



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. 2020. hal-02425477v2

HAL Id: hal-02425477

<https://hal.science/hal-02425477v2>

Preprint submitted on 6 Aug 2020 (v2), last revised 12 Dec 2020 (v3)

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 for robustness analysis in uncertainty quantification of computer codes

Clement GAUCHY^{1*}

Jerome STENGER¹²

Roman SUEUR¹

Bertrand IOOSS^{12†}

¹EDF R&D, Département PRISME, 6 Quai Watier, 78401, Chatou, France

²Institut de Mathématiques de Toulouse, 31062 Toulouse, France

August 6, 2020

Abstract

Robustness analysis is an emerging field in the domain of uncertainty quantification. It consists of analysing the response of a computer model with uncertain inputs to the perturbation of one or several of its input distributions. Thus, a practical robustness analysis methodology should rely on a coherent definition of a distribution perturbation. This paper addresses this issue by exposing a rigorous way of perturbing densities. The proposed methodology is based on the Fisher distance on manifolds of probability distributions. A numerical method to calculate perturbed densities in practice is presented. This method comes from Lagrangian mechanics and consists of solving an ordinary differential equations system. This perturbation definition is then used to compute quantile-oriented robustness indices. The resulting Perturbed-Law based Indices (PLI) are illustrated on several numerical models. This methodology is also applied to an industrial study (simulation of a loss of coolant accident in a nuclear reactor), where several tens of the model physical parameters are uncertain with limited knowledge concerning their distributions.

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

1 Introduction

During the last decades, two major trends in industrial and research practices have led to a rise in importance of uncertainty quantification (UQ) methodologies (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 consists in accounting for the risks in an increasing number of industrial activities, this implies that those risks should be evaluated from a quantitative point of view. In both situations, the quantification of uncertainties can be conducted by considering as a vector of random variables, named $\mathbf{X} = (X_1, \dots, X_d)$, the uncertain inputs of the computer code represented by a function $G(\cdot)$. The most widespread approach consists of running $G(\cdot)$ with different combinations of inputs in accordance with their range of variation, in order to study the related uncertainty on the output $Y = G(X_1, \dots, X_d)$ or to estimate a specific quantity of interest (QoI). A QoI is a statistical quantity derived from Y , e.g. a performance as the mean of Y or a risk criterion as a high-level quantile of Y .

As an example, the nuclear industry faces major issues as facilities age and regulatory authorities' requirements strengthen (Bucalossi et al., 2010; Mousseau and Williams, 2017). For example, the operators have to study the "Loss of Coolant Accident" (LOCA) resulting in a break on the primary loop of pressurized water nuclear reactors. This scenario can be simulated using system thermal-hydraulic computer codes, which include tens of physical parameters such as condensation or heat transfer coefficients (Mazgaj et al., 2016; Sanchez-

Saez et al., 2018). Yet, the values of these parameters are known with a limited precision (Larget, 2019) as they are calculated by the way of other quantities measured via small-scale physical experiments. Some other variables are only observed during periodic inspections, such as the characteristics of pumps in hydraulic systems.

Various methods coming from the UQ domain are useful in considering these uncertainties in the system safety analysis. First of all, some methods aim at improving the exploration of the input domain \mathcal{X} by using specific designs of experiments, such as the space filling designs (Fang et al., 2006). Such a design allows to cover an input domain as evenly as possible with a fixed number of code runs as well to limit unexplored areas as much as possible. For the estimation of some specific QoI, such as a probability of threshold exceedance by the output or an α -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 to rank influential parameters according to their impact on the QoI (Iooss and Lemaître, 2015; Iooss and Marrel, 2019).

All these approaches are useful to deal with the existence of uncertainties in applied problems. However, industrial (e.g. nuclear facilities) operators have to face the difficulty of justifying their risk assessment methodologies not merely by providing simulation results. Such a justification has to demonstrate that the computed values overestimate the actual risks which most of the time cannot be calculated. This principle of conservatism, which

can be easily implemented when dealing with very simple monotonic physical models, can be hard to be adapted to computer codes simulating complex and non monotonic physical phenomena. It is also not always straightforward to apply this principle when implementing UQ methods based on a set of computer experiments providing a whole range of values for the output quantity Y .

To address this issue, the new UQ branch of robustness analysis has emerged during the recent years in the field of sensitivity analysis. It consists of evaluating the impact of the choice of the inputs' distributions and, more precisely, by analyzing the QoI variations with respect to this choice. A first solution would consider a whole set of input laws and analysing the related output distributions. For global sensitivity analysis, Hart and Gremaud (2019) uses "optimal perturbations" of the probability density functions to analyze the robustness of the variance-based sensitivity indices (called Sobol indices (Sobol, 1993)). Meynaoui et al. (2019) and Chabridon et al. (2018) propose approaches to deal with the so-called second-level uncertainty, i.e. uncertainty on the parameters of the input distributions. Another approach, called optimal uncertainty quantification, avoids specifying the input probability distributions, turning the problem to the definition of constraints on moments (Owhadi et al., 2013; Stenger et al., 2019). This solution is out of scope of the present work which considers that the initial input probability, that has been defined by the user, is of practical importance.

In practical engineering uncertainty quantification studies, input distributions are truncated as it corresponds to physical parameters with known domain of validity. It is therefore natural to assume no uncertainty on the support of the input random variables. In this

paper, we also assume their mutual independence. Keeping in mind that our goal is to directly deal with the input distributions (without considering 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)) with the so called 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 of a moment (e.g. the mean or the variance). Amongst all densities with shifted mean or variance of a δ value, $f_{i\delta}$ is defined as the one minimizing the Kullback-Leibler divergence from f_i . This method has been applied on the computation of a probability of failure (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) as the QoI.

However, this method is not fully satisfactory. Indeed, the minimal Kullback-Leibler divergence can significantly vary between different inputs' distribution of even two different parameters of the same density, so that some densities are more perturbed than others. Moreover, some distributions do not have defined moments. As in Perrin and Defaux (2019), an iso-probabilistic operator can be applied to transform all the input random variables into centered normalized Gaussian ones. It allows to make perturbations comparable when applied in this standard space, but it remains difficult to translate this interpretation in the initial physical space which is the one of interest for the practitioners. Note that another type of robustness analysis has been proposed in quantitative finance by Cont et al. (2010). These authors investigate whether the estimated QoI is sensitive to a small pertur-

bation of the empirical distribution function. For this purpose, they define the robustness of a QoI as its continuity with respect to the Prokhorov distance on the set of integrable random variables.

The goal of this paper is to propose a novel approach for perturbing probability distribution. It relies on density perturbation based on the Fisher distance (Costa et al., 2012) as a measure of dissimilarity between the initial density f_i and the perturbed one $f_{i\delta}$. This distance defines a geometry on spaces of probability measures called information geometry (Nielsen, 2013). The statistical interpretation of the Fisher distance provides an equivalence between perturbation of non-homogeneous quantities and consequently a coherent framework for robustness analysis. To present this approach, we first review the existing density perturbation methods in Section 2. Section 3 is then dedicated to the description of our method and the discussion of our numerical tools. Section 4 illustrates our methodology of density perturbation on the practical robustness index PLI. An analytical application and an industrial case study are presented in Section 5. The last section gives conclusions and some research perspectives.

2 Previous approaches of density perturbation for UQ robustness analysis

The method of Lemaître et al. (2015) has been later called PLI by Sueur et al. (2016), as it is based on the idea of perturbing the inputs' densities. It aims at providing 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 one) is changed. This can be seen as a way to take into account an “error term” one could add to an imperfectly known input distribution.

2.1 Kullback-Leibler divergence minimization

To build a perturbed distribution $f_{i\delta}$ from a distribution f_i , the approach of Lemaître et al. (2015) is non-parametric. It is mainly thought to analyze perturbations on the most common characteristics of input laws which are the mean and variance. To illustrate it in the case of a mean perturbation, we assume the random variable $X_i \sim f_i$ has mean $\mathbb{E}[X_i] = \mu$. By definition, the perturbed density will have a $\mu + \delta$ mean. But this is obviously not sufficient to fully determine the perturbed law and especially to explicitly access the value of $f_{i\delta}$ on the whole domain of X_i . Amongst all densities with a mean equal to $\mu + \delta$, $f_{i\delta}$ is defined as the solution of the minimization problem

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

where \mathcal{P} is the set of all probability measures absolutely continuous with respect to f_i . This approach basically consists of perturbing the chosen parameter while changing the initial model as little as possible. With this definition, “changing” the model is understood as an increase of entropy, the Kullback-Leibler divergence between two densities f and π being

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

This method can be applied on higher order moments (for instance moments of order 2, to define variance perturbation) and, more generally, to constraints that can be expressed

as a function of the perturbed density, as quantiles (Lemaître, 2014). Notice that, in the case of an initial Gaussian distribution, the perturbed distribution remains Gaussian with a mean shift of δ .

In the general case, this method has several drawbacks: First of all, the likelihood ratio between $f_{i\delta}$ and f_i might not have an analytic form, which leads to numerical difficulties. Moreover, this method requires defined moments for the initial density. Finally, the main difficulty concerns the interpretation of the results obtained from this PLI method. Indeed, each uncertain input of the UQ model is perturbed with a range of δ values. To interpret the QoI shift resulting of these perturbations in the standard space, a clear understanding of the physical meaning of each perturbation is necessary. Low interpretability of the perturbed density can appear for some physical parameters, e.g. for uncertainties on the state of the system coming from a variability of the quantity throughout the operating process. In this case, the probability distribution of the uncertain quantity can be regarded in terms of relative frequency of occurrence. But it can be more difficult when it comes to constant physical parameters known with a limited accuracy.

We recall that all input random variables are assumed mutually independent. Nonetheless, the effect of perturbations can be considered only for each variable individually and in absolute terms (as a same δ shift might have completely different impacts for different input densities). This methodology thus yields difficulty to compare the relative impact of perturbations between different inputs.

2.2 Standard space transformation

To interpret the δ shift on the input distribution and especially to allow a comparison between inputs according to the impact on the QoI of a 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 laws are identical, making all perturbations equivalent. Finally, the perturbed densities are obtained by applying the reverse transformation as the one used to transform inputs in the standard space.

In the case of independent inputs, the required distribution transformation is a simple inverse probability transform. Given a random vector \mathbf{X} with cumulative distribution function F , the transform is the random vector $\mathbf{S} = \Phi^{-1}(F(\mathbf{X}))$, where Φ is the cumulative distribution function 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 simply consists of a mean shift without changing the standard deviation. Hence this leads to an analytical expression for the perturbed density $f_{i\delta}$ thanks to the variable change formula (Stirzaker, 2003, p.318):

$$f_{\delta}(x) = e^{\frac{-\delta^2 + 2\delta\Phi^{-1}(F(x))}{2}} f(x) . \quad (3)$$

This simple formula makes the perturbed density and the likelihood ratio easy to compute.

However, similar perturbations in the standard space implies very different ones in the physical space according to the initial distribution. As an example, Figure 1 depicts two Kullback-Leibler divergences (approximated with Simpson's rule (Abramowitz and Stegun,

1974)) between a particular distribution (the Triangular $\mathcal{T}(-1, 0, 1)$ ¹ and the Uniform one $\mathcal{U}[-1, 1]$) and its associated distribution in the standard space. The results show that the Kullback-Leibler divergence behaves very differently in the physical space, depending of the original distribution, even though the same perturbation is applied in the standard space. For example, there is no general rule to estimate the mean of the physical perturbed input for a given mean perturbation in the standard space. Such difficulties are even more significant when considering perturbations on other parameters than the mean. For instance, there is no general equivalence in the physical space between perturbations applied on the mean and on the standard deviation of the same input in the standard one. Hence, it seems generally impossible to convert in a simple way the results given by this method into a relationship between input and output physical quantities, making these results difficult to interpret.

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 an alternative hypothesis “ X_i follows distribution $f_{i\delta}$ ” (Eguchi and Copas, 2006). For this reason, it seems to be an appropriate tool to measure how far a perturbed density is from its initial reference and thus to provide a formal counterpart to the dim idea of “uncertainty on the distribution”. It is especially well

¹the triangular distribution $\mathcal{T}(-1, 0, 1)$ is parametrized by its minimum a , mode b and maximum c

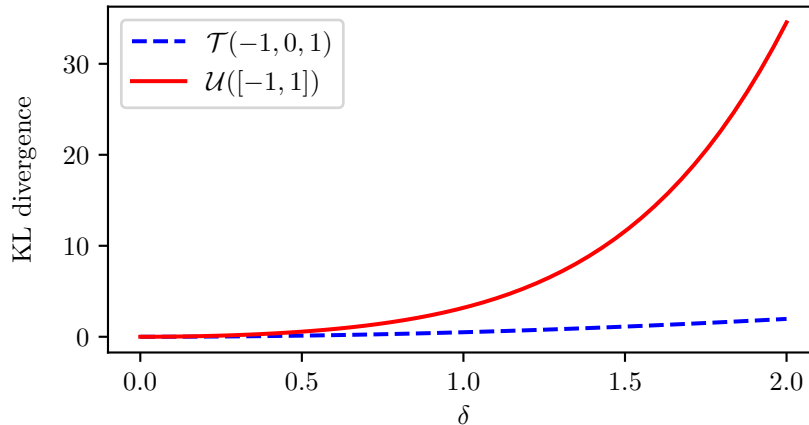


Figure 1: Kullback-Leibler divergence between the initial distribution and the perturbed one for perturbation levels $\delta \in [0, 2]$. Further description on the perturbed distribution can be found in Lemaître (2014) and Lemaître et al. (2015).

suited to compare Gaussian distributions, which requires, in a robustness analysis context, an additional transformation to embed inputs in a standard space as these are physical quantities with potentially non Gaussian distributions. This additional operation, which also provides an equivalence between non-homogeneous input variables, makes it impossible to interpret in terms of physical inputs the perturbations of the related standard ones.

3.1 Fisher distance

To allow intuitive understandings of the consequence of these perturbations on the output distribution, it is necessary to base our perturbation method on a metric which allows at the same time to compare perturbations on different parameters of the same distribution and on different inputs of the UQ model. In particular it should not depend on the representation of the input distribution, which means being independent of the parametrization. The

Fisher distance has all these advantages. 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 parametric density family $\mathcal{S} = \{f_\theta, \theta \in \Theta \subset \mathbb{R}^r\}$. We recall that every input variables represent physical parameters with known domain of validity, therefore for all θ in Θ , the support of f_θ is assumed fixed. The metric associated to the coordinate function θ , called the Fisher (or Fisher - Rao) metric, is defined as:

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

where $I(\theta)$ is the Fisher information matrix evaluated in θ for this statistical model. The Fisher information, well known for instance in 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 θ 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_{θ_1} and f_{θ_2} in the manifold \mathcal{S} , a path from f_{θ_1} to f_{θ_2} is a piecewise smooth map $q : [0, 1] \rightarrow \Theta$ satisfying $q(0) = \theta_1$ and $q(1) = \theta_2$. Its length $l(q)$ 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 . Alike, the energy $E(q)$ 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 distance between f_{θ_1} and f_{θ_2} , called the Fisher distance, is defined as the minimal length over the set of paths from f_{θ_1} to f_{θ_2} , denoted by $\mathcal{P}(f_{\theta_1}, f_{\theta_2})$:

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

The path γ 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 θ , 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_θ , the maximum likelihood estimator $\hat{\theta}_n$ of θ converges weakly to a normal law with mean θ and covariance $\frac{I(\theta)^{-1}}{n}$. The density of $\hat{\theta}_n$ denoted by $p(\hat{\theta}_n, \theta)$ writes

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

When n is large, this probability density is proportional to $(\hat{\theta}_n - \theta)^T I(\theta)(\hat{\theta}_n - \theta)$ which is the local inner product defined in equation (4). Therefore, the Fisher distance between two distributions with parameters θ and θ' can be constructed as a measure of the risk of confusion between them. In other words, the Fisher distance between two distributions f_θ and $f_{\theta'}$ represents the separability of the two distributions by a finite sample of independent observations sampled from the f_θ distribution.

We illustrate the Fisher distance on 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):

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

We can apply the change of coordinate $\phi(\mu, \sigma) \rightarrow (\frac{\mu}{\sqrt{2}}, \sigma)$, so that 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, the geodesics are the vertical lines and the half circle centered on the line $\sigma = 0$.

Further details on the interpretation 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 Gaussian distributions C and D are more difficult to be distinguished than the distributions A and B although in both cases the KL divergence 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 respectively between points A and B , and points C and D . We observe that the Fisher distance between A and B is greater than the distance between C and D . This illustrates how information geometry provides a proper framework to measure statistical dissimilarities in a space of probability measures.

The Fisher distance provides a satisfactory grounding to our notion of density perturbation. We define a perturbation of a density f to be of magnitude δ if the Fisher distance

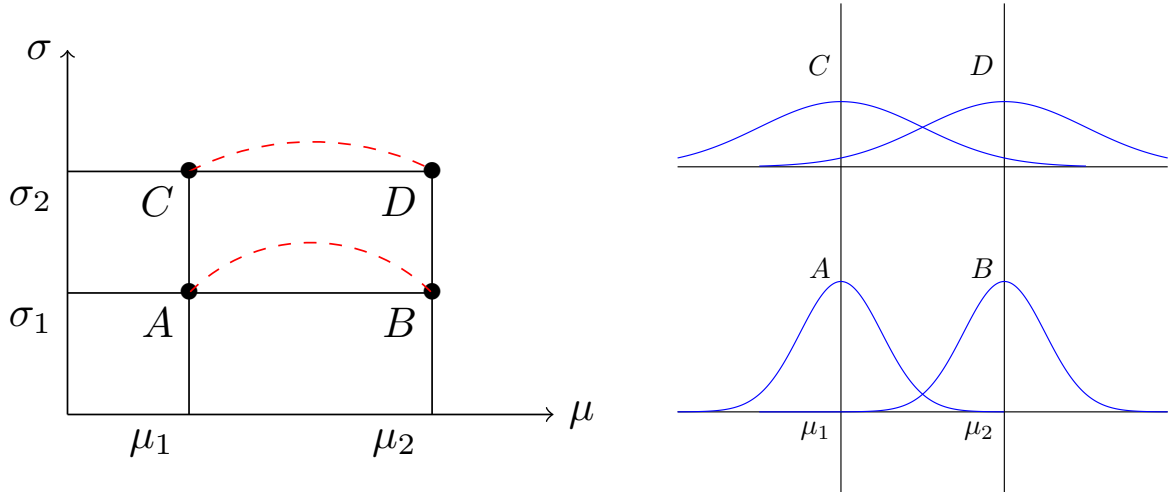


Figure 2: Representation of four Gaussian distributions in the parameter space on the left, and their respective distributions on the right. Although $KL(A||B) = KL(C||D)$, it is easier to distinguish A from B than C from D . The dashed curved lines are two geodesics in $(\frac{\mu}{\sqrt{2}}, \sigma)$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

between f and the perturbed density f_δ is equal to δ . The set of all perturbations of f at level δ is then the Fisher sphere of radius δ centered in 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 a non finite set of probability densities. The next section is dedicated to the development of a numerical method to compute the Fisher spheres of radius δ centered in f .

3.2 Computing Fisher spheres

As detailed in Section 3.1, geodesics are defined as the solution of a minimization problem. More 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 $L(t, q, \dot{q}) = \frac{1}{2} \langle \dot{q}(t), \dot{q}(t) \rangle_{q(t)}$ and L is called 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 derivation 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

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

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 Equation (ODE) called Hamilton's equations:

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

The objective is to determine any geodesics q satisfying $q(0) = \theta_0$ and $d_F(f, q(1)) = \delta$, it corresponds to computing the Fisher sphere centered in f_{θ_0} with radius δ . The only

degree of freedom left to fully solve the ODE system (14) is the initial velocity $p(0)$. Notice that the Hamiltonian is equal to the kinetic energy as $p = I(q)\dot{q}$. As the Hamiltonian is constant on a geodesic, we have for all t :

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

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

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

so that $\delta = \sqrt{2k}$. Therefore, Eq. (13) rewrites:

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

Taking equation (17) at initial state $t = 0$, we can determine all the initial velocity such that $d_F(q(0), q(1)) = \delta$. Those velocities are needed to solve the ODE system (14) and compute the geodesics.

Generally, computing the geodesic between two given distributions is a challenging problem. Methods relying on shooting algorithms have been developed in that matters. Our framework overcomes this problem as we compute the entire Fisher sphere. In the next section, we focus on numerical methods for computing geodesics by solving the systems of ODE (14). These methods are illustrated by computing Fisher spheres in the Gaussian manifold $\mathcal{S} = \{\mathcal{N}(\mu, \sigma^2), (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*\}$.

3.3 Numerical results

The Hamilton equations (14) are solved with numerical approximation methods. Figure 3 illustrates our numerical resolution method in the Gaussian case, that is when

$\mathcal{S} = \{\mathcal{N}(\mu, \sigma^2), (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+^*\}$. In order to solve (14), we compare two different numerical methods: namely, the explicit Euler algorithm and the Adams-Moulton algorithm. We recall that in the Gaussian case we dispose of an exact analytical expression of the Fisher sphere detailed in Costa et al. (2012). The Fisher sphere is centered in $\mathcal{N}(0, 1)$ with radius $\delta = 1$. Notice that there is no observable difference between the two methods in Figure 3. Hence, a better way to estimate the numerical error is required. We recall that the Hamiltonian value is conserved along the geodesics. Therefore, 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 $t \in [0, 1]$. Δ represents the relative variation of the Hamiltonian along the path q computed with our numerical methods.

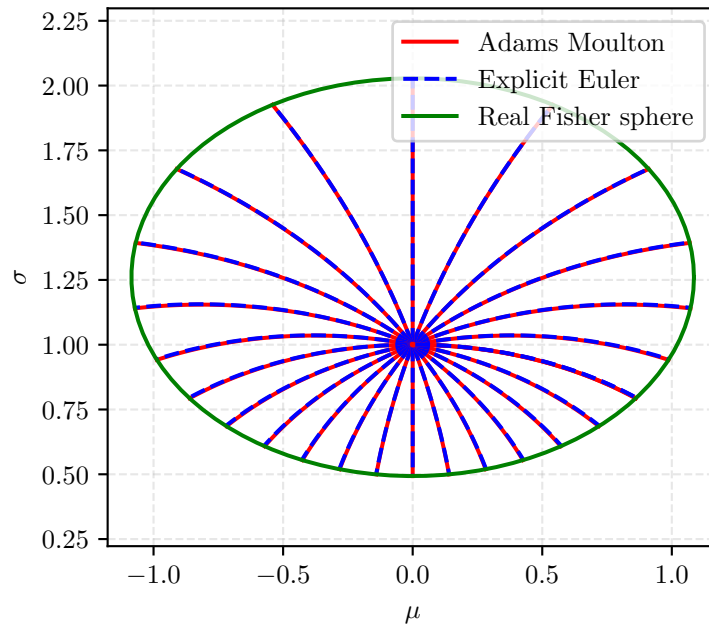


Figure 3: Geodesics in the Gaussian information geometry computed with Euler explicit and Adams Moulton methods. The radius δ is equal to 1.

Figure 4 displays the value of $\Delta(t)$ for $t \in [0, 1]$ for one arbitrary geodesic shown in Figure 3. 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 preferred. Nevertheless, some instabilities have been observed in practice mainly due to the truncation of the distribution support which impair the Hamiltonian consistency. Symplectic method (Amari and Nagaoka, 2000; Leimkuhler and Reich, 2005) and more particularly symplectic Euler algorithm could help to assess this problem by forcing the Hamiltonian constant. This will be the subject of a future work. Moreover, the truncation can lead to other numerical errors when the radius δ is too large. Indeed, the normalization factor of some truncated distribution can become smaller than the computer machine precision.

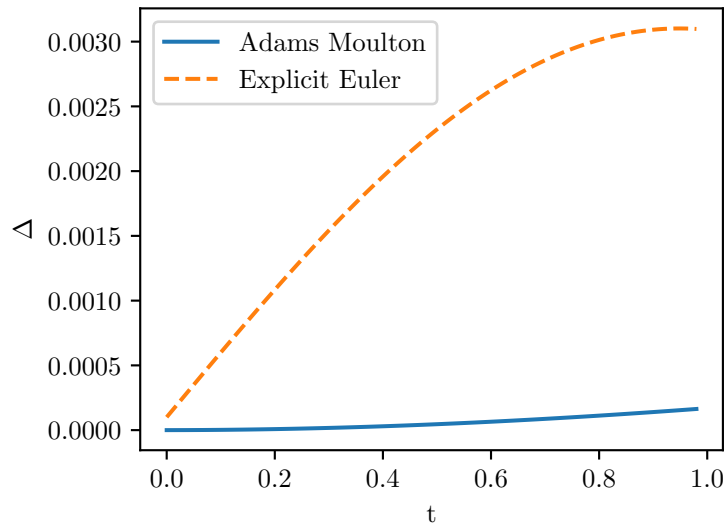


Figure 4: Relative variation of the Hamiltonian Δ along a geodesic for two different numerical scheme.

4 Application to Perturbed-Law based Indices

The UQ robustness analysis explained in Section 1 and Section 2 aims at quantifying the impact of a lack of knowledge on an input distribution in UQ of model outputs. In Section 3, a coherent formal definition of density perturbation has been proposed. We now illustrate the interest of this solution for the definition of a practical robustness analysis methodology. Analyzing the effect of perturbing an input density first requires defining an index which summarizes this effect on the QoI.

4.1 Definition of the Perturbed-Law based Index

A PLI aims to measure the impact of the modification of an input density on some events affecting the 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 α , which is often used in practical applications as a risk measure (Mousseau and Williams, 2017; Delage et al., 2018; Larget, 2019).

Given the random vector $\mathbf{X} = (X_1, \dots, X_d) \in \mathcal{X}$ of our d independent uncertain input variables, $G(\cdot)$ our numerical model and $Y = G(\mathbf{X}) \in \mathbb{R}$ the model output, the quantile of order α of Y is:

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

where F_Y is the cumulative distribution function of the random variable Y . In order to compute the i -th PLI, we change the density f_i of X_i into a density $f_{i\delta}$, where $\delta \in \mathbb{R}^+$

represents the level of the perturbation. The perturbed quantile then writes:

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

where $F_{Y,i\delta}$ is the cumulative distribution function corresponding to the input variable X_i sampled from $f_{i\delta}$. The PLI is then simply defined as the relative change in the output quantile generated by the perturbation :

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

This definition slightly differs from the one proposed in previous studies (Lemaître et al., 2015; Sueur et al., 2017). Indeed, after several applications of the PLI, it has been found more convenient to compute directly the relative variation of the quantile when submitted to a density perturbation. The interpretation is straightforward.

In a lot of applications, for instance in nuclear safety exercises, the computer models are costly in terms of CPU time and memory. Only a limited number of N code runs is then available for the estimation of all the PLIs. We then 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 based on the empirical quantile estimator denoted $\hat{q}_N^\alpha = \inf\{t \in \mathbb{R}, \hat{F}_Y^N(t) \leq \alpha\}$ where $\hat{F}_Y^N(t) = \frac{1}{N} \sum_{n=1}^N \mathbf{1}_{(y^{(n)} \leq t)}$ is the empirical estimator of the cumulative density function of Y . In order to estimate the perturbed quantile $\hat{q}_{N,i\delta}^\alpha$ from the same sample \mathcal{X}_N , we use the so-called reverse importance sampling mechanism from Hesterberg (1996) (see the online supplementary material) 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

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

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

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

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

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

Among all perturbed distributions at level δ , we investigate the one that deviates the quantile the most from its original value. Thus, these two quantities $S_{i\delta}^+$ and $S_{i\delta}^-$ measure the robustness of the numerical code taking into account uncertainties tainting the input distribution.

4.2 Theoretical properties of the estimator

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

estimator $(\widehat{q}_N^\alpha, \widehat{q}_{N,i\delta}^\alpha)$. Detailed proof of the following results are reported in the Appendix A.

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

$$\sigma_i^2 = \frac{\alpha(1-\alpha)}{F_Y'(q^\alpha)^2}, \quad (25)$$

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

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

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

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

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

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

$$\sqrt{N}(\widehat{S}_{N,i\delta} - S_{i\delta}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, d_s^T \Sigma d_s) \text{ with } d_s = \begin{pmatrix} -q^\alpha/q_{i\delta}^{\alpha 2} \\ 1/q^\alpha \end{pmatrix}. \quad (26)$$

Notice that the asymptotic variance relies on the α initial quantile and perturbed quantile, which are precisely what we want to estimate. Hence, Theorem 2 cannot be used for

building asymptotic confidence intervals. However, the convergence properties are important for the method credibility and acceptance. In practice, the estimation error can be measured using bootstrapping (Efron, 1979).

4.3 Practical implementation of the methodology

As already discussed, in practical applications, the computer model is often costly and cannot be reevaluated. The main limitation of the previously exposed estimator $\widehat{S}_{N,i\delta}$ arises from the available sample size which is finite. Therefore, at a certain level of perturbation, there might not be enough sample points to correctly compute the perturbed quantile (and its confidence interval). One of the key issue of the methodology is to determine how far the input distribution should be perturbed. We propose to adapt the empirical criterion from Iooss et al. (2020) in order to establish a maximal perturbed level δ_{max} . The number of points in the output sample \mathcal{Y}_N , smaller or larger than the δ -perturbed quantile has to be sufficient. A value of $N_{\mathcal{Y}} = 10$ has been chosen (from several numerical tests) as the smallest size for computing the PLI-quantile. As soon as a distribution on the Fisher sphere exceeds this criteria, the corresponding radius is taken as δ_{max} .

The estimation of the quantity of interest $S_{i\delta}^+$ and $S_{i\delta}^-$ is summarized as follows:

- Choose a level of perturbation δ , an input number $i \in \llbracket 1; d \rrbracket$ and a sample of K points on the Fisher sphere of radius δ centered in f_i using the numerical method of Section 4.2.
- For each $\{f_{i\delta}^{(k)}\}_{1 \leq k \leq K}$ sampled on the Fisher sphere, estimate $q_{i\delta}^{\alpha, (k)}$ using the reverse importance sampling technique based on the sample \mathcal{X}_N . Verify that the number of

point in the output sample below or above the perturbed quantile $q_{i\delta}^{\alpha,(k)}$ satisfies the stopping criteria N_y . Then, compute the PLI estimator $\widehat{S}_{N,i\delta}^{(k)}$.

- The estimators $\widehat{S}_{N,i\delta}^+$ and $\widehat{S}_{N,i\delta}^-$ of the quantity of interest $S_{i\delta}^+$ and $S_{i\delta}^-$ are taken as the maximal and minimal value of the PLI sampled on the Fisher sphere $\{\widehat{S}_{N,i\delta}^{(k)}\}$.

We emphasize that this approach only restricts to expensive computer models. Indeed, the bootstrap variance of the estimated quantile with reverse importance sampling tends to become very large as illustrated in Iooss et al. (2020). This is due to the likelihood ratio that punctually explodes. 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 δ_{max} .

The code for the new version of the PLI, called OF-PLI (for Optimal Fisher-based PLI) in the following, is available at <https://github.com/JeromeStenger/PLI-Technometrics>. In future works, the OF-PLI confidence intervals (computed *via* bootstrap) will provide valuable additional information such as confidence intervals. They are not pictured in the following application as it requires at this stage further investigations. The code for computing the old version of the PLI, called E-PLI (for Entropy-based PLI) in the following, is available in the sensitivity package of the R software.

5 Perturbed-Law based Indices in engineering studies

The PLI, as defined in the previous sections, allow to assess to what extent the output quantile can be impacted by an error of magnitude δ in the characterization of an input

distribution. In the next subsection, we compare in a toy example the newly introduced methodology (OF-PLI) to the previous one (E-PLI). Moreover, as the PLI are based on a change in the input distribution, it differs from global sensitivity measures (Iooss and Lemaître, 2015) which evaluate the effect of input variability for a fixed probabilistic model. To study the potential coherence and divergence between the two approaches, we compare Sobol’ indices and OF-PLI results in Section 5.2 on an analytical model. In the third subsection, we illustrate the use of the OF-PLI as a support in nuclear safety analysis of a pressurized water nuclear reactor.

5.1 A toy example: the Ishigami function

The Ishigami function (Ishigami and Homma, 1990) is 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 4.3 to estimate OF-PLI and compare our results to the E-PLI. The Ishigami function, which takes three input random variables (X_1, X_2, X_3) normally distributed $\mathcal{N}(0, 1)$, is defined with the following analytical formula:

$$G(x_1, x_2, x_3) = \sin(x_1) + 7 \sin(x_2)^2 + 0.1x_3^4 \sin(x_1) . \quad (27)$$

We intend to evaluate the impact of a perturbed input distribution to 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 4.3. We rather draw new samples of size $N = 2000$ directly from the perturbed input distributions in order to compute the output perturbed quantile. We chose a number of $K = 100$ trajectories over

each Fisher sphere for computing the minimum and maximum of the OF-PLI. The OF-PLI are computed for perturbation levels δ varying in $[0, 0.9]$. We emphasize that the choice $\delta_{max} = 0.9$ is arbitrary. Indeed, there is here no actual limit for the maximal perturbation level as the OF-PLI are computed by resampling from the perturbed distribution. 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 5. It appears that the third input has most impact in particular for shifting the quantile to the right. On the other hand, the second input has more impact for perturbing the quantile to the left. Our results coincide to the well known behavior of the Ishigami function in terms both of non-linearity of the model and primary influence of the third input.

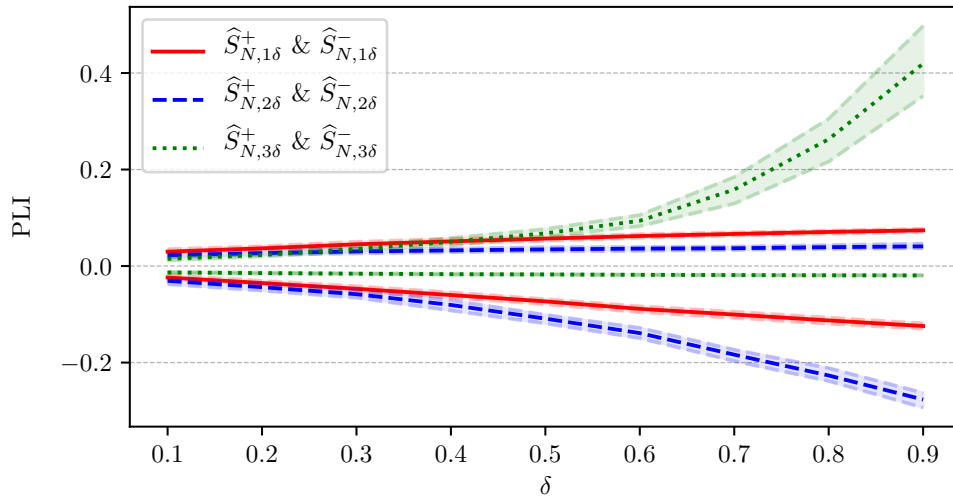


Figure 5: Minimum and maximum of the OF-PLI over the Fisher sphere over $K = 100$ trajectories for δ varying in $[0, 0.9]$, and their 95%-confidence intervals.

Because, the maximum and minimum of the OF-PLI are taken over the Fisher sphere, we depict in Figure 6 the distribution of the OF-PLI over the Fisher sphere with radius

$\delta = 0.9$ of the third input. One can see that in this situation the maximum and minimum are found for respectively high variance and low variance with no change of the mean.

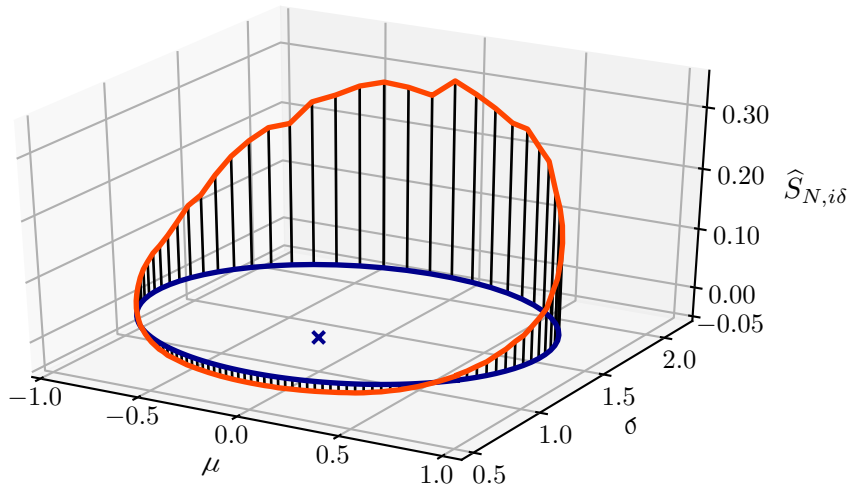


Figure 6: Value of the OF-PLI $\hat{S}_{N,i\delta}$ (red line) for the third input of the Ishigami model ($N = 100$, $i = 3$) over a Fisher sphere of radius $\delta = 0.9$ (blue line).

These results are to be compared to the E-PLI (see Section 2). We recall that the inputs are all normally distributed so that there is no need to apply the inversion distribution function. Therefore, perturbing the mean (respectively the variance) of the input variable is equivalent to drawing straight horizontal trajectories (respectively vertical) in the parameters space (see Fig. 3). Results are depicted in Figure 7, 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 compare the third input between the two methodologies, we detected in Figure 6 that the maximal OF-PLI was reached for high variance and no mean perturbation which is coherent with the results in Figure 7. However, one misses the true impact a perturbed density can induce in any

situation where the maximal and minimal OF-PLI are not obtained in these two axes, such as, for instance, the first variable. Hence, the E-PLI, restricted to two directions of the Fisher sphere, have limited interpretation.

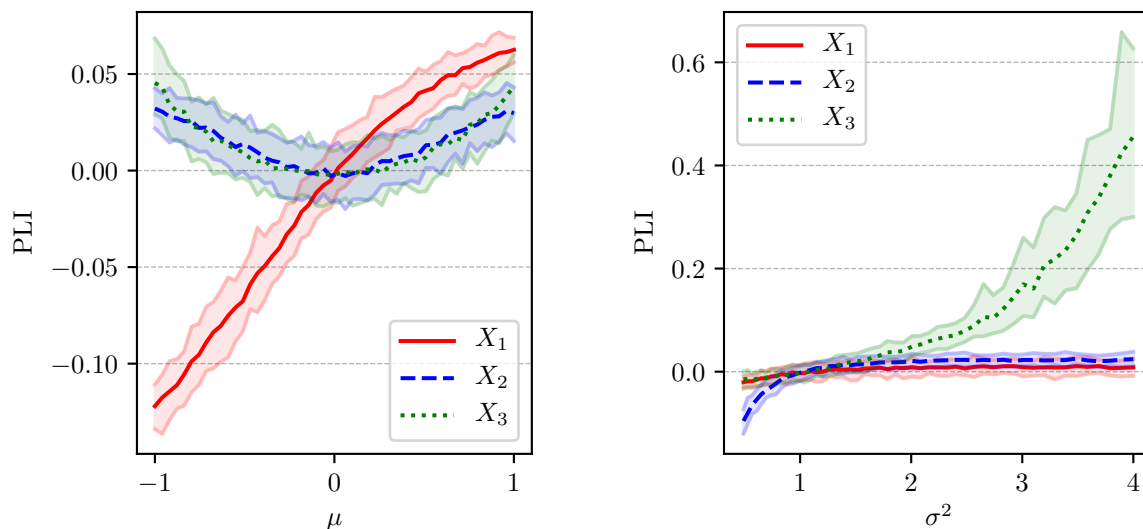


Figure 7: Computation of the E-PLI. Left: perturbation of the mean of the Gaussian distribution. Right: perturbation of the variance of the Gaussian distribution.

5.2 An analytical model: flood risk of an industrial site

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 nearby a river. To this purpose, we use a model implementing a simplified version of the 1D hydro-dynamical equations of Saint Venant. 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 uncertain:

$$H = \left(\frac{Q}{300K_s \sqrt{2 \cdot 10^{-4} (Z_m - Z_v)}} \right)^{0.6}. \quad (28)$$

The inputs are modeled as random variables with associated truncated distributions given in Table 1 (Iooss and Lemaître, 2015).

Table 1: 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 analysis, 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 assumption of independent inputs (Sobol’, 2001; Saltelli and Tarantola, 2002; Prieur and Tarantola, 2017). They will be then compared to the results of our robustness analysis framework in order to illustrate their difference. However, 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. (28)), the asymptotically efficient pick-freeze estimator (Prieur and Tarantola, 2017) is used with elementary Monte Carlo matrix of size 10^6 . It 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 2, 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 between them in order the output exceeds the threshold).

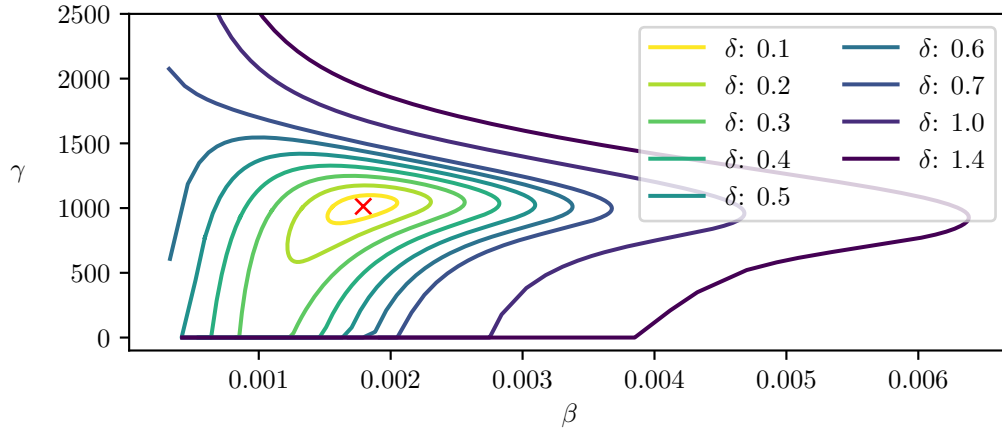
Table 2: Sobol' indices estimates of the flood model inputs.

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

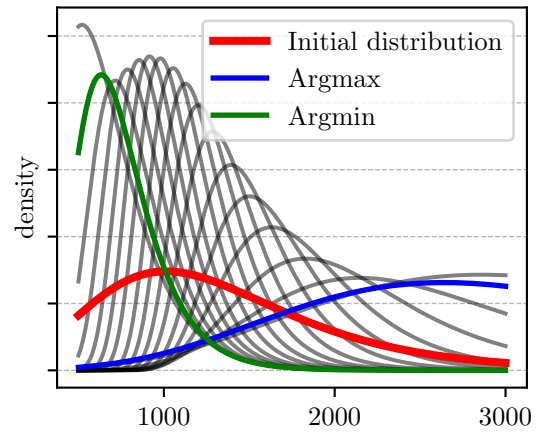
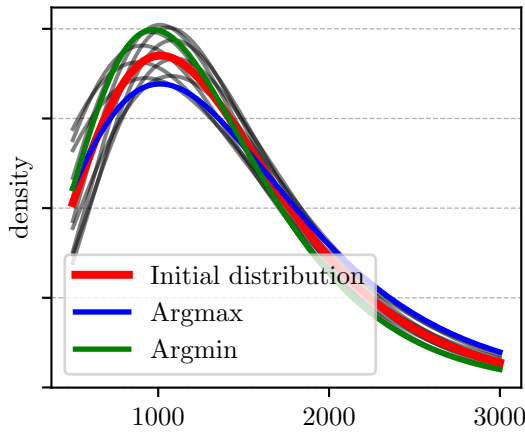
We compute the OF-PLI (w.r.t. a quantile of order $\alpha = 0.95$) for the flood model inputs with the methodology of Section 4.3 for increasing Fisher spheres radius $\delta \in [0, 1.4]$ with step 0.1. The spheres are respectively centered on the distributions of Table 1. On each of these spheres, 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 $\delta_{max} = 1.4$ derives from the stopping criteria in Section 4.3. More precisely, the criterion is reached for the first input Q at perturbation level $\delta > 1.4$, meaning there are lower than $N_y = 10$ sample points above the maximal perturbed quantile. The Figure 8 depicts how the Fisher sphere centered in the variable Q deforms and how the perturbed densities spread around the initial distribution. Figures 8b and 8c indicates that the maximal value

of the OF-PLI is obtained by putting weight to the right hand side of the distribution queue (the distributions minimizing and maximizing the OF-PLI are colored green and blue). This behavior was here predictable as the the height river is a growing function of the river flow (see Eq. (28)). However, this analysis can give substantial information in an real world engineering study. At last, one can observe (Fig. 8a) that the Fisher sphere flatten to the boundary of the parameters' domain. This characteristic is peculiar to each probability distribution, for instance it never not happen for the non-truncated normal distribution.

The results of the OF-PLI, displayed in Figure 9, confirm those of the target Sobol' indices (see Table 2): 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 . Moreover, perturbations of Q and K_s seem to have comparable effects on the 95%-quantile of H although they have significantly different contributions to the output variance. On the other hand, compared to target Sobol' indices, OF-PLI provide more informative results with their evolution as a function of δ . This clearly shows how a lack of knowledge on an input uncertainty can have a low or high impact on the value of a risk measure. In conclusion, this example confirms the interest of the OF-PLI as it conveys complementary information compared to existing sensitivity indices. Notice that the flat parts visible on some curves are due to approximation errors attributed to the low number of sample points N and high quantile level (0.95).



(a) Deformation of the Fisher sphere for increasing radius δ .



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

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

Figure 8: Analysis of the Fisher metric based perturbation of the truncated Gumbel distribution of the variable Q (see Table 1).

5.3 A nuclear safety case

This industrial application concerns the study of the peak cladding temperature (PCT) of fuel rods in case of loss of coolant accident caused by an intermediate-size break in the primary loop (IB-LOCA) in a nuclear pressurized water reactor. According to operation rules,

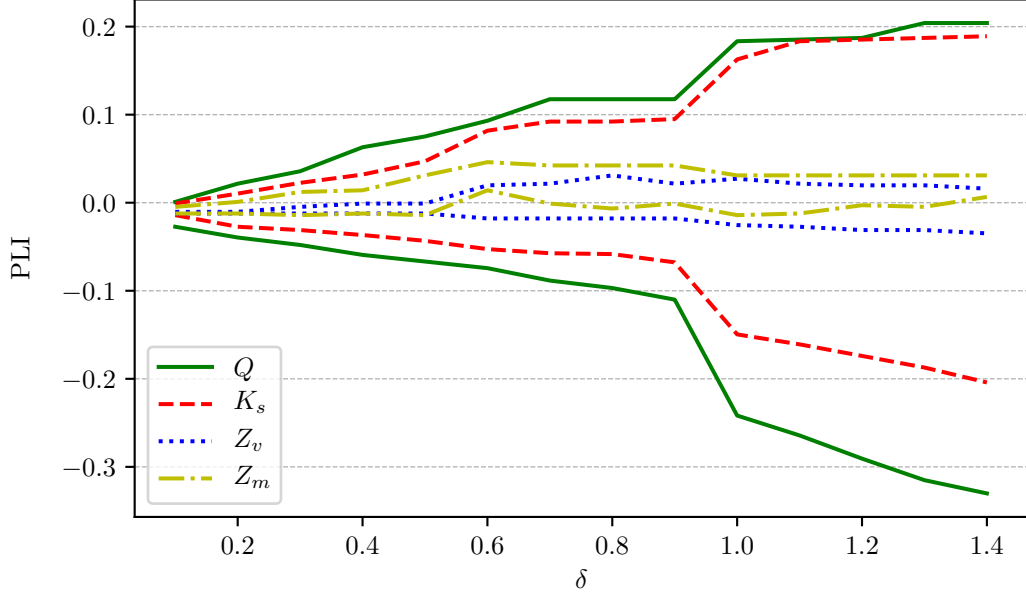


Figure 9: Maximum and minimum estimated value of the OF-PLI $\hat{S}_{N,i\delta}^+$ and $\hat{S}_{N,i\delta}^-$ for the different variables of the flood model.

this temperature must remain below a threshold to prevent any deterioration of the reactor state. 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, from expert knowledge or recovered by solving inverse problems on an experimental database (Baccou et al., 2019).

The input uncertainties are propagated inside this model and the UQ objective consists of estimating a high-order quantile of the PCT (model output). This α -quantile is

interpreted as a pessimistic estimate of the PCT. Like any scientific approach, this methodology is based on hypotheses, which regulatory authorities ask to evaluate the impact on exhibited results. Indeed, nuclear power operators are required to conduct studies in such a way to ensure that actual risks are overestimated. By this “conservatism principle” they are bound 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 values for each of the input variables. This way, the resulting computation is supposed to simulate a worst case scenario for the examined risk. It is, however, not straightforward to implement such a principle when the numerical code is complex with interactions between inputs and non-monotonic effects of inputs. It is even harder to extend this rationale to a UQ framework aiming to represent all potential scenarios with related occurrence plausibility. Recent works (Larget, 2019) have shown that the E-PLI can be useful to support a discussion on the choice of the input distributions.

In our case, we study a reduced scale mock-up of a pressurized water reactor with 7 uncertain inputs given in Table 3 (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 inputs’ 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 10 presents the maximum and minimum values of our two estimators $\widehat{S}_{N,i\delta}^+$

Table 3: 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 $[0, 12.8]$

and $\widehat{S}_{N,i\delta}^-$. We compute Fisher spheres with radius δ sampled uniformly in $[0.1, 0.5]$, all respectively centered on the initial input distributions. On every sphere, $K = 100$ perturbed densities are sampled. The OF-PLIs are finally estimated on a 1000-sized dataset. The stopping criterion of 4.3 gives a maximal admissible OF-PLI of 4%, this value is determined from the maximal admissible quantile such that there is $N_y = 10$ sample points above it. Actually, one can see that $\widehat{S}_{N,7\delta}^+$ is close to this maximal admissible value.

Studies previously conducted on the same application (Delage et al., 2018) lead to similar results concerning the most influential inputs on the quantile of the PCT: strong impact of variables 3 and 4 and weak influence of variables 1, 2 and 5. In comparison with these studies based on the standard space transformation, our information geometry

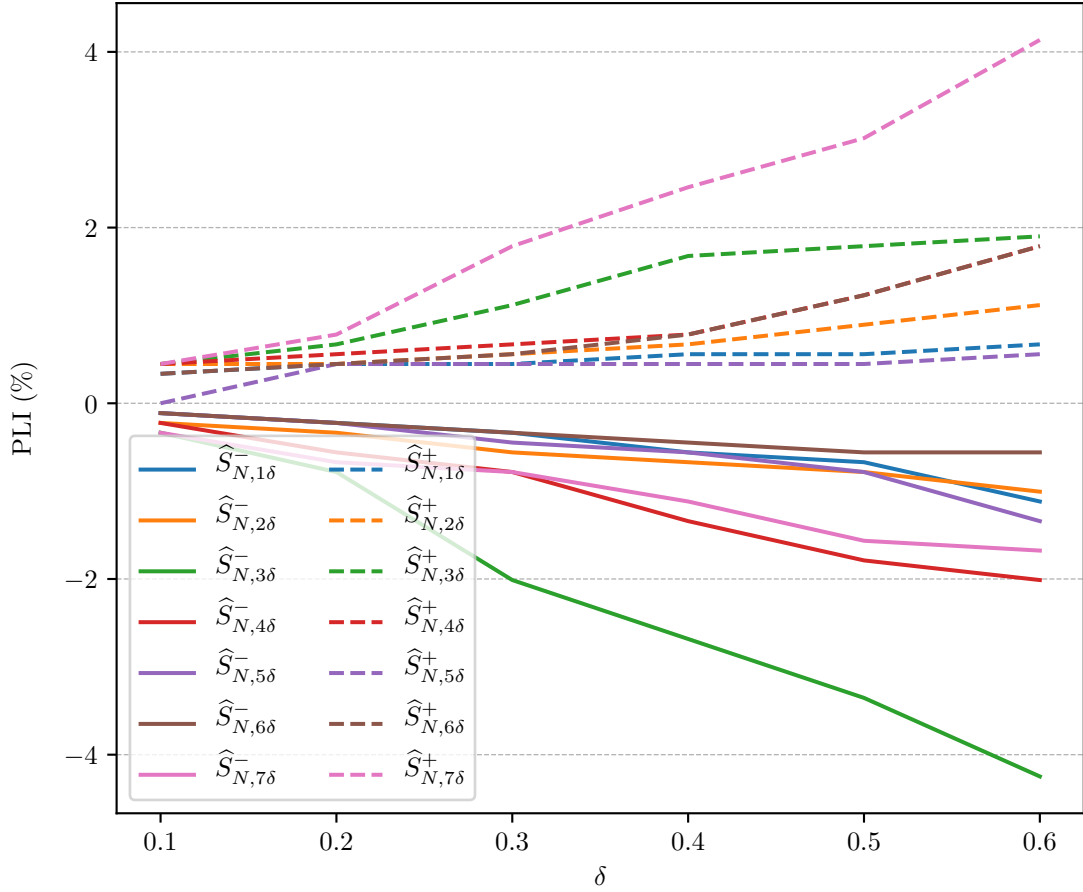


Figure 10: Bootstrap mean of the maximum and minimum of the OF-PLI $S_{i\delta}$ for the CATHARE2 code. The confidence interval are not shown for the sake of clarity.

perturbation methodology leads to a reduced evaluated influence of variable 7. In fact, as it is the only Gaussian distribution, the reverse transformation from the standard space to the physical one operates differently for this input than for the others. Finally, 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 quantile of the PCT. This behavior, which was not observed with the standard space transformation, is probably due to the fact that the standard space approach allows

perturbing only one of the probability distribution parameters (for example the expected value). Contrarily, our estimator corresponds to the maximal quantile deviation over a whole set of equivalent perturbations. This shows two main advantages of our newly developed methodology: it prevents the interpretation bias induced by the standard space transformation and it allows for an exhaustive exploration of density perturbations for a given δ .

6 Conclusion

Based on the Fisher distance, we have defined an original methodology to perturb input probability distributions in the peculiar case of mutual independent input random variables. The Fisher information is an intrinsic characteristic of probability measure and in particular does not depend on a specific chosen parametric representation. This fundamental property makes it the proper mathematical tool to compare perturbations on different uncertain physical inputs of a computer model, but also on different parameters of the same input distribution. It is even possible to get rid of all references to a parametric sub-domain of the set of probability measures on \mathcal{X} , as a non-parametric extension of the Fisher distance is proposed by Holbrook et al. (2017). However, this last perspective is limited by practical issues as it is supposed to rely on a finite dimension 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 base. This approximation will then be poor for probabilities which are very different from those of the chosen base. This fact shows that in practice it is not easy to eliminate

the reference to a particular parametric model, even in a non-parametric framework.

Nevertheless, based on the PLI, our method provides useful information on the most influential uncertainties regarding the distributions of input variables, or the 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 poorly reliable input laws identification or when an improved level of robustness is demanded about the choice of input distributions. In the target application (nuclear licensing), our aim is not only to exhibit safety margin values for the simulated accidents but also to prove the methodology as a whole does not induce any risk of underestimating these values. Hence we do not only look for a worst case assessment method, but for a more global understanding of how a potential error on an input’s distribution affects the output. In that perspective, a practical option to increase the conservatism of UQ studies is to replace one or several input distributions by penalized deterministic values or by a penalized versions of the distributions themselves. This nevertheless implies to justify the choice of the variables for which this penalization is done (see, e.g., Larget and Gautier (2020)).

Further investigations are still to be completed as this method increases the numerical complexity and the computational time compared to the previous method of Lemaître et al. (2015). Indeed, several Monte Carlo loops are needed to compute the maximal and minimal PLI over Fisher spheres. There is ongoing work about the improvement of the estimation of the maximum and the minimum of the PLI on a Fisher sphere. There is known numerical

issue with the reverse important sampling strategy as the likelihood ratio tends to explode as well as the confidence intervals. Moreover the method consists in sampling trajectories over the Fisher sphere, but one could benefit of a more advanced strategy by optimizing directly the PLI over the sphere, *via* gradient descent along this manifold for instance. The crucial problem of probabilistic dependencies between inputs should also be explored to extend our framework to the non independent-input case, works in robustness analysis dealing with dependent input can be found for instance in Pesenti et al. (2019). Moreover, using a distance in a complex space such as the space of probability density functions instead of a moment perturbation makes our methodology harder to interpret from a physicist’s perspective. Thus, it is crucial to clearly define the statistical interpretation of the Fisher distance, i.e. the link with the statistical tests theory. Last but not least, the numerical difficulties illustrated in Section 3.3 prevents us from having a complete degree of freedom on the δ value.

A Proof of Theorem 1

We study the consistency and asymptotic normality of specific M and Z -estimators in order to establish the proof. We suppose this theory is known so that the details can be kept to the bare minimum. Further readings 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

$$\begin{aligned}
\eta &= \frac{\alpha}{1-\alpha}, \\
m_\theta(x) &= -(G(x) - \theta)\mathbf{1}_{(G(x)\leq\theta)} + \eta(G(x) - \theta)\mathbf{1}_{(G(x)>\theta)}, \\
M_N(\theta_1, \theta_2) &= \frac{1}{N} \sum_{n=1}^N m_{\theta_1}(\mathbf{X}^{(n)}) + \frac{f_{i\delta}(X_i^{(n)})}{f_i(X_i^{(n)})} m_{\theta_2}(\mathbf{X}^{(n)}), \\
\hat{\theta}_N &= \arg \max M_N(\theta_1, \theta_2).
\end{aligned} \tag{29}$$

$\hat{\theta}_N$ is defined such that its two components correspond respectively to the estimators \hat{q}_N^α 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 θ 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}_θ is also differentiable in $\theta_0 = \arg \max_{\theta \in \Theta} \mathbb{E}[\bar{m}_\theta(X)]$ with gradient:

$$\nabla_{\theta_0} \bar{m}_{\theta_0}(x) = ((1 + \eta)\mathbf{1}_{(G(\mathbf{x})\leq\theta_1)} - \eta, \frac{f_{i\delta}(x_i)}{f_i(x_i)}((1 + \eta)\mathbf{1}_{(G(\mathbf{x})\leq\theta_2)} - \eta))^T. \tag{30}$$

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

$$V_{\theta_0} = \begin{pmatrix} (1 + \eta)F'_Y(q^\alpha) & 0 \\ 0 & (1 + \eta)F'_{Y,i\delta}(q_{i\delta}^\alpha) \end{pmatrix}, \tag{31}$$

which is symmetric definite non negative 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. It proves the asymptotic normality of the estimator $(\hat{q}^\alpha, \hat{q}_{i\delta}^\alpha)^T$.

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, 44564463.
- 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.
- 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 uncertainty 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.
- 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.
- 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.
- 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 International 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.