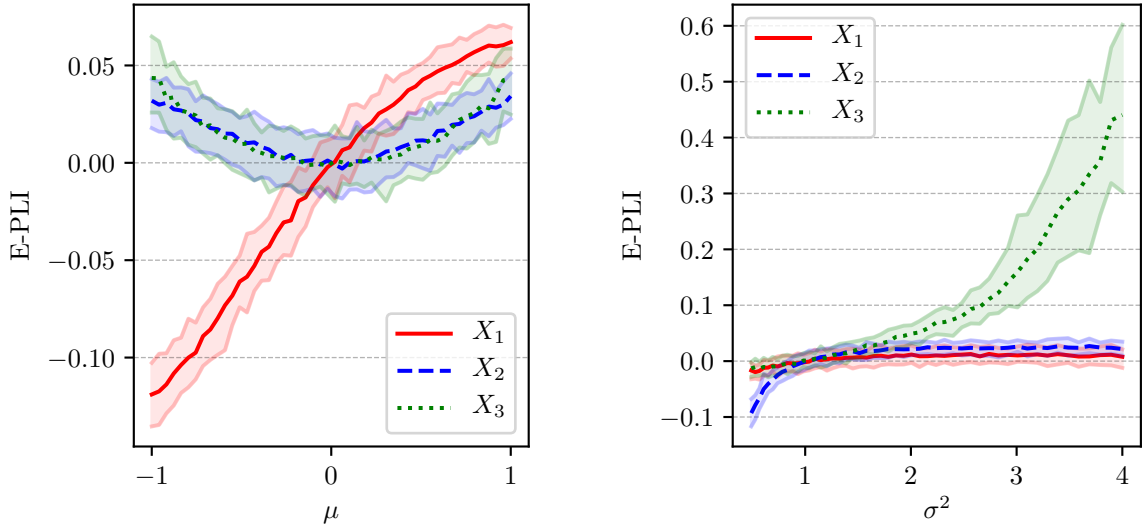


Figure 5: Computation of the E-PLI. Left: perturbation of the mean of the Gaussian distribution. Right: perturbation of the variance of the Gaussian distribution.

the state of the reactor. The thermal-hydraulic transient caused by this accidental scenario is simulated with the CATHARE2 code (Geffraye et al., 2011), providing a temperature profile throughout time for the surface of the nuclear core assemblies (Mazgaj et al., 2016). The thermal-hydraulic model involves boundary and initial conditions, and many physical parameters (heat transfer coefficient, friction coefficient, etc.) whose exact values are unknown. The probability distributions of these inputs can be obtained from data and expert knowledge, or by solving inverse problems on experimental data (Baccou et al., 2019).

Input uncertainties are propagated inside this model and the goal of UQ is to estimate a high-order quantile of the PCT (model output). This  $\alpha$ -quantile is interpreted as a pessimistic estimate of the PCT. Like any scientific approach, this methodology is based on hypotheses; regulatory authorities require an evaluation of the effect of these hypotheses on claimed results. Nuclear power operators are required to conduct studies in such a way

to ensure that actual risks are overestimated. Under this “conservative principle” they are required to choose the most pessimistic assumption each time a modeling decision has to be made. In deterministic studies, this simply consists of taking the most penalizing value for each input variable. In this way, the resulting computation is seen as simulating the worst case scenario for the risk being examined. It is, however, not straightforward to implement such a principle when the numerical code is complex, with both non-monotonic effects of inputs and interactions between inputs. It is even harder to extend this rationale to a UQ framework that aims to represent all potential scenarios with related occurrence likelihoods. Nevertheless, recent works (Larget, 2019) have shown that the E-PLI can be useful as evidence in discussions on the choice of input distributions.

In the present application, we study a reduced-scale mock-up of a pressurized water reactor with the 7 imperfectly known inputs given in Table 1 (Delage et al., 2018). To compute the OF-PLI, an input-output sample of size  $N = 1000$  is available, coming from a space filling design of experiments (Fang et al., 2006) (whose points in  $[0, 1]^d$  have been transformed to follow the input probability distributions). More precisely, a Latin hypercube sample minimizing the  $L^2$ -centered discrepancy criterion (Jin et al., 2005) has been used. The OF-PLI (with respect to a quantile of order  $\alpha = 0.95$ ) will then be estimated without any additional code run (see Section 4.1).

Figure 6 presents the maximum and minimum values of our two estimators  $\widehat{S}_{N,i\delta}^+$  and  $\widehat{S}_{N,i\delta}^-$ . We compute Fisher spheres with radius  $\delta$  sampled uniformly in  $[0.1, 0.5]$ , all respectively centered on the initial input distributions. The maximal perturbation level  $\delta_{max} = 0.5$  is derived from the stopping criterion introduced in Section 4.2. On every sphere,  $K = 100$

Table 1: Input variables of the CATHARE2 code with their associated probability distributions.

Variable number	Input name	Probability distribution
1	STMFSCO	Uniform $\mathcal{U}([-44.9, 63.5])$
2	STBAEBU	Truncated Log Normal $\mathcal{LN}(0, 0.76)$ on $[0.1, 10]$
3	STOIBC1	Truncated Log Normal $\mathcal{LN}(0, 0.76)$ on $[0.1, 10]$
4	STOIBC3	Truncated Log Normal $\mathcal{LN}(0, 0.76)$ on $[0.1, 10]$
5	STOIDC	Truncated Log Normal $\mathcal{LN}(0, 0.76)$ on $[0.1, 10]$
6	STOICO	Truncated Log Normal $\mathcal{LN}(-0.1, 0.45)$ on $[0.23, 3.45]$
7	CLFBR	Truncated Normal $\mathcal{N}(6.4, 4.27)$ on $[0, 12.8]$

perturbed densities are sampled. The OF-PLIs are finally estimated on a dataset of size 1000. The stopping criterion of Section 4.2 gives a maximal admissible OF-PLI of 4%; this value is determined from the maximal admissible quantile for which there are  $N_y = 10$  sample points above it. In fact, we see that  $\widehat{S}_{N,7\delta}^+$  is close to this maximal admissible value.

Studies conducted earlier for the same application (Delage et al., 2018) lead to similar results concerning the most influential inputs on the PCT’s quantile: strong influence of variables 3 and 4, weak influence of variables 1, 2, and 5. With respect to these previous studies based on the standard space transformation, our information geometry perturbation method evaluates the influence of variable 7 as being smaller. In fact, as it is the only Gaussian distribution, the reverse transformation from the standard space to the physical

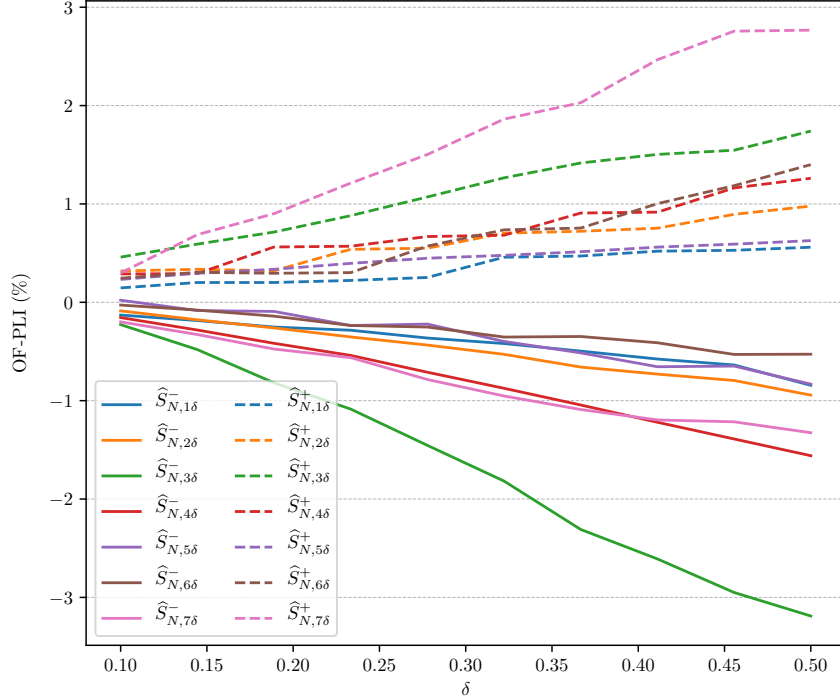


Figure 6: Bootstrap mean of the maximum and minimum of the OF-PLI  $S_{i\delta}$  for the CATHARE2 code. Confidence intervals are not shown for the sake of clarity.

one operates differently for this input than for the others. Overall, according to the values of  $\hat{S}_{N,3\delta}^+$  and  $\hat{S}_{N,7\delta}^+$ , the variables 3 and 7 appear to be the most influential inputs on the value of the PCT's quantile. This was not observed using the standard space transformation, and is probably due to the fact that the standard space approach allows for perturbing only one of the probability distribution parameters (for example the expected value) at a time. In contrast, our estimator corresponds to the maximal quantile deviation over an entire set of equivalent perturbations. This suggests two major advantages of our newly developed method: it (i) prevents interpretation bias induced by the standard space transformation, and (ii) allows for an exhaustive exploration of density perturbations for a given  $\delta$ .

## 6 Conclusion

Based on the Fisher distance, we have defined an original methodology for perturbing input probability distributions in the specific case of mutually independent input random variables. The Fisher information is an intrinsic characteristic of probability measures and, in particular, does not depend on the specific choice of parametric representation. This fundamental property makes it the proper mathematical tool to compare perturbations on different imperfectly known physical inputs of a computer model, and also on different parameters of the same input distribution. It is even possible to remove all reference to a parametric sub-domain of the set of probability measures on  $\mathcal{X}$ , given the nonparametric extension of the Fisher distance proposed by Holbrook et al. (2017). However, the latter is limited by practical issues as it supposes a finite-dimensional representation of the densities, for example by means of projection onto an orthonormal basis of the probability space. This implies truncating the infinite sum of the projections of a given probability on all elements of the basis. This approximation will be poor for probabilities which are very different from those of the chosen basis. This shows that in practice it is not easy to eliminate the reference to a particular parametric model, even in a nonparametric framework.

Nevertheless, based on PLIs, our method provides useful information on the most influential uncertainties related to the distributions of input variables, so-called “epistemic uncertainties”. This is in particular crucial not only in making decisions concerning further research programs aiming at gaining better knowledge about these variables, but also to bring strong backing arguments to operators’ safety demonstrations. Indeed, we argue that this methodology is adequate for uncertainty studies with unreliable input distribution



identification, or when an improved level of robustness is demanded for the choice of input distributions. In our target application (nuclear licensing), the aim is not only to exhibit safety margin values for simulated accidents, but also to prove that the methodology as a whole does not induce any risk of underestimating these values. Hence we are not only looking for a worst case assessment method, but also for a more global understanding of how a potential error in an input's distribution affects the output. In this sense, a practical option to increase conservatism in UQ studies is to replace one or several input distributions by penalized deterministic values, or by penalized versions of the distributions themselves. This nevertheless implies justifying the choice of the variables for which this penalization is done (see, e.g., Larget and Gautier (2020)).

Further investigations remain to be completed, due in part to this method increasing the numerical complexity and computational time (by several hours) required compared to the previous method of Lemaître et al. (2015). Several Monte Carlo loops are needed to compute the maximal and minimal PLI over Fisher spheres and we have ongoing work on the improvement of the estimation of the maximum and the minimum of the PLI on a Fisher sphere. In particular, using a better-adapted programming language, a large computational gain (of hours) could be expected. There is a known numerical issue with the reverse importance sampling strategy whereby the likelihood ratio tends to explode, as do the confidence intervals. Moreover, the method consists in sampling trajectories over the Fisher sphere, but one could benefit from a more advanced strategy by directly optimizing the PLI over the sphere via, for instance, gradient descent along this manifold. The crucial problem of probabilistic dependency between inputs needs also to be studied

in order to extend our framework to the non-independent input case; works in robustness analysis dealing with dependent inputs can be found for instance in Pesenti et al. (2019). Moreover, using a distance based on the Fisher information might be less intuitive for non-statisticians than simply shifting moments. In order to promote our methodology in engineering applications, it is essential to explicit its link with statistical test theory, thus helping practitioners interpreting the results it provides. Finally, future works will improve the estimation accuracy of the OF-PLI, especially for high values of  $\delta$  where numerical difficulties occurred, as mentioned in Section 1 of the online supplementary material.

# Online Supplementary Materials for the Paper “An Information Geometry Approach to Robustness Analysis for the Uncertainty Quantification of Computer Codes”

## 7 Computing Fisher spheres: Numerical results

Hamilton’s equations (see Section 3.2 of the main article),

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p} = I^{-1}(q)p, \\ \dot{p} = -\frac{\partial H}{\partial q} = \frac{\partial L(t, q, I^{-1}(q)p)}{\partial q}, \end{cases} \quad (26)$$

can be solved using numerical approximation methods. Figure 7 illustrates our numerical solutions in the Gaussian case, i.e., when  $\mathcal{S} = \{\mathcal{N}(\mu, \sigma^2), (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*\}$ . We compare the solution given by the explicit Euler algorithm to that given by the Adams-Moulton algorithm. We recall that in the Gaussian case we have at our disposal an exact analytical expression of the Fisher sphere, as detailed in Costa et al. (2012). The Fisher sphere is centered at  $\mathcal{N}(0, 1)$  with radius  $\delta = 1$ . Notice that there is no observable difference between the two methods in Figure 7. Hence, the numerical error is estimated from the variation of the Hamiltonian value that should be theoretically conserved along the geodesics. As a consequence, it is possible to quantify the performance of the numerical approximation by computing the value  $\Delta(t) = \frac{H(p(t), q(t)) - H(p(0), q(0))}{H(p(0), q(0))}$  for  $t \in [0, 1]$ . Here,  $\Delta$  represents the relative variation of the Hamiltonian along the path  $q$  computed with our numerical methods.

Figure 8 displays the value of  $\Delta(t)$  for  $t \in [0, 1]$  for one arbitrary geodesic from Figure

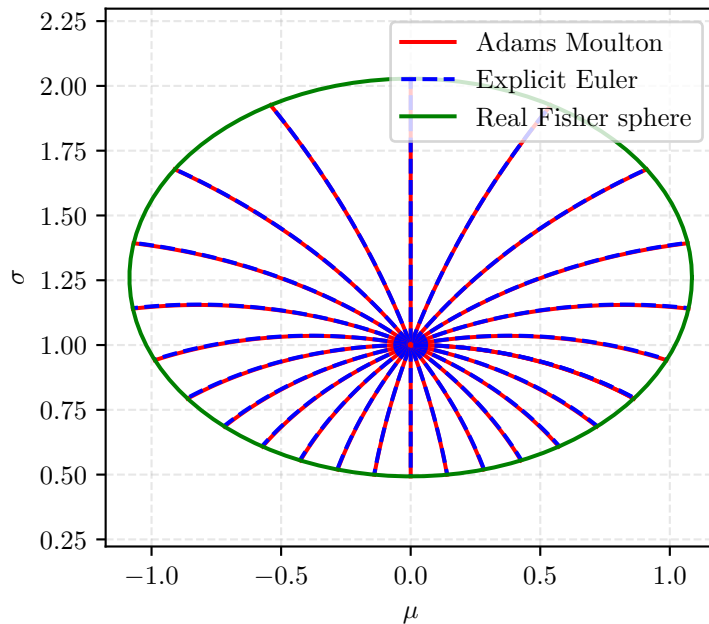


Figure 7: Geodesics in Gaussian information geometry computed with Euler’s explicit method and the Adams-Moulton method. The radius  $\delta$  is equal to 1.

7. The relative error for the Adams-Moulton method is negligible, while the maximum relative error for the explicit Euler scheme is around 0.3%. Hence, in the Gaussian case the Adams-Moulton scheme is preferable. However, instabilities have been observed in practice. Symplectic methods (Amari and Nagaoka, 2000; Leimkuhler and Reich, 2005), and in particular a symplectic Euler algorithm, could help to alleviate this problem by forcing the Hamiltonian to be constant. This will be the subject of future work. Note also that truncation can lead to other numerical errors when the radius  $\delta$  is too large. Indeed, the normalization constant of some truncated distributions can sometimes become smaller than a computer’s machine precision.

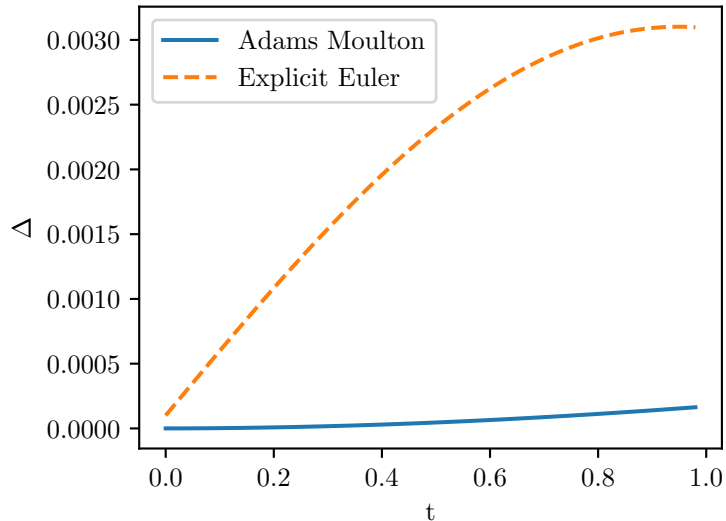


Figure 8: Relative variation of the Hamiltonian  $\Delta$  along a geodesic for two different numerical methods.

## 8 The reverse importance sampling method

Importance sampling (IS) is a Monte Carlo technique to compute statistical quantities such as  $\mathbb{E}[\phi(X)]$  ( $\mathbf{X} = (X_1, \dots, X_d)$ ) with a sample  $\tilde{\mathcal{X}}_N = \{\tilde{X}^{(1)}, \dots, \tilde{X}^{(N)}\}$  following a different distribution from that of  $X$ . If we call  $\pi$  the density of  $X$  and  $\tilde{\pi}$  the sampling one, we have

$$\mathbb{E}_\pi[\phi(X)] = \int \phi(x)\pi(x)dx = \int \phi(x)\tilde{\pi}(x)\frac{\pi(x)}{\tilde{\pi}(x)}dx = \mathbb{E}_{\tilde{\pi}}[\phi(X)L(X)],$$

where  $L = \pi/\tilde{\pi}$  is the likelihood ratio allowing us to convert an integral under the distribution of  $X$  into an integral under the sampling one. This leads to estimate of  $M = \mathbb{E}[\phi(X)]$  via:

$$\hat{M}_{IS} = \frac{1}{N} \sum_{n=1}^N \phi(\tilde{X}^{(n)})L(\tilde{X}^{(n)}), \quad (27)$$

with an estimation error given by  $\text{Var}[\hat{M}_{IS}] = N^{-1}\text{Var}_{\tilde{\pi}}[\phi L]$ .

The use of an alternative density (called an *instrumental density*) is most often a vari-

ance reduction method (Kahn and Marshall, 1953), as the variance of  $\hat{M}$  can be much lower than that of the usual empirical mean for a well chosen  $\tilde{\pi}$ . Indeed, an optimal variance reduction can be obtained for a sampling distribution  $\propto \pi\phi$ , which however implies that we know the value of  $\mathbb{E}[\phi]$  we are aiming to estimate.

The principle is similar if we aim to estimate quantiles corresponding to a perturbed density, though from a sample generated from the initial density of  $X$  (Hesterberg, 1996). Here the IS is called “reverse IS” as we do not change the sampling density to achieve a better estimation of the initial quantity, but instead change the quantity to be estimated with the same initial sample. Our purpose is to estimate a perturbed quantile of  $Y$ , e.g., a quantile of the distribution of  $Y_{i\delta} = G(X_1, \dots, X_{i-1}, \tilde{X}_i, X_{i+1}, \dots, X_d)$ , where the  $i$ -th input  $\tilde{X}_i$  is perturbed and has a density of  $f_{i\delta}$  instead of the initial one.

Hence the initial density of  $X$  can be seen as the instrumental density of a typical IS-Monte Carlo scheme, and the perturbed density is the one under which we must estimate the cdf  $F_{Y,i\delta}$ . If we denote  $\pi(x_1, \dots, x_d) = \prod_{i=1}^d f_i(x_i)$  the initial density and  $\tilde{\pi}(x_1, \dots, x_d) = f_1(x_1) \times \dots \times f_{i-1}(x_{i-1}) \times f_{i\delta}(x_i) \times f_{i+1}(x_{i+1}) \times \dots \times f_d(x_d)$  the perturbed one as regards the  $i$ -th input, and recalling that  $F_{Y,i\delta}(y) = \mathbb{E}_{\tilde{\pi}}[\mathbb{1}_{(G(X) \leq y)}]$ , we have the following estimator for the perturbed cdf:

$$\hat{F}_{Y,i\delta}^N(y) = \frac{1}{N} \sum_{n=1}^N L_i^{(n)} \mathbb{1}_{(y^{(n)} \leq y)}. \quad (28)$$

In this estimator, illustrated in Figure 9, we recognize that of formula (27) with  $\phi(\cdot) = \mathbb{1}_{(G(\cdot) \leq y)}$  and  $L(x_1^{(n)}, \dots, x_d^{(n)}) = \pi(x_1^{(n)}, \dots, x_d^{(n)}) / \tilde{\pi}(x_1^{(n)}, \dots, x_d^{(n)}) = f_i(x_i^{(n)}) / f_{i\delta}(x_i^{(n)}) = L_i^{(n)}$ .

The estimator  $\hat{q}_{N,i\delta}^\alpha$  of the  $\alpha$ -order quantile of  $Y_{i\delta}$  is then obtained by plugging  $\hat{F}_{Y,i\delta}^N$  into

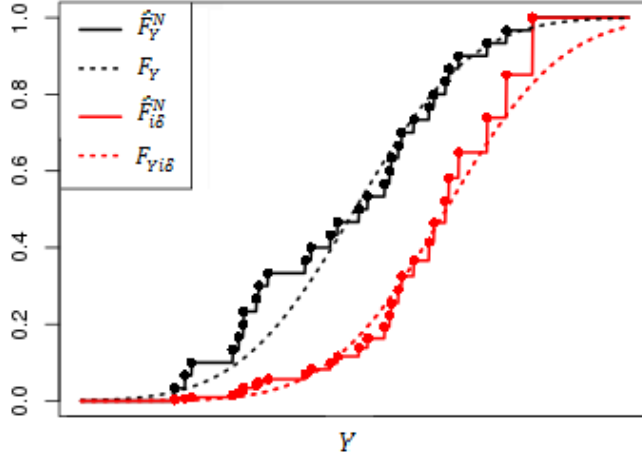


Figure 9: Empirical (solid lines) and theoretical (dotted lines) cumulative distribution functions. The samples (at the points' locations) used to estimate the initial distribution function  $F_Y$  (black lines) and the perturbed one  $F_{Y,i\delta}$  (red lines) are the same. Steps are all equal to  $1/N$  for the initial cumulative distribution function, whereas they are weighted by the likelihood ratio in  $\widehat{F}_{Y,i\delta}^N$ .

the definition of the theoretical quantile  $q_{i\delta}^\alpha = \inf\{t \in \mathbb{R}, F_{Y,i\delta}(t) \geq \alpha\}$ , which gives:

$$\widehat{q}_{N,i\delta}^\alpha = \inf\{t \in \mathbb{R}, \widehat{F}_{Y,i\delta}^N(t) \geq \alpha\}.$$

Note that the estimator (28) of  $F_{Y,i\delta}$  is normalized by  $1/N$ , which slightly differs from the one presented here, which is normalized by  $1/\sum_{n=1}^N L_i^{(n)}$ , as well as the one plotted in Figure 9. This guarantees  $\widehat{F}_{Y,i\delta}^N(y) \rightarrow 1$  when  $y \rightarrow +\infty$  for finite values of  $N$  without changing the asymptotic properties of the quantile estimator (Glynn, 1996).

## 9 Theoretical properties of the PLI-quantile estimator

In this section, we investigate some theoretical aspects of the PLI estimator  $\widehat{S}_{N,i\delta}$ . As it is based on quantile estimators, we first focus on the asymptotic properties of the estimator  $(\widehat{q}_N^\alpha, \widehat{q}_{N,i\delta}^\alpha)$ .

**Theorem 1.** *Suppose that  $F_Y$  is differentiable at  $q^\alpha = F_Y^{-1}(\alpha)$  with  $F'_Y(q^\alpha) > 0$  and that  $F_{Y,i\delta}$  is differentiable at  $q_{i\delta}^\alpha = F_{Y,i\delta}^{-1}(\alpha)$  with  $F'_{Y,i\delta}(q_{i\delta}^\alpha) > 0$ . We define:*

$$\sigma_i^2 = \frac{\alpha(1-\alpha)}{F'_Y(q^\alpha)^2}, \quad (29)$$

$$\tilde{\sigma}_{i\delta}^2 = \frac{\mathbb{E} \left[ \left( \frac{f_{i\delta}(X_i)}{f_i(X_i)} \right)^2 (\mathbb{1}_{(G(\mathbf{X}) \leq q_{i\delta}^\alpha)} - \alpha)^2 \right]}{F'_{Y,i\delta}(q_{i\delta}^\alpha)^2},$$

$$\tilde{\theta}_i = \frac{\mathbb{E} \left[ \frac{f_{i\delta}(X_i)}{f_i(X_i)} \mathbb{1}_{(G(\mathbf{X}) \leq q^\alpha)} \mathbb{1}_{(G(\mathbf{X}) \leq q_{i\delta}^\alpha)} \right] - \alpha \mathbb{E}[\mathbb{1}_{(G(\mathbf{X}) \leq q_{i\delta}^\alpha)}]}{F'_Y(q^\alpha) F'_{Y,i\delta}(q_{i\delta}^\alpha)}.$$

Let us write the covariance matrix  $\Sigma$  as:

$$\Sigma = \begin{pmatrix} \sigma^2 & \tilde{\theta}_i \\ \tilde{\theta}_i & \tilde{\sigma}_{i\delta}^2 \end{pmatrix},$$

and suppose that it is invertible and  $\mathbb{E} \left[ \left( \frac{f_{i\delta}(X_i)}{f_i(X_i)} \right)^3 \right] < +\infty$ . Then

$$\sqrt{N} \left( \begin{pmatrix} \widehat{q}_N^\alpha \\ \widehat{q}_{N,i\delta}^\alpha \end{pmatrix} - \begin{pmatrix} q^\alpha \\ q_{i\delta}^\alpha \end{pmatrix} \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma).$$

*Proof.* We study the consistency and asymptotic normality of specific  $M$  and  $Z$ -estimators in order to establish the proof. We suppose the theory around these estimators is known so



that details can be kept to the bare minimum. Further details can be found in Chapters 5.2 and 5.3 of Van der Vaart (2000). Given a sample  $(\mathbf{X}^{(n)})_{n \in \{1, \dots, N\}}$  where  $\mathbf{X}$  is a  $d$ -dimensional random vector, we define

$$\begin{aligned}
\eta &= \frac{\alpha}{1 - \alpha}, \\
m_\theta(x) &= -(G(x) - \theta)\mathbf{1}_{(G(x) \leq \theta)} + \eta(G(x) - \theta)\mathbf{1}_{(G(x) > \theta)}, \\
M_N(\theta_1, \theta_2) &= \frac{1}{N} \sum_{n=1}^N m_{\theta_1}(\mathbf{X}^{(n)}) + \frac{f_{i\delta}(X_i^{(n)})}{f_i(X_i^{(n)})} m_{\theta_2}(\mathbf{X}^{(n)}), \\
\hat{\theta}_N &= \arg \max M_N(\theta_1, \theta_2).
\end{aligned} \tag{30}$$

$\hat{\theta}_N$  is defined such that its two components correspond respectively to the estimators  $\hat{q}_N^\alpha$  and  $\hat{q}_{N,i\delta}^\alpha$  of the quantile and the perturbed quantile. The map  $\theta \mapsto \nabla_\theta M_N(\theta)$  with  $\theta = (\theta_1, \theta_2)^T$  has two non-decreasing components (it is a sum of non-decreasing maps). Now, by definition of  $\hat{\theta}_N$  and concavity of  $M_n(\theta)$ , it holds that  $\nabla_\theta M_N(\hat{\theta}_N) = 0$ . Furthermore, we have that  $\nabla_\theta M_N(\theta) \xrightarrow{P} [(1 + \eta)F_Y(\theta_1) - \eta, ((1 + \eta)F_{Y,i\delta}(\theta_2) - \bar{L}_N \eta)]^T$  with  $\bar{L}_N = \frac{1}{N} \sum_{n=1}^N \frac{f_{i\delta}(X_i^{(n)})}{f_i(X_i^{(n)})}$ , and this limit is a strictly non-decreasing function. Therefore, the assumptions of Lemma 5.10 in (Van der Vaart, 2000, p.47) are satisfied, proving the consistency of the estimator  $\hat{\theta}_N \xrightarrow{P} (q^\alpha, q_{i\delta}^\alpha)^T$ .

The asymptotic normality is studied via the map  $\bar{m}_\theta(x) \mapsto m_{\theta_1}(x) + \frac{f_{i\delta}(x)}{f_i(x)} m_{\theta_2}(x)$  which is Lipschitz for the variable  $\theta$  with Lipschitz constant  $h(x) = \max(1, \eta) \left(1 + \frac{f_{i\delta}(x_i)}{f_i(x_i)}\right)$ . The function  $h$  belongs in  $L^2$  if  $\mathbb{E} \left[ \left( \frac{f_{i\delta}(X_i)}{f_i(X_i)} \right)^2 \right] < +\infty$ . The map  $\bar{m}_\theta$  is also differentiable at  $\theta_0 = \arg \max_{\theta \in \Theta} \mathbb{E}[\bar{m}_\theta(X)]$  with gradient:

$$\nabla_{\theta_0} \bar{m}_{\theta_0}(x) = ((1 + \eta)\mathbf{1}_{(G(\mathbf{x}) \leq \theta_1)} - \eta, \frac{f_{i\delta}(x_i)}{f_i(x_i)} ((1 + \eta)\mathbf{1}_{(G(\mathbf{x}) \leq \theta_2)} - \eta))^T. \tag{31}$$

Moreover, the map  $\theta \rightarrow \mathbb{E}[\bar{m}_\theta(\mathbf{X})]$  has the following Hessian:

$$V_{\theta_0} = \begin{pmatrix} (1 + \eta)F'_Y(q^\alpha) & 0 \\ 0 & (1 + \eta)F'_{Y,i\delta}(q_{i\delta}^\alpha) \end{pmatrix}, \quad (32)$$

which is non-negative definite symmetric whenever  $F'_Y(q^\alpha) > 0$  and  $F'_{Y,i\delta}(q_{i\delta}^\alpha) > 0$ . Hence, Theorem 5.23 in (Van der Vaart, 2000, p.53) applies. This proves the asymptotic normality of the estimator  $(\hat{q}^\alpha, \hat{q}_{i\delta}^\alpha)^T$ .  $\square$

The PLI  $S_{i\delta}$  is a straightforward transformation of  $(q^\alpha, q_{i\delta}^\alpha)^T$ . To obtain the almost sure convergence of  $\hat{S}_{N,i\delta}$  to  $S_{i\delta}$ , it suffices to apply the continuous mapping theorem to the function  $s(x, y) = \frac{y - x}{x}$ .

**Theorem 2.** *Under the assumptions of theorem 1, we have*

$$\sqrt{N}(\hat{S}_{N,i\delta} - S_{i\delta}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, d_s^T \Sigma d_s), \text{ with } d_s = \begin{pmatrix} -q^\alpha / q_{i\delta}^{\alpha^2} \\ 1/q^\alpha \end{pmatrix}. \quad (33)$$

Notice that the asymptotic variance relies on the initial  $\alpha$ -quantile and the perturbed quantile, which are precisely what we want to estimate. Hence, Theorem 2 cannot be used for building asymptotic confidence intervals. However, its convergence properties are important for our method's credibility and acceptance. In practice, the estimation error can be measured using bootstrapping (Efron, 1979).

## 10 Applying PLI to an analytic flood risk model

The model of interest concerns a flooded river simulation, which is especially useful in assessing the risk of submergence of a dike protecting industrial sites near a river. To this

end, we use a model implementing a simplified version of the 1D hydro-dynamic Saint-Venant equations. This model computes  $H$ , the maximal annual water level of the river, from four parameters  $Q$ ,  $K_s$ ,  $Z_m$  and  $Z_v$ , which are considered imperfectly known:

$$H = \left( \frac{Q}{300K_s\sqrt{2.10^{-4}(Z_m - Z_v)}} \right)^{0.6}. \quad (34)$$

The inputs are modeled as random variables with the associated truncated distributions given in Table 2 (Iooss and Lemaître, 2015).

Table 2: Input variables of the flood model with their associated probability distributions.

Input n°	Name	Description	Probability distribution	Truncation
1	$Q$	Maximal annual flowrate	Gumbel $\mathcal{G}(1013, 558)$	[500, 3000]
2	$K_s$	Strickler coefficient	Normal $\mathcal{N}(30, 7.5)$	[15, $+\infty$ ]
3	$Z_v$	River downstream level	Triangular $\mathcal{T}(50)$	[49, 51]
4	$Z_m$	River upstream level	Triangular $\mathcal{T}(55)$	[54, 56]

In global sensitivity analyses, Sobol indices are the most popular sensitivity measures because they are easy to interpret; each Sobol index represents a share of the output variance, and all indices sum to 1 under the assumption of independent inputs (Sobol', 2001; Saltelli and Tarantola, 2002; Prieur and Tarantola, 2017). We will compare these to the results of our robustness analysis framework in order to illustrate differences. Note however that these conventional Sobol indices focus on the central part of the distribution (variance of the output). We then also compute the target Sobol indices (Marrel and Chabridon, 2020), i.e., Sobol indices applied to the indicator function of exceeding a given threshold (chosen here as the 95%-quantile of the output). To compute the first order and

total Sobol indices of the inputs of the flood model (Eq. (34)), the asymptotically efficient pick-freeze estimator (Priour and Tarantola, 2017) is used with an elementary Monte Carlo matrix of size  $10^6$ . This gives a total cost of  $N = 6 \times 10^6$  model runs and a standard deviation of the indices' estimation error smaller than  $10^{-3}$ . As shown in Table 3, in the central part of the distribution (conventional Sobol indices), we observe that the variable  $Q$  is clearly more influential than the variable  $K_s$ , whereas  $Z_v$  and  $Z_m$  appear to have almost no influence on the output. From the target Sobol indices, we observe that in the extreme part of the distribution (close to the 95%-quantile),  $Q$  and  $K_s$  have the same total effect (due to a strong interaction effect between them, which comes from the fact that threshold exceedence is principally sensitive to the joint values of these two inputs).

Table 3: Sobol index estimates of the flood model inputs.

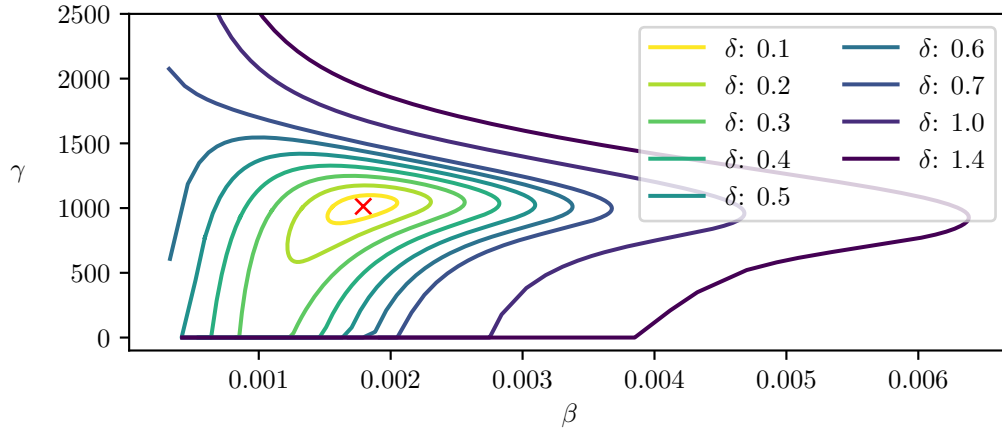
Inputs	$Q$	$K_s$	$Z_v$	$Z_m$
First-order Sobol indices	0.713	0.254	0.006	0.006
Total Sobol indices	0.731	0.271	0.008	0.008
First-order target Sobol indices	0.242	0.125	0.002	0.002
Total target Sobol indices	0.867	0.739	0.119	0.121

We compute the OF-PLI (w.r.t. a quantile of order  $\alpha = 0.95$ ) for the flood model inputs by increasing the Fisher spheres radii  $\delta \in [0, 1.4]$  by steps of 0.1. The spheres are respectively centered on the distributions shown in Table 2. On each of these spheres, we compute the OF-PLI for  $K = 100$  different perturbed distributions using a sample of  $N = 2000$  points distributed according to the initial distribution. The maximal radius  $\delta_{max} = 1.4$

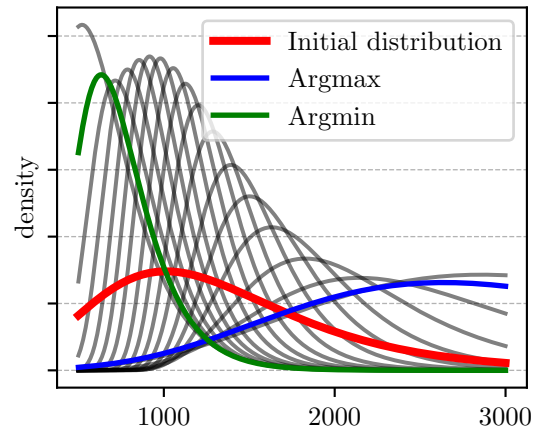
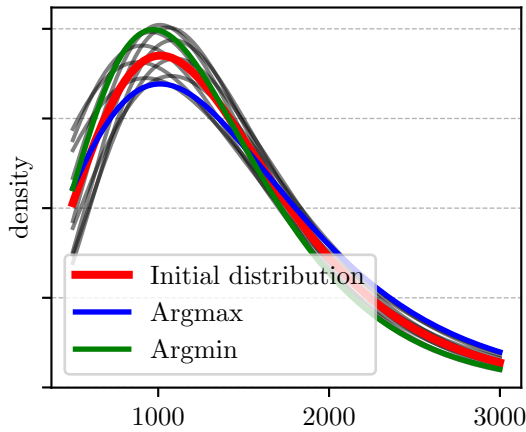
comes from the stopping criteria explained in Section 4.2 (“Practical implementation”) of the paper. More precisely, the criterion is hit for the first input  $Q$  at perturbation level  $\delta > 1.4$ , meaning there are less than  $N_{\gamma} = 10$  sample points above the maximal perturbed quantile.

Figure 10 depicts how the Fisher sphere centered at the variable  $Q$  deforms, and how the perturbed densities are spread out around the initial distribution. Figures 10b and 10c indicate that the maximal values of the OF-PLI are obtained for densities allocating the largest probabilities to large input values (blue curves), whereas the minimal values are obtained for densities with the least weighted tails (green curves). This behavior was predictable here since the river height is an increasing function of river flow (see Eq. (34)), though this type of analysis can in practice provide substantial information in real-world engineering studies. Finally, we also observe that Fisher spheres sometimes have nontrivial shapes, as illustrated in Fig. 10a. In this plot, we see that Fisher spheres are non-closed curves for radius  $0.3 \leq \delta$ , the parameter’s domain being bounded by 0 for  $\beta$  and  $\gamma$ . Such behaviour is specific to each parametric family of densities and, for instance, is not observed for non-truncated normal distributions.

The results of the OF-PLI, displayed in Figure 11, confirm those of the target Sobol indices (see Table 3): the variables 3 and 4, corresponding to  $Z_v$  and  $Z_m$ , are much less influential on the output quantile of level  $\alpha = 0.95$  than the variables 1 and 2, corresponding to  $Q$  and  $K_s$ . It turns out that in this case study, perturbations of  $Q$  and  $K_s$  appear to have comparable effects on the 95%-quantile of  $H$ , although they contribute quite differently to the output variance. On the other hand, compared to target Sobol indices, OF-PLI provide



(a) Deformation of the Fisher sphere for increasing radius  $\delta$ .



(b) Densities over the Fisher sphere ( $\delta = 0.1$ ).

(c) Densities over the Fisher sphere ( $\delta = 1.4$ ).

Figure 10: Analysis of the Fisher metric-based perturbation of the truncated Gumbel distribution of the variable  $Q$  (see Table 2).

more informative results with their evolution as a function of  $\delta$ . This clearly shows how a lack of knowledge on an imperfectly known input may turn out to have anything from a high to low impact on the value of a risk measure. Note also that the flat segments visible on some curves are due to approximation errors attributed to the low number of

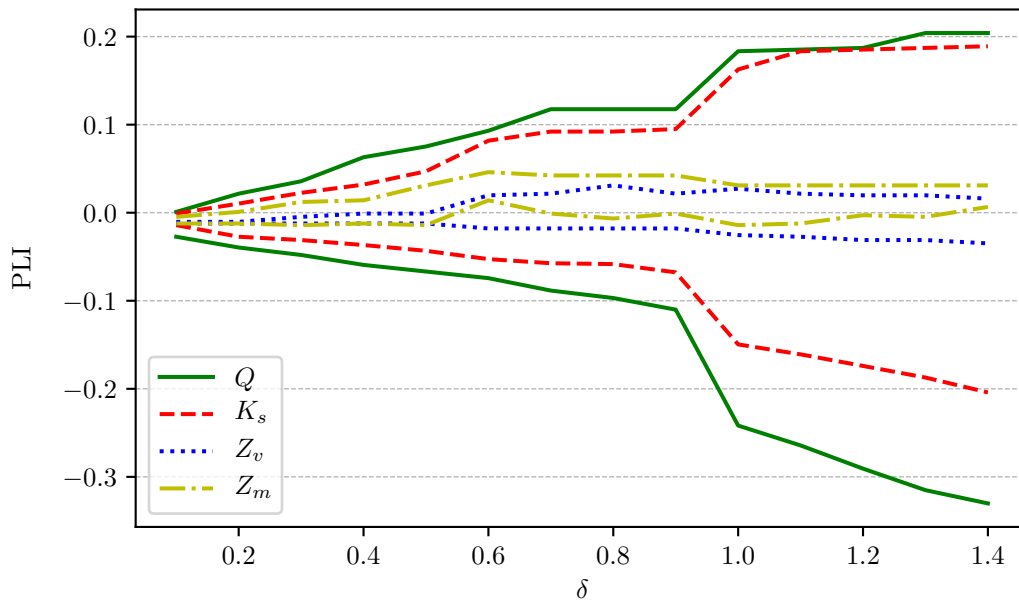


Figure 11: Maximum and minimum estimated value of the OF-PLI  $\widehat{S}_{N,i\delta}^+$  and  $\widehat{S}_{N,i\delta}^-$  for the different variables of the flood model.

sample points  $N$  and the high quantile level (0.95). In conclusion, this example confirms the interest of OF-PLI in that it conveys useful complementary information to existing sensitivity indices.

## References

Abramowitz, M. and Stegun, I. (Editors) (1974), *Handbook of Mathematical Functions*, Dover Publications, Inc. New York.

Amari, S. (1985), *Differential-Geometrical Methods in Statistics*, New York, NY: Springer New York.

- Amari, S. and Nagaoka, H. (2000), *Methods of Information Geometry*, Oxford University Press.
- Arnold, V. I. (1997), *Mathematical Methods of Classical Mechanics (Graduate Texts in Mathematics, Vol. 60)*, Springer.
- Baccou, J., Zhang, J., Fillion, P., Damblin, G., Petruzzi, A., Mendizábal, R., Reventós, F., Skorek, T., Couplet, M., Iooss, B., Oh, D., and Takeda, T. (2019), “Development of good practice guidance for quantification of thermal-hydraulic code model input uncertainty,” *Nuclear Engineering and Design*, 354, 110173.
- Bucalossi, A., Petruzzi, A., Kristof, M., and D’Auria, F. (2010), “Comparison between best-estimate-plus-uncertainty methods and conservative tools for nuclear power plant licensing,” *Nuclear Technology*, 172, 29–47.
- Chabridon, V., Balesdent, M., Bourinet, J.-M., Morio, J., and Gayton, N. (2018), “Reliability-based sensitivity estimators of rare event probability in the presence of distribution parameter uncertainty,” *Reliability Engineering & System Safety*, 178, 164–178.
- Cont, R., Deguest, R., and Scandolo, G. (2010), “Robustness and sensitivity analysis of risk measurement procedures,” *Quantitative Finance*, 10, 593 – 606.
- Costa, S. I. R., Santos, S. A., and Strapasson, J. E. (2012), “Fisher information distance: a geometrical reading,” *Discrete Applied Mathematics*, 197, 59–69.
- de Rocquigny, E., Devictor, N., and Tarantola, S. (Editors) (2008), *Uncertainty in industrial practice*, Wiley.



- Delage, T., Sueur, R., and Iooss, B. (2018), “Robustness analysis of epistemic uncertainties propagation studies in LOCA assessment thermal-hydraulic model,” in *Proceedings of ANS Best Estimate Plus Uncertainty International Conference (BEPU 2018)*, Lucca, Italy.
- Efron, B. (1979), “Bootstrap Methods: Another Look at the Jackknife,” *The Annals of Statistics*, 7, 1–26.
- Eguchi, S. and Copas, J. (2006), “Interpreting Kullback-Leibler Divergence with the Neyman-Pearson Lemma,” *Journal of Multivariate Analysis*, 97, 2034–2040.
- Fang, K.-T., Li, R., and Sudjianto, A. (2006), *Design and modeling for computer experiments*, Chapman & Hall/CRC.
- Geffraye, G., Antoni, O., Farvacque, M., Kadri, D., Laviolle, G., Rameau, B., and Ruby, A. (2011), “CATHARE2 V2.5\_2: A single version for various applications,” *Nuclear Engineering and Design*, 241, 4456–4463.
- Gelfand, I. and Fomin, S. (2012), *Calculus of Variations*, Dover Publications.
- Ghanem, R., Higdon, D., and Owhadi, H. (Editors) (2017), *Springer Handbook on Uncertainty Quantification*, Springer.
- Glynn, P. W. (1996), “Importance Sampling for Monte Carlo Estimation of Quantiles,” in *Proceedings of Second International Workshop on Mathematical Methods in Stochastic Simulation and Experimental Design*, Publishing House of Saint Petersburg University.

- Hart, J. and Gremaud, P. (2019), “Robustness of the Sobol’ indices to marginal distribution uncertainty,” *SIAM/ASA Journal on Uncertainty Quantification*, 7, 1224–1244.
- Hesterberg, T. (1996), “Estimates and confidence intervals for importance sampling sensitivity analysis,” *Math. Comput. Modelling*, 23, 79–85.
- Holbrook, A., Lan, S., Streets, J., and Shahbaba, B. (2017), “The nonparametric Fisher geometry and the chi-square process density prior,” *arXiv e-prints*, arXiv:1707.03117.
- Iooss, B. and Le Gratiet, L. (2019), “Uncertainty and sensitivity analysis of functional risk curves based on Gaussian processes,” *Reliability Engineering and System Safety*, 187, 58–66.
- Iooss, B. and Lemaître, P. (2015), “A review on global sensitivity analysis methods,” in Meloni, C. and Dellino, G. (editors), *Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications*, Springer.
- Iooss, B. and Marrel, A. (2019), “Advanced methodology for uncertainty propagation in computer experiments with large number of inputs,” *Nuclear Technology*, 205, 1588–1606.
- Iooss, B., Vergès, V., and Larget, V. (2020), “BEPU robustness analysis via perturbed-law based sensitivity indices,” in *Accepted to the ANS Best Estimate Plus Uncertainty International Conference (BEPU 2020)*, Giardini Naxos, Italy, URL <https://hal.archives-ouvertes.fr/hal-02864053>.
- Ishigami, T. and Homma, T. (1990), “An importance quantification technique in uncer-

- tainty analysis for computer models,” in [1990] *Proceedings. First International Symposium on Uncertainty Modeling and Analysis*, IEEE Comput. Soc. Press.
- Jin, R., Chen, W., and Sudjianto, A. (2005), “An efficient algorithm for constructing optimal design of computer experiments,” *Journal of Statistical Planning and Inference*, 134, 268–287.
- Kahn, H. and Marshall, A. W. (1953), “Methods of Reducing Sample Size in Monte Carlo Computations,” *Journal of the Operations Research Society of America*, 1, 263–278.
- Kiureghian, A. D. and Ditlevsen, O. (2009), “Aleatory or epistemic? Does it matter?” *Structural Safety*, 31, 105 – 112, URL <http://www.sciencedirect.com/science/article/pii/S0167473008000556>. Risk Acceptance and Risk Communication.
- Larget, V. (2019), “How to bring conservatism to a BEPU analysis,” in *NURETH-18*, Portland, USA.
- Larget, V. and Gautier, M. (2020), “Increasing conservatism in BEPU IB LOCA safety studies using complementary and industrially cost effective statistical tools,” in *Accepted to the ANS Best Estimate Plus Uncertainty International Conference (BEPU 2020)*, Giardini Naxos, Italy.
- Le Brigant, A. (2019), “A Discrete Framework to Find the Optimal Matching Between Manifold-Valued Curves,” *Journal of Mathematical Imaging and Vision*, 61, 40–70.
- Leimkuhler, B. and Reich, S. (2005), *Simulating Hamiltonian Dynamics*, Cambridge University Press.

- Lemaître, P. (2014), *Analyse de sensibilité en fiabilité des structures - Sensitivity analysis in structural reliability*, Thèse de l'Université Bordeaux I.
- Lemaître, P., Sergienko, E., Arnaud, A., Bousquet, N., Gamboa, F., and Iooss, B. (2015), “Density modification based reliability sensitivity analysis,” *Journal of Statistical Computation and Simulation*, 85, 1200–1223.
- Marrel, A. and Chabridon, V. (2020), “Statistical developments for target and conditional sensitivity analysis: application on safety studies for nuclear reactor,” *Preprint*, <https://hal.archives-ouvertes.fr/hal-02541142>.
- Mazgaj, P., Vacher, J.-L., and Carnevali, S. (2016), “Comparison of CATHARE results with the experimental results of cold leg intermediate break LOCA obtained during ROSA-2/LSTF test 7,” *EPJ Nuclear Sciences & Technology*, 2.
- Meynaoui, A., Marrel, A., and Laurent, B. (2019), “New statistical methodology for second level global sensitivity analysis,” URL <https://hal.archives-ouvertes.fr/hal-02019412>. Working paper or preprint.
- Morio, J. and Balesdent, M. (2016), *Estimation of rare event probabilities in complex aerospace and other systems*, Woodhead Publishing.
- Mousseau, V. and Williams, B. (2017), “Uncertainty quantification in a regulatory environment,” in Ghanem, R., Higdon, D., and Owhadi, H. (editors), *Springer Handbook on Uncertainty Quantification*, Springer.
- Newey, W. K. and McFadden, D. (1994), “Large sample estimation and hypothesis testing,”

- in *Handbook of Econometrics*, volume 4, chapter 36, Elsevier, 2111 – 2245. ISSN: 1573-4412.
- Nielsen, F. (2013), “Cramér-Rao Lower Bound and Information Geometry,” in Bhatia, R., Rajan, C. S., and Singh, A. I. (editors), *Connected at Infinity II: A Selection of Mathematics by Indians*, Gurgaon: Hindustan Book Agency.
- Owhadi, H., Scovel, C., Sullivan, T. J., McKerns, M., and Ortiz, M. (2013), “Optimal Uncertainty Quantification,” *SIAM Review*, 55, 271–345.
- Perrin, G. and Defaux, G. (2019), “Efficient estimation of reliability-oriented sensitivity indices,” *Journal of Scientific Computing*, 80.
- Pesenti, S., Millosovich, P., and Tsanakas, A. (2019), “Cascade Sensitivity Measures,” *Available at SSRN*.
- Prieur, C. and Tarantola, S. (2017), “Variance-Based Sensitivity Analysis: Theory and Estimation Algorithms,” in Ghanem, R., Higdon, D., and Owhadi, H. (editors), *Springer Handbook on Uncertainty Quantification*, Springer.
- Rao, C. (1945), “Information and the accuracy attainable in the estimation of statistical parameters,” *Bull. Calcutta Math. Soc.*, 37.
- Saltelli, A. and Tarantola, S. (2002), “On the relative importance of input factors in mathematical models: Safety assessment for nuclear waste disposal,” *Journal of American Statistical Association*, 97, 702–709.

- Sanchez-Saez, F., Sánchez, A., Villanueva, J., Carlos, S., and Martorell, S. (2018), “Uncertainty analysis of large break loss of coolant accident in a pressurized water reactor using non-parametric methods,” *Reliability Engineering and System Safety*, 174, 19–28.
- Smith, R. (2014), *Uncertainty quantification*, SIAM.
- Sobol, I. (1993), “Sensitivity estimates for non linear mathematical models,” *Mathematical Modelling and Computational Experiments*, 1, 407–414.
- Sobol’, I. (2001), “Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates,” *Mathematics and Computers in Simulation*, 55, 271 – 280. The Second IMACS Seminar on Monte Carlo Methods.
- Stenger, J., Gamboa, F., Keller, M., and Iooss, B. (2019), “Optimal Uncertainty Quantification of a risk measurement from a thermal-hydraulic code using Canonical Moments,” *International Journal for Uncertainty Quantification*, 10, 35–53.
- Sternberg, S. (1999), *Lectures on Differential Geometry*, AMS Chelsea Publishing Series, American Mathematical Society.
- Stillwell, J. (1997), *Numbers and Geometry (Undergraduate Texts in Mathematics)*, Springer.
- Stirzaker, D. (2003), *Elementary Probability*, Cambridge University Press.
- Sueur, R., Bousquet, N., Iooss, B., and Bect, J. (2016), “Perturbed-Law based sensitivity indices for sensitivity analysis in structural reliability,” in *Proceedings of the 8th Interna-*

*tional Conference on Sensitivity Analysis of Model Output (SAMO 2016)*, Le Tampon, Réunion Island, France.

Sueur, R., Iooss, B., and Delage, T. (2017), “Sensitivity analysis using perturbed-law based indices for quantiles and application to an industrial case,” in *Proceedings of the 10th International Conference on Mathematical Methods in Reliability (MMR 2017)*, Grenoble, France.

Van der Vaart, A. W. (2000), *Asymptotic Statistics*, Cambridge University Press.

Villani, C. (2009), *Optimal transport*, Springer.