



**HAL**  
open science

## Emulators for stochastic simulation codes

Vincent Moutoussamy, Simon Nanty, Benoît Pauwels

► **To cite this version:**

Vincent Moutoussamy, Simon Nanty, Benoît Pauwels. Emulators for stochastic simulation codes. ESAIM: Proceedings and Surveys, 2015, 48, pp.116 - 155. 10.1051/proc/201448005 . hal-01011770

**HAL Id: hal-01011770**

**<https://hal.science/hal-01011770v1>**

Submitted on 24 Jun 2014

**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.

## EMULATORS FOR STOCHASTIC SIMULATION CODES

VINCENT MOUTOUSSAMY<sup>1,2</sup>, SIMON NANTY<sup>3,4</sup> AND BENOÎT PAUWELS<sup>2,5</sup>

**Abstract.** Numerical simulation codes are very common tools to study complex phenomena, but they are often time-consuming and considered as *black boxes*. For some statistical studies (*e.g.* asset management, sensitivity analysis) or optimization problems (*e.g.* tuning of a molecular model), a high number of runs of such codes is needed. Therefore it is more convenient to build a fast-running approximation – or metamodel – of this code based on a design of experiments. The topic of this paper is the definition of metamodels for *stochastic* codes. Contrary to deterministic codes, stochastic codes can give different results when they are called several times with the same input. In this paper, two approaches are proposed to build a metamodel of the probability density function of a stochastic code output. The first one is based on *kernel regression* and the second one consists in decomposing the output density on a basis of well-chosen probability density functions, with a metamodel linking the coefficients and the input parameters. For the second approach, two types of decomposition are proposed, but no metamodel has been designed for the coefficients yet. This is a topic of future research. These methods are applied to two analytical models and three industrial cases.

**Résumé.** Les codes de simulation numérique sont couramment utilisés pour étudier des phénomènes complexes. Ils sont cependant souvent coûteux en temps de calcul. Pour certaines études statistiques (*e.g.* gestion d'actifs, analyse de sensibilité) ou problèmes d'optimisation (*e.g.* réglage des paramètres d'un modèle moléculaire), un grand nombre d'appels à de tels codes est nécessaire. Il est alors plus approprié d'élaborer une approximation – ou *métamodèle* – peu coûteuse en temps de calcul à partir d'un plan d'expérience. Le sujet de cet article est la définition de métamodèles pour des codes *stochastiques*. Contrairement aux codes *déterministes*, les codes stochastiques peuvent renvoyer des résultats différents lorsqu'ils sont appelés plusieurs fois avec le même jeu de paramètres. Dans cet article deux approches sont proposées pour la construction d'un métamodèle de la densité de probabilité de la sortie d'un code stochastique. La première repose sur la *régression à noyau* et la seconde consiste à décomposer la densité de la sortie du code dans une base de densités de probabilité bien choisies, avec un métamodèle exprimant les coefficients en fonction des paramètres d'entrée. Pour la seconde approche, deux types de décomposition sont proposés, mais aucun métamodèle n'a encore été élaboré pour les coefficients: c'est un sujet de recherches futures. Ces méthodes sont appliquées à deux cas analytiques et trois cas industriels.

---

<sup>1</sup> EDF R&D, MRI, Chatou, France; e-mail: [vincent.moutoussamy@edf.fr](mailto:vincent.moutoussamy@edf.fr)

<sup>2</sup> Université Paul Sabatier, Toulouse, France

<sup>3</sup> CEA, DEN, DER, F-13108 Saint-Paul-Lez-Durance, France; e-mail: [simon.nanty@cea.fr](mailto:simon.nanty@cea.fr)

<sup>4</sup> Université Joseph Fourier, Grenoble, France

<sup>5</sup> IFPEN, DTIMA, Rueil-Malmaison, France; e-mail: [benoit.pauwels@ifpen.fr](mailto:benoit.pauwels@ifpen.fr)

## INTRODUCTION

Numerical simulation codes are very common tools to study complex phenomena. However these computer codes can be intricate and time-consuming: a single run may take from a few hours to a few days. Therefore, when carrying out statistical studies (*e.g.* sensitivity analysis, reliability methods) or optimization processes (*e.g.* tuning of a molecular model), it is more convenient to build an approximation – or *metamodel* – of the computer code. Since this metamodel is supposed to take less time to run than the actual computer code, it is substituted for the numerical code. A numerical code is said *deterministic* if it gives the same output each time it is called with a given set of input parameters. For this case, the design of metamodels has already been widely studied. The present work deals with the formulation of a metamodel for a numerical code which is not deterministic, but *stochastic*. Two types of numerical codes can be covered by this framework. First, the code under consideration may be *intrinsically* stochastic: if the code is called with the same input several times, the output may differ. Thus, the output conditionally to the input is itself a random variable. Second, the computer code may be deterministic with complex inputs. For example, these inputs may be functions or spatiotemporal fields. These complex inputs are not treated explicitly and are called *uncontrollable parameters*, while the other parameters are considered as *controllable parameters*. When the code is run several times with the same set of controllable parameters, the uncontrollable parameters may change, and thus different output values may be obtained: the output is a random variable conditionally to the controllable parameters.

Two types of approaches can be found in the literature for stochastic code metamodeling. In the first one, strong hypotheses are made on the shape of the conditional distribution of the output. In [24] the authors consider that the output is normally distributed and estimate its mean and variance, in [23] the authors assume that the distribution of the output of the code is a mixture of normal distributions and estimate its mean. In the second approach, the mean and variance of the output are fitted jointly. In [13, 31] the authors use joint generalized linear models and joint generalized additive models respectively to fit these quantities. Joint metamodels based on Gaussian processes have been developed in [3, 17]. As we can observe, so far in the literature the metamodels proposed only provide an estimate for one or two of the output distribution moments, sometimes under strong hypotheses on the distribution of the output.

Our approach does not require such a priori hypotheses. Furthermore we study the whole distribution of the output, allowing to derive estimations for the mean, variance and other moments, and quantiles as well. The metamodeling is not carried out on the stochastic computer code under consideration itself, but rather on the code  $G$  that has same input and that returns the probability density function of the computer code output rather than just a single random value:

$$G : \mathcal{X} \subset \mathbb{R}^d \rightarrow \mathcal{F} \\ \mathbf{x} \mapsto G(\mathbf{x}, \bullet),$$

where  $\mathcal{F}$  is the space of probability density functions with support in an interval  $I$  of  $\mathbb{R}$ :

$$\mathcal{F} = \left\{ f : I \rightarrow \mathbb{R}_+ : \int_I f = 1 \right\}.$$

Let  $\mathbf{X}_N = \{(\mathbf{x}_i, f_i) : i = 1, \dots, N\} \subset (\mathcal{X} \times \mathcal{F})^N$  be a *training set* ( $N \in \mathbb{N} \setminus \{0\}$ ): we suppose that we have access to the probability density functions  $f_i$  of the outputs associated with  $N$  given sets of input parameters  $\mathbf{x}_i$ , *i.e.*

$$f_i = G(\mathbf{x}_i, \bullet) \quad (i = 1, \dots, N).$$

In practice, for a given  $i$  in  $\{1, \dots, N\}$ , the output of a stochastic computer code is not a probability density function  $f_i$ . We only have access to a finite sample of output values  $y_{i1}, \dots, y_{iM}$  which can be written as a vector  $[y_{i1} \ \cdots \ y_{iM}]$  representing the function. In real applications (*cf.* Sections 3.1, 3.2 and 3.3) we chose to

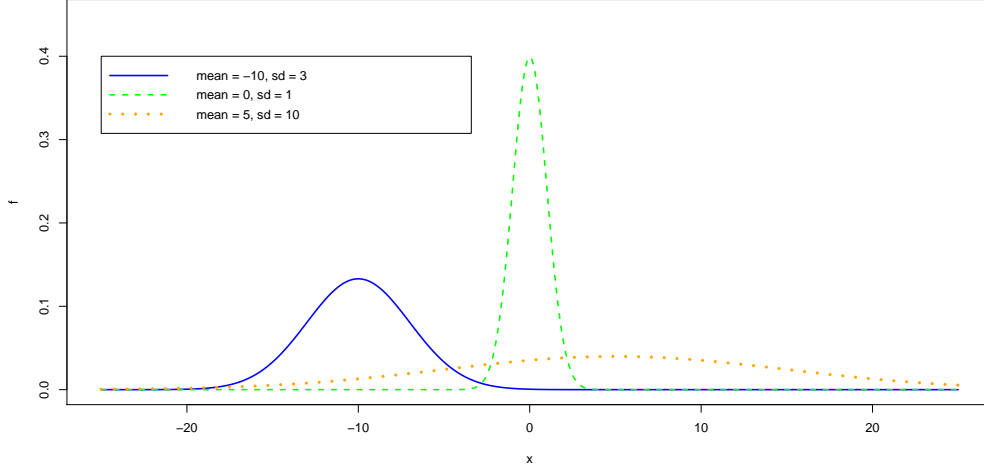


FIGURE 1. Probability density function given by the numerical code  $G$  for different values of the input parameters.

derive each probability density function  $f_i$  thanks to kernel density estimation. In this non-parametric method developed by Rosenblatt [25] and Parzen [21], the estimator of the probability density function  $f_i$ ,  $i = 1, \dots, N$  at a point  $y$  in  $I$  is defined as follows:

$$\hat{f}_i(y) = \frac{1}{M} \sum_{j=1}^M K_H(y_{ij} - y),$$

where  $H \in \mathbb{R}$  is the bandwidth and  $K_H$  is the kernel. Let  $\mathbf{x}_0 \in \mathcal{X}$  be a new set of input parameters, we want to estimate the probability density function  $f_0$  of the corresponding output:

$$\begin{aligned} f_0 &= G(\mathbf{x}_0, \bullet) : I \subset \mathbb{R} \rightarrow \mathbb{R}_+ \\ t &\mapsto f_0(t) = G(\mathbf{x}_0, t). \end{aligned}$$

For example, let us suppose that  $\mathcal{X} = \mathbb{R} \times \mathbb{R}_+^*$ ,  $\mathbf{x}_0 = (\mu_0, \sigma_0)$  and  $G((\mu_0, \sigma_0), \bullet) = \frac{e^{-\frac{1}{2}\left(\frac{\bullet - \mu_0}{\sigma_0}\right)^2}}{\sqrt{2\pi\sigma_0^2}}$ . Figure 1 shows the output of  $G$  for different values of  $\mu_0$  and  $\sigma_0$ .

Section 1 expounds two types of metamodeling for stochastic computer codes: the first one provides a metamodel with functional outputs based on *kernel regression*, the second one consists in choosing relevant probability density functions to express the probability density of the output as a convex combination of them, the coefficients being linked to the input parameters by another metamodel. In Section 2, the previous methods are applied to two analytical test cases. Finally, the same methods are carried out on three industrial applications in Section 3.

## 1. METAMODELS FOR PROBABILITY DENSITY FUNCTIONS

In this section we expound our two approaches of formulating a metamodel for  $G$ . The first relies on *kernel regression* (cf. Section 1.1). The second consists in writing  $f_0$  as a convex combination of well-chosen probability

density functions (cf. Section 1.2). In the following we denote  $\|\bullet\|_{L^2}$  and  $\langle \bullet, \bullet \rangle_{L^2}$  the norm and inner product of  $L^2(I)$  respectively: for all  $u$  and  $v$  in  $L^2(I)$ ,

$$\|u\|_{L^2} = \left( \int_I u^2(t) dt \right)^{1/2} \quad \text{and} \quad \langle u, v \rangle_{L^2} = \int_I u(t)v(t) dt.$$

### 1.1. Kernel regression-based metamodel

In this subsection, first we apply the classical kernel regression method to the problem under consideration, then a new kernel regression estimator involving the Hellinger distance is introduced and, finally, some perspectives of improvement are suggested.

#### 1.1.1. Classical kernel regression

Let  $\mathbf{x}_0 \in \mathcal{X}$  and  $f_0 = G(\mathbf{x}_0, \bullet)$ . We suppose that the sample set  $\mathbf{X}_N = \{(\mathbf{x}_i, f_i) : i = 1, \dots, N\} \subset (\mathcal{X} \times \mathcal{F})^N$  is available. We want to estimate  $f_0$  by  $\hat{f}_0 \in \mathcal{F}$ , knowing  $\mathbf{X}_N$ . However, forcing  $\hat{f}_0$  to be in  $\mathcal{F}$  can be difficult in practice. We propose a first estimator given by

$$\hat{f}_0 = \sum_{i=1}^N \alpha_i f_i$$

where  $\alpha_i \in \mathbb{R}$  ( $i = 1, \dots, N$ ). In order to have  $\hat{f}_0$  in  $\mathcal{F}$ , we impose

$$\begin{aligned} \alpha_i &\geq 0 \quad (i = 1, \dots, N), \\ \sum_{i=1}^N \alpha_i &= 1. \end{aligned}$$

The *non-parametric kernel regression* verifies the previous constraints. This estimator, introduced in [20, 30], can be written as follows:

$$\hat{f}_0 = \sum_{i=1}^N \frac{K_H(\mathbf{x}_i, \mathbf{x}_0)}{\sum_{j=1}^N K_H(\mathbf{x}_j, \mathbf{x}_0)} f_i, \quad (1)$$

where  $K_H : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  is a *kernel function*. In the following, the *Gaussian kernel* is used:

$$K_H(\mathbf{x}, \mathbf{y}) = \frac{1}{\sqrt{2\pi \det(H)}} e^{-(\mathbf{x}-\mathbf{y})^\top H^{-1}(\mathbf{x}-\mathbf{y})} \quad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^d),$$

where

$$H = \text{diag}(h_1, \dots, h_d)$$

is the (diagonal) *bandwidth matrix*, with  $h_1, \dots, h_d \in \mathbb{R}_+$ . The higher  $h_j$  is, the more points  $\mathbf{x}_i$  are taken into account in the regression in the direction  $j$ . The method relies on the intuition that if  $\mathbf{x}_0$  is close to  $\mathbf{x}_i$  then  $f_0$  will be more influenced by the probability density function  $f_i$  (see Figure 2). We want that the bandwidth matrix  $H^*$  minimizes the global error of estimation. That means finding  $H^* = \text{diag}(h_1^*, \dots, h_d^*)$  such that

$$\text{diag}(h_1^*, \dots, h_d^*) \in \arg \min_{h_1, \dots, h_d \in \mathbb{R}_+} \int_{\mathcal{X}} \|\hat{f}_0 - f_0\|_{L^2}^2 d\mathbf{x}_0.$$

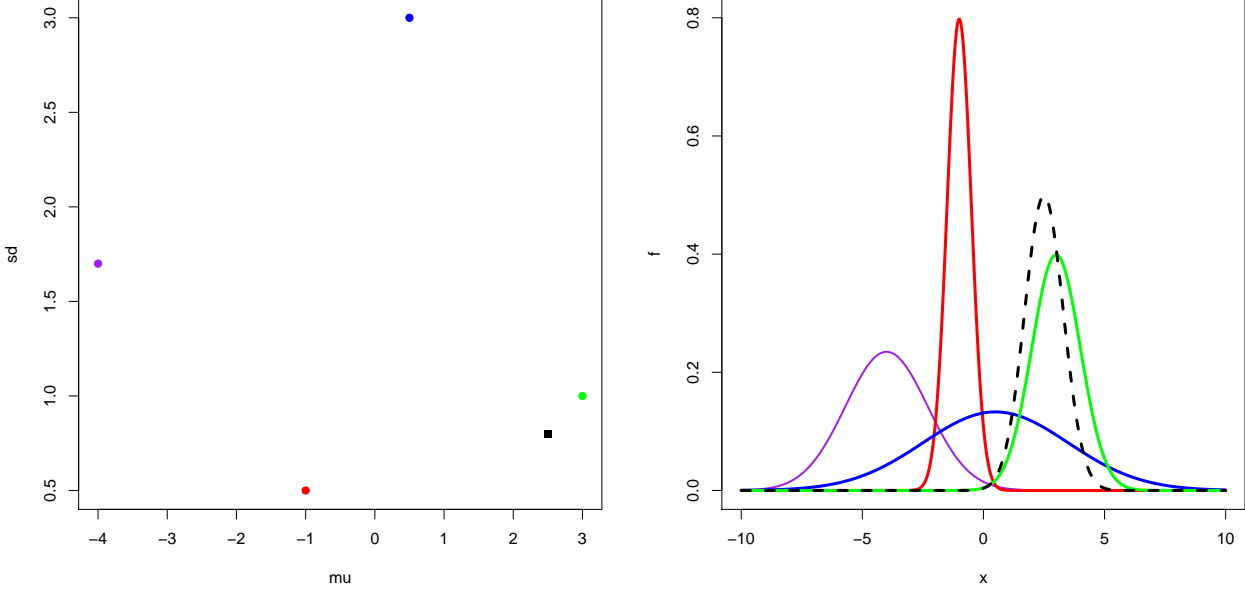


FIGURE 2. Left: the spots correspond to design points in  $\mathcal{X}$ . Right: the associated probability density functions given by the code  $G$  defined in the introduction. The probability density function associated to the black square (dashed line) in left Figure seems to be closer to the probability density function represented in green than to the others.

In order to get an approximation of  $H^*$ , one uses the *leave-one-out cross-validation*, as proposed in Hardle and Marron [11]. Then

$$H^* \in \arg \min_{h_1, \dots, h_d \in \mathbb{R}_+} \sum_{i=1}^N \left\| \hat{f}_{-i} - f_i \right\|_{L^2}^2,$$

with

$$\hat{f}_{-i} = \sum_{\substack{j=1 \\ j \neq i}}^N \alpha_j^{-i} f_j$$

$$\alpha_j^{-i} = \frac{e^{-(\mathbf{x}_i - \mathbf{x}_j)^t H^{-1} (\mathbf{x}_i - \mathbf{x}_j)}}{\sum_{\substack{l=1 \\ l \neq i}}^N e^{-(\mathbf{x}_i - \mathbf{x}_l)^t H^{-1} (\mathbf{x}_i - \mathbf{x}_l)}}.$$

This optimization problem is solved with the L-BFGS-B algorithm developed by Byrd et al. [5]. Finally, the kernel regression metamodel is given by

$$\hat{f}_0 = \sum_{i=1}^N \alpha_{i, H^*} f_i, \quad (2)$$

with

$$\alpha_{i,H^*} = \frac{e^{-(\mathbf{x}_0 - \mathbf{x}_i)^T (H^*)^{-1} (\mathbf{x}_0 - \mathbf{x}_i)}}{\sum_{j=1}^N e^{-(\mathbf{x}_0 - \mathbf{x}_j)^T (H^*)^{-1} (\mathbf{x}_0 - \mathbf{x}_j)}}.$$

### 1.1.2. Kernel regression with the Hellinger distance

The kernel regression introduced in equation (1) can be considered as a minimization problem. The estimator of the function  $f_0$  corresponding to the point  $\mathbf{x}_0$  is as follows:

$$\hat{f}_0 = \arg \min_{f \in \mathcal{F}} \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) \int_I (f_i - f)^2. \quad (3)$$

The kernel regression is therefore a locally weighted constant regression, which is fitted by minimizing weighted least squares. The equivalence between estimators (1) and (3) is proven in Appendix A.1. We observe that the  $L^2$  distances between the sample functions and the unknown function appear in the objective function of the minimization problem (3).

Another classical example of distance between probability density functions is the Hellinger distance. We derived a new kernel regression estimator with respect to this distance replacing the  $L^2$  distance by the Hellinger distance in equation (3). The kernel estimator thus takes the following form:

$$\hat{f}_0 = \arg \min_{f \in \mathcal{F}} \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) \int_I \left( \sqrt{f_i(t)} - \sqrt{f(t)} \right)^2 dt. \quad (4)$$

The following analytical expression can be derived for this Hellinger distance-based estimator:

$$\hat{f}_0 = \frac{\left( \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) \sqrt{f_i} \right)^2}{\int_I \left( \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) \sqrt{f_i} \right)^2}. \quad (5)$$

The calculation that leads to (5) is provided in Appendix A.2.

From now on, we will denote  $\hat{f}_{0,L^2}$  and  $\hat{f}_{0,He}$  the estimators given by equations (2) and (5) respectively. Let us illustrate the Hellinger kernel regression on the four gaussian density functions given on Figure 2. On Figure 3 the same  $N = 4$  learning curves (in dark blue, red, light blue and green) are represented, along with an unknown output  $f_0$  (in dashed black) and its associated Hellinger kernel estimator  $\hat{f}_{0,He}$  (in orange). We can observe that, although the output is not very well estimated, most of the weight is given to the curve corresponding to the closest vector of parameters (in green).

### 1.1.3. Perspectives

One of the well-known major drawbacks of the kernel regression is due to the curse of dimensionality. The bandwidth estimation quality is deteriorating as the dimension  $d$  of the problem increases for a constant sample size. First, as the sample points are sparsely distributed in high dimension, the local averaging performs poorly. Second, the optimization problem to find the optimal bandwidth is more complex in higher dimension. In the presented work, kernel regression has been used with a fixed bandwidth for all sample points. The Adaptive or "Variable-bandwidth" kernel estimation techniques have been proposed in the literature to use

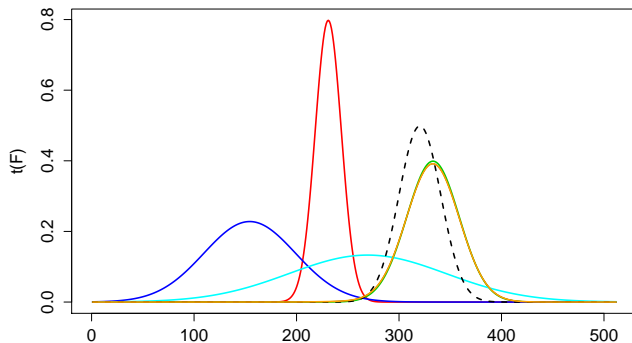


FIGURE 3. The orange curve is the Hellinger kernel prediction for the (unknown) output in black, built from the dark blue, red, light blue and green learning probability density functions ( $N = 4$ ).

varying bandwidth over the domain of the parameters  $\mathcal{X}$ . See Muller and Stadtmuller [19] and Terrell and Scott [27] for more details on these methods. Terrell and Scott [27] review several adaptive methods and compare their asymptotic mean squared error, while Muller and Stadtmuller [19] propose a varying bandwidth selection method. In numerous cases, using non-constant bandwidths leads to better estimates as it enables to better capture the different features of the curves. Hence, the quality of the proposed methods could be improved by using adaptive kernel regression. However, the use of kernel regression is not advised in high dimension.

## 1.2. Functional decomposition-based metamodel

In this section, we aim at building a set of *basis functions*  $w_1, \dots, w_q$  in  $\mathcal{F}$  to approximate  $f_0$  by an estimator of the form

$$\hat{f}_0 = \sum_{k=1}^q \psi_k(\mathbf{x}_0) w_k, \quad (6)$$

where  $\psi_k$  are functions from  $\mathcal{X}$  to  $\mathbb{R}$  ( $k = 1, \dots, q$ ), called *coefficient functions*, satisfying

$$\begin{cases} \psi_k(\mathbf{x}) \geq 0 \\ \sum_{k=1}^q \psi_k(\mathbf{x}) = 1 \end{cases} \quad (\mathbf{x} \in \mathcal{X}).$$

Kernel regression and functional decomposition have different goals but have similar forms. The former provides an estimator which is a convex combination of all the  $N$  sampled probability density functions  $f_1, \dots, f_N$ . The latter gives a combination of fewer density functions built from the experimental data, thus lying in a smaller space, in order to *reduce the dimension* of the problem. We observe that the estimator provided by classical kernel regression (1) is of the form (6) with  $q = N$  (no dimension reduction),  $w_k = f_k$  for  $k = 1, \dots, q$  and

$$\psi_k(\mathbf{x}_0) = \frac{K_H(\mathbf{x}_k, \mathbf{x}_0)}{\sum_{j=1}^N K_H(\mathbf{x}_j, \mathbf{x}_0)} \quad (k = 1, \dots, q).$$



We will build the basis functions and the coefficient functions so that they fit the available data  $\mathbf{X}_N = \{(\mathbf{x}_i, f_i) : i = 1, \dots, N\}$ . Thus the functions  $w_1, \dots, w_q$  will be designed so that, for all  $i$  in  $\{1, \dots, N\}$ ,  $f_i$  is approximated by

$$\hat{f}_i = \sum_{k=1}^q \psi_{ik} w_k$$

$$\text{where } \begin{cases} \psi_{ik} = \psi_k(\mathbf{x}_i) & (k = 1, \dots, q) \\ \psi_{ik} \geq 0 & (k = 1, \dots, q) \\ \sum_{k=1}^q \psi_{ik} = 1. \end{cases} \quad (7)$$

For all  $i$  in  $\{1, \dots, N\}$ , we require the approximation  $\hat{f}_i$  of  $f_i$  to be a probability density function as well; therefore  $\hat{f}_i$  must be non-negative and have its integral equal to 1.

We study three ways of choosing the basis functions. In Section 1.2.1 we discuss the adaptation of *Functional Principal Components Analysis* (FPCA) to compute the basis functions:  $w_1, \dots, w_q$  are then orthonormal with integral equal to 1 and their negative values are interpreted as 0, hence the predictions on the experimental design are probability density functions. In the following two subsections we propose two new approaches to probability density function decomposition. The non-negativity constraint is taken into account directly rather than through a post-treatment of FPCA results (setting negative values to zero and renormalizing). In Section 1.2.2, we propose to adapt the *Empirical Interpolation Method* or *Magic Points method* [16] to our problem: we obtain an algorithm to select the basis functions (not necessarily orthogonal) among the sample distributions  $f_1, \dots, f_N$  (while losing the interpolation property). It follows that the predictions on the experimental design are probability density functions. In Section 1.2.3 a third approach is proposed, consisting in computing directly the basis functions that minimize the  $L^2$ -approximation error on the experimental design (without imposing orthogonality):

$$\sum_{i=1}^N \left\| f_i - \sum_{k=1}^q \psi_{ik} w_k \right\|_{L^2}^2. \quad (8)$$

We introduce an iterative procedure to ease the numerical computation. The basis functions obtained are probability density functions and so are the predictions on the experimental design. Finally, in Section 1.2.4 we discuss the formulation of the coefficient functions.

### 1.2.1. Constrained Principal Component Analysis (CPCA)

Among all the *dimension reduction* techniques, one of the best known is the *Functional Principal Component Analysis* (FPCA), developed by Ramsay and Silverman [22], which is based on *Principal Component Analysis* (PCA). Let us denote  $\bar{f}$  the *mean estimator* of  $f_1, \dots, f_N$ :

$$\bar{f} = \frac{1}{N} \sum_{i=1}^N f_i.$$

The goal of FPCA is to build an orthonormal basis  $w_1, \dots, w_q$  for the *centered functions*  $f_i^c = f_i - \bar{f}$  ( $i = 1, \dots, N$ ) so that the projected variance of every centered function onto the functions  $w_1, \dots, w_q$  is maximized. The following maximization problem is thus considered:

$$\begin{aligned} & \text{Minimize}_{w_1, \dots, w_q: I \rightarrow \mathbb{R}} \sum_{i=1}^N \left\| f_i^c - \sum_{k=1}^q \langle f_i^c, w_k \rangle_{L^2} w_k \right\|_{L^2}^2 \\ & \text{subject to } \langle w_k, w_l \rangle_{L^2} = \delta_{kl} \quad (k, l = 1, \dots, q), \end{aligned}$$

where  $\delta_{kl}$  is 1 when  $k = l$  and 0 otherwise. Therefore the estimator of  $f_i$  ( $i = 1, \dots, N$ ) is

$$\hat{f}_i = \bar{f} + \sum_{k=1}^q \psi_{ik} w_k,$$

with  $\psi_{ik} = \langle f_i^c, w_k \rangle_{L^2}$  ( $k = 1, \dots, q$ ). To solve this minimization problem, Ramsay and Silverman [22] proposes to express the sample functions on a spline basis. Then PCA can be applied to the coefficients of the functions on the spline basis. They also propose to apply PCA directly to the discretized functions. The FPCA decomposition ensures that  $\int_I \hat{f}_i = \int_I f_i$  for  $i = 1, \dots, N$  so that the integrals of the sample probability density functions approximations  $\hat{f}_i$  are equal to one. However, the FPCA decomposition does not ensure that the estimators are non-negative.

Delicado [8] proposes to apply FPCA to the logarithm transform  $g_i = \log f_i$  to ensure the positivity of the prediction  $\hat{f}_i = \exp \hat{g}_i$  ( $i = 1, \dots, N$ ). However this approach does not ensure that the approximations are normalized.

Affleck-Graves et al. [1] have proposed to put a non-negativity constraint on the first basis function  $w_1$  and let the other basis functions free. Indeed they claim that in general it is not possible to put such a constraint on the basis functions  $w_2, \dots, w_q$ . Hence the first basis function is forced to be non-negative but not the other ones. Therefore, it cannot be guaranteed that the approximations are non-negative.

Another method has been proposed by Kneip and Utikal [14] to ensure that the FPCA approximations are non-negative. The FPCA method is applied to the sample functions. Then the negative values of the probability density functions approximations are interpreted as 0 and they are normalized to one. In the numerical studies of Sections 2 and 3, we will refer to this method as CPCA, for Constrained Principal Component Analysis.

### 1.2.2. Modified Magic Points (MMP)

The *Empirical Interpolation Method* (EIM) or *Magic points* (MP) [16] is a *greedy* algorithm that builds interpolators for a set of functions. Here the set of functions under consideration is  $G(\mathcal{X}, \bullet) = \{G(\mathbf{x}, \bullet) : \mathbf{x} \in \mathcal{X}\}$ . The algorithm iteratively picks a set of basis functions in  $G(\mathcal{X}, \bullet)$  and a set of *interpolation points* in  $I$  to build an interpolator with respect to these points. Let  $\mathbf{x}$  be in  $\mathcal{X}$  and  $f_{\mathbf{x}} = G(\mathbf{x}, \bullet)$ . At step  $q - 1$ , the current basis functions and interpolation points are denoted  $w_1, \dots, w_{q-1} : I \rightarrow \mathbb{R}$  and  $t_{i_1}, \dots, t_{i_{q-1}} \in I$  respectively. The interpolator  $I_{q-1}[f_{\mathbf{x}}]$  of  $f_{\mathbf{x}}$  is defined as a linear combination of the basis functions:

$$I_{q-1}[f_{\mathbf{x}}] = \sum_{k=1}^{q-1} \psi_k(\mathbf{x}) w_k,$$

where the values of the coefficient functions  $\psi_1, \dots, \psi_{q-1} : \mathcal{X} \rightarrow \mathbb{R}$  at  $\mathbf{x}$  are uniquely defined by the following *interpolation equations* (see [16] for existence and uniqueness):

$$I_{q-1}[f_{\mathbf{x}}](t_{i_k}) = f_{\mathbf{x}}(t_{i_k}) \quad (k = 1, \dots, q - 1).$$

At step  $q$ , one picks the parameter  $\mathbf{x}_{i_q}$  such that  $f_{\mathbf{x}_{i_q}} (= G(\mathbf{x}_{i_q}, \bullet))$  is the element of  $G(\mathcal{X}, \bullet)$  whose  $L^2$ -distance from its own current interpolator  $I_{q-1}[f_{\mathbf{x}_{i_q}}]$  is the highest. The next interpolation point  $t_{i_q}$  is the one that maximizes the gap between  $f_{\mathbf{x}_{i_q}}$  and  $I_{q-1}[f_{\mathbf{x}_{i_q}}]$ . Then the new basis function  $w_q$  is defined as a particular linear combination of  $f_{\mathbf{x}_{i_q}}$  and  $I_{q-1}[f_{\mathbf{x}_{i_q}}]$ . Let us be more specific by summarizing the algorithm [4] hereafter.

**Algorithm 1.1** (MP).

- Set  $q \leftarrow 1$  and  $I_0 \leftarrow 0$ .
- While  $\varepsilon > \text{tol}$  do
  - (a) choose  $\mathbf{x}_{i_q}$  in  $\mathcal{X}$ :

$$\mathbf{x}_{i_q} \leftarrow \arg \sup_{\mathbf{x} \in \mathcal{X}} \|f_{\mathbf{x}} - I_{q-1}[f_{\mathbf{x}}]\|_{L^2}$$

and the associated interpolation point  $t_{i_q}$ :

$$t_{i_q} \leftarrow \arg \sup_{t \in I} \left| f_{\mathbf{x}_{i_q}}(t) - I_{q-1}[f_{\mathbf{x}_{i_q}}](t) \right|,$$

(b) define the next element of the basis:

$$w_q \leftarrow \frac{f_{\mathbf{x}_{i_q}}(\bullet) - I_{q-1}[f_{\mathbf{x}_{i_q}}](\bullet)}{f_{\mathbf{x}_{i_q}}(t_{i_q}) - I_{q-1}[f_{\mathbf{x}_{i_q}}](t_{i_q})},$$

(c) compute the interpolation error:

$$\varepsilon \leftarrow \|err_q\|_{L^\infty(\mathcal{X})}, \text{ where } err_q : \mathbf{x} \mapsto \|f_{\mathbf{x}} - I_q[f_{\mathbf{x}}]\|_{L^2},$$

(d) set  $q \leftarrow q + 1$ .

This method has been successfully applied *e.g.* to a heat conduction problem, crack detection and damage assessment of flawed materials, inverse scattering analysis [6], parameter-dependent convection-diffusion problems around rigid bodies [28], in biomechanics to describe a stenosed channel and a bypass configuration [26].

In general, the interpolators provided by the MP algorithm are not probability density functions. Therefore we modified the MP method so that the approximations are non-negative with integral equal to one. In the derived new method the interpolation is lost, but the basis functions are picked in a similar greedy way. In the following we denote  $A_{q-1}[f_{\mathbf{x}}]$  the interpolator at step  $q - 1$  of the function  $f_{\mathbf{x}}$  associated with a parameter  $\mathbf{x}$  in  $\mathcal{X}$ :

$$A_{q-1}[f_{\mathbf{x}}] = \sum_{k=1}^{q-1} \psi_k(\mathbf{x}) w_k,$$

where the values of the coefficient functions at  $\mathbf{x}$  are now defined as solutions of the following convex quadratic program:

$$\begin{aligned} & \underset{\psi_1(\mathbf{x}), \dots, \psi_{q-1}(\mathbf{x}) \in \mathbb{R}}{\text{Minimize}} && \left\| f_{\mathbf{x}} - \sum_{k=1}^{q-1} \psi_k(\mathbf{x}) w_k \right\|_{L^2}^2 \\ & \text{subject to} && \begin{cases} \psi_k(\mathbf{x}) \geq 0 & (k = 1, \dots, q-1) \\ \sum_{k=1}^{q-1} \psi_k(\mathbf{x}) = 1. \end{cases} \end{aligned} \quad (9)$$

At step  $q$ , the parameter  $\mathbf{x}_{i_q}$  is chosen in the same way as in the MP algorithm. The new basis function  $w_q$  is then  $f_{\mathbf{x}_{i_q}}$ . The *Modified Magic Points* (MMP) algorithm is detailed below.

**Algorithm 1.2** (MMP).

- Set  $q \leftarrow 1$  and  $I_0 \leftarrow 0$ .
- While  $\varepsilon > tol$  do
  - (a) choose  $\mathbf{x}_{i_q}$  in  $\mathcal{X}$ :

$$\mathbf{x}_{i_q} \leftarrow \arg \sup_{\mathbf{x} \in \mathcal{X}} \|f_{\mathbf{x}} - A_{q-1}[f_{\mathbf{x}}]\|_{L^2},$$

(b) define the next element of the basis:

$$w_q \leftarrow f_{\mathbf{x}_{i_q}},$$

(c) compute the estimation error:

$$\varepsilon \leftarrow \|err_q\|_{L^\infty(\mathcal{X})}, \text{ where } err_q : \mathbf{x} \mapsto \|f_{\mathbf{x}} - A_q[f_{\mathbf{x}}]\|_{L^2},$$

(d) set  $q \leftarrow q + 1$ .

**Remark 1.3.** In step (a) of the MMP algorithm, the  $L^2$  norm can be replaced by the Hellinger distance. This yields a new algorithm which is tested along with the previous one in Section 2.

In practice, we do not apply the MMP algorithm to the whole set  $G(\mathcal{X}, \bullet)$  since the computation of the functions is time-consuming. It is rather applied to the available sample set  $\{f_1, \dots, f_N\}$ . Thus the MMP algorithm is a way to select the most relevant functions among the sample  $f_1, \dots, f_N$ . Then the step (a) of the algorithm is only a finite maximization. Furthermore we build a regular grid of  $M$  points  $t_1, \dots, t_M$  in the interval  $I$ :

$$\begin{aligned} t_j &= t_1 + (j-1)\Delta t & (j = 1, \dots, M) \\ \text{with } \Delta t &\in \mathbb{R}_+, \end{aligned} \tag{10}$$

and we discretize the functions  $f_i$  ( $i = 1, \dots, N$ ) and  $w_k$  ( $k = 1, \dots, q$ ) on this grid:

$$\begin{aligned} f_{ij} &= f_i(t_j) \\ w_{kj} &= w_k(t_j) \end{aligned} \quad (j = 1, \dots, M).$$

Thus we can rewrite the problem (9) in the following way for all  $i \in \{1, \dots, N\}$ :

$$\begin{aligned} &\text{Minimize}_{\psi_{i1}, \dots, \psi_{iq}} \sum_{j=1}^M \left[ f_{ij} - \sum_{k=1}^q \psi_{ik} w_{kj} \right]^2 \Delta t \\ &\text{subject to} \quad \begin{cases} \psi_{i1}, \dots, \psi_{iq} \geq 0 \\ \sum_{k=1}^q \psi_{ik} = 1. \end{cases} \end{aligned} \tag{11}$$

These  $N$  problems are convex quadratic programs with  $q$  unknowns and  $q + 1$  linear constraints, they are easily solvable for basis sizes smaller than  $10^4$ .

**Remark 1.4.** Let us suppose  $q = 1$ , *i.e.* the basis contains only one function. The equality constraint in the minimization (11) reduces to  $\psi_{i1} = 1$ : the approximation is the same for each function  $f_i$  in the sample.

### 1.2.3. Alternate Quadratic Minimizations (AQM)

Our third functional decomposition approach consists in tackling directly the problem consisting in minimizing the  $L^2$  norm of the approximation error (8) with basis functions  $w_1, \dots, w_q$  in  $\mathcal{F}$  and coefficients  $\psi_{ik}$  satisfying the constraints (7). It differs from PCA because no orthonormality condition is imposed, and the decomposition is carried out directly on the raw data, *i.e.* without centering. This functional minimization program can be written as follows.

$$\begin{aligned} &\text{Minimize}_{\substack{w_k, \psi_{ik} \\ k=1, \dots, q, i=1, \dots, N}} \frac{1}{2} \sum_{i=1}^N \left\| f_i - \sum_{k=1}^q \psi_{ik} w_k \right\|_{L^2}^2 \\ &\text{subject to} \quad \begin{cases} w_k \in \mathcal{F} & (k = 1, \dots, q) \\ \psi_{ik} \geq 0 & (i = 1, \dots, N)(k = 1, \dots, q) \\ \sum_{k=1}^q \psi_{ik} = 1 & (i = 1, \dots, N). \end{cases} \end{aligned} \tag{12}$$

Just as before, instead of working on actual functions, we consider discretizations of them. We discretize the interval  $I$  in the same way (10) and we set

$$\begin{aligned}\mathbf{f}_i &= [f_i(t_1) \quad \cdots \quad f_i(t_M)] \quad (i = 1, \dots, N) \\ &= [\mathbf{f}_{i1} \quad \cdots \quad \mathbf{f}_{iM}] \\ \mathbf{w}_k &= [w_k(t_1) \quad \cdots \quad w_k(t_M)] \quad (k = 1, \dots, q) \\ &= [\mathbf{w}_{k1} \quad \cdots \quad \mathbf{w}_{kM}].\end{aligned}$$

The functional minimization program (12) is then replaced by the following vectorial minimization program (in this section  $\|\bullet\|_{\mathbb{R}^M}$  designates the euclidean norm on  $\mathbb{R}^M$ ).

$$\begin{aligned}\text{Minimize}_{\substack{\mathbf{w}_k, \psi_{ik} \\ k=1, \dots, q, i=1, \dots, N}} & \frac{1}{2} \sum_{i=1}^N \left\| \mathbf{f}_i - \sum_{k=1}^q \psi_{ik} \mathbf{w}_k \right\|_{\mathbb{R}^M}^2 \\ \text{subject to} & \begin{cases} \mathbf{w}_{kj} \geq 0 & (k = 1, \dots, q)(j = 1, \dots, M) \\ \sum_{j=1}^M \mathbf{w}_{kj} \Delta t = 1 & (k = 1, \dots, q) \\ \psi_{ik} \geq 0 & (i = 1, \dots, N)(k = 1, \dots, q) \\ \sum_{k=1}^q \psi_{ik} = 1 & (i = 1, \dots, N). \end{cases} \end{aligned} \quad (13)$$

Let us introduce additional notations to reformulate the objective function and the variables in a more compact way.

$$\Psi = [\psi_{ik}]_{i=1, \dots, N; k=1, \dots, q} = \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_N \end{bmatrix} \quad \mathbf{W} = [\mathbf{w}_{kj}]_{k=1, \dots, q; j=1, \dots, M} = \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_q \end{bmatrix}$$

We recall that the Frobenius norm of a matrix  $\mathbf{A} = [a_{ij}]_{i=1, \dots, N; j=1, \dots, M} \in \mathbb{R}^{N \times M}$  is defined as

$$\|\mathbf{A}\|_{\mathbb{F}} = \sqrt{\sum_{i=1}^N \sum_{j=1}^M a_{ij}^2}.$$

This being said, the vectorial minimization program (13) can be written in matrix form.

$$\begin{aligned}\text{Minimize}_{\substack{\Psi \in \mathbb{R}_+^{N \times q}, \mathbf{W} \in \mathbb{R}_+^{q \times M}}} & \mathcal{O}(\Psi, \mathbf{W}) = \frac{1}{2} \|\mathbf{F} - \Psi \mathbf{W}\|_{\mathbb{F}}^2 \\ \text{subject to} & \begin{cases} \psi_i \mathbf{1} = 1 & (i = 1, \dots, N) \\ \mathbf{w}_k \mathbf{1} = 1/\Delta t & (k = 1, \dots, q). \end{cases} \end{aligned} \quad (14)$$

The objective function of the program (13) is second-order polynomial, hence twice continuously differentiable (but not necessarily convex). Its first-order derivatives are expressed as

$$\begin{aligned}\partial_{\psi_{ik}} \mathcal{O}(\Psi, \mathbf{W}) &= -\mathbf{w}_k (\mathbf{f}_i - \psi_i \mathbf{W})^\top & (i = 1, \dots, N)(k = 1, \dots, q), \\ \partial_{w_{kj}} \mathcal{O}(\Psi, \mathbf{W}) &= -\Psi_{\bullet, k}^\top (\mathbf{F}_{\bullet, j} - \Psi \mathbf{W}_{\bullet, j}) & (k = 1, \dots, q)(j = 1, \dots, M),\end{aligned}$$

where  $\Psi_{\bullet,k}$  is the  $k^{\text{th}}$  column of  $\Psi$ ,  $\mathbf{F}_{\bullet,j}$  and  $\mathbf{W}_{\bullet,j}$  are the  $j^{\text{th}}$  columns of  $\mathbf{F}$  and  $\mathbf{W}$  respectively, and here are the second-order derivatives which will be useful further down:

$$\begin{aligned} \partial_{\psi_{ik}\psi_{ik'}}^2 \mathcal{O}(\Psi, \mathbf{W}) &= \mathbf{w}_k \mathbf{w}_{k'}^T & (i = 1, \dots, N)(k, k' = 1, \dots, q), \\ \partial_{w_{kj}w_{kj'}}^2 \mathcal{O}(\Psi, \mathbf{W}) &= \delta_{jj'} \|\Psi_{\bullet,k}\|_{\mathbb{R}^N}^2 & (k = 1, \dots, q)(j, j' = 1, \dots, M). \end{aligned}$$

The program (14) has  $q(M+N) \approx 10^4$  design variables,  $q(M+N) + q + N \approx 10^4$  constraints and may not be convex. Hence its resolution through the use of a numerical solver may be very time-consuming (it turns out we could not manage to get a satisfactory feasible solution). In order to circumvent this issue we chose to implement (14) as successive convex quadratic minimization programs. The value of the program can be written as follows.

$$\inf_{\substack{(\Psi, \mathbf{W}) \in (\mathbb{R}_+^{N \times M})^2 \\ \psi_i \mathbf{1} = 1 \ (i=1, \dots, N) \\ \mathbf{w}_k \mathbf{1} = 1/\Delta t \ (k=1, \dots, q)}} \mathcal{O}(\Psi, \mathbf{W}) = \inf_{\substack{\mathbf{w}_1 \in \mathbb{R}_+^M \\ \mathbf{w}_1 \mathbf{1} = 1/\Delta t}} \cdots \inf_{\substack{\mathbf{w}_q \in \mathbb{R}_+^M \\ \mathbf{w}_q \mathbf{1} = 1/\Delta t}} \inf_{\substack{\psi_1 \in \mathbb{R}_+^q \\ \psi_1 \mathbf{1} = 1}} \cdots \inf_{\substack{\psi_N \in \mathbb{R}_+^q \\ \psi_N \mathbf{1} = 1}} \mathcal{O}(\Psi, \mathbf{W}).$$

The idea is to minimize the criterion  $\mathcal{O}(\Psi, \mathbf{W})$  working on one line-vector at a time:  $\psi_1$ , then  $\psi_2$ , and so on until  $\psi_N$ , and then  $\mathbf{w}_1$ ,  $\mathbf{w}_2$  and so forth until  $\mathbf{w}_q$ . This process being repeated as many times as necessary – and affordable – to get some kind of convergence.

Let us make one more remark before giving out the algorithm. Let  $i \in \{1, \dots, N\}$ . The two following programs are equivalent – in the sense that their feasible and optimal solutions are the same (we just reduced the criterion to minimize).

$$\begin{aligned} \text{Minimize}_{\psi_i} \quad & \mathcal{O}(\Psi, \mathbf{W}) & \iff & \text{Minimize}_{\psi_i} \quad \frac{1}{2} \|\mathbf{f}_i - \psi_i \mathbf{W}\|_{\mathbb{R}^M}^2 \\ \text{subject to} \quad & \begin{cases} \psi_i \in \mathbb{R}_+^q \\ \psi_i \mathbf{1} = 1 \end{cases} & & \text{subject to} \quad \begin{cases} \psi_i \in \mathbb{R}_+^q \\ \psi_i \mathbf{1} = 1. \end{cases} \end{aligned}$$

The algorithm we implemented to numerically solve (13) is the following.

**Algorithm 1.5** (AQM).

- Initialize  $\Psi$  and  $\mathbf{W}$  with uniform values.

$$\begin{aligned} \psi_{ik} &= 1/q & (i = 1, \dots, N)(k = 1, \dots, q) \\ \mathbf{w}_{kj} &= 1/(M\Delta t) & (k = 1, \dots, q)(j = 1, \dots, M). \end{aligned}$$

- While some stopping criterion is not reached (i.e. a maximum number  $iter_{max} \approx 20$  of iterations),
  - do, for  $i = 1, \dots, N$ ,

$$\begin{aligned} \text{Minimize}_{\psi_i} \quad & \frac{1}{2} \|\mathbf{f}_i - \psi_i \mathbf{W}\|_{\mathbb{R}^M}^2 \\ \text{subject to} \quad & \begin{cases} \psi_i \in \mathbb{R}_+^q \\ \psi_i \mathbf{1} = 1. \end{cases} \end{aligned}$$

- do, for  $k = 1, \dots, q$ ,

$$\begin{aligned} \text{Minimize}_{\mathbf{w}_k} \quad & \mathcal{O}(\Psi, \mathbf{W}) \\ \text{subject to} \quad & \begin{cases} \mathbf{w}_k \in \mathbb{R}_+^M \\ \mathbf{w}_k \mathbf{1} = 1/\Delta t. \end{cases} \end{aligned}$$

Each of the minimization problems addressed in this algorithm has a convex quadratic criterion depending on at most  $M \approx 2000$  variables constrained by the same amount of lower bounds plus one linear equality. These optimization problems were solved with a dual method of the *active set* type, detailed in Goldfarb and Idnani [10], specifically designed to address strictly convex quadratic programs. For a given quadratic program  $P$ , the *active set* of a point  $\mathbf{x}$  is the set of (linear) inequality constraints satisfied with equality (*active constraints*). The algorithm starts from a point  $\mathbf{x}$  such that the set  $A$  of its active constraints (possibly empty) is linearly independent, and the following steps are repeated until all constraints are satisfied. First, a violated constraint  $\mathbf{c}$  is picked. Second, the feasibility of the subproblem  $P(A \cup \{\mathbf{c}\})$  of  $P$  constrained only by the constraint  $\mathbf{c}$  and those of the current active set  $A$  is tested. If  $P(A \cup \{\mathbf{c}\})$  is infeasible then  $P$  is infeasible as well. Otherwise  $\mathbf{x}$  is updated with a point whose active set is a linearly independent subset of  $A \cup \{\mathbf{c}\}$  including  $\mathbf{c}$ . The loop terminates with an optimal solution.

#### 1.2.4. Metamodelling of functional decomposition coefficients

In Sections 1.2.1, 1.2.2 and 1.2.3, three methods have been studied to approximate a sample of probability density functions on basis such that:

$$\hat{f}_i = \sum_{k=1}^q \psi_{ik} w_k.$$

The functions are characterized by their coefficients. The problem dimension is therefore reduced to the basis size  $q$ . A metamodel must now be designed to link the input parameters and the coefficients: a function  $\hat{\boldsymbol{\psi}} : \mathcal{X} \rightarrow \mathbb{R}^q$ , where  $\hat{\boldsymbol{\psi}} = (\hat{\psi}_1, \dots, \hat{\psi}_q)$ . The metamodel is built thanks to the known points:  $\psi_{ik} = \psi_k(\mathbf{x}_i)$  ( $i = 1, \dots, N, k = 1, \dots, q$ ). For a new  $\mathbf{x} \in \mathcal{X}$ , we look for the estimation  $\hat{f}_{\mathbf{x}}$  of the output  $f_{\mathbf{x}}$  of  $G$ :

$$\hat{f}_{\mathbf{x}} = \sum_{k=1}^q \hat{\psi}_k(\mathbf{x}) w_k.$$

For MMP and AQM methods, the searched coefficients must respect the following constraints: for all  $\mathbf{x} \in \mathcal{X}$ ,

$$\begin{cases} \sum_{k=1}^q \psi_k(\mathbf{x}) = 1 \\ \psi_k(\mathbf{x}) \geq 0 \quad (k = 1, \dots, q). \end{cases}$$

Therefore, the outputs of the metamodel to be built must be positive and have their sum equal to one.

So far, the main approach investigated has been to build separate metamodels on each coefficient without constraining the output of the metamodel. When the metamodel is used to make predictions on new points in the input space, the predicted coefficients which are lower than zero are put equal to zero, and the coefficients are then renormalized such that their sum is equal to one. Any type of metamodel can be used in this approach. Polynomial regression, generalized additive models [12] and Gaussian process metamodels have been tested, but the obtained results on different test cases are not convincing. This approach on the presented test cases does not lead to good results. The metamodel error is bigger than the error due to the decomposition on a functional basis.

Future research should be dedicated on the construction of more efficient metamodels taking into account these constraints. Metamodels which ensure the positivity of the output exist in the literature (see for example [7] for Gaussian process with inequality constraints). The main difficulty is to ensure that the sum of the coefficients is equal to one. A solution can be to consider the coefficients as realizations of a random variable on a simplex space. Aitchison [2] proposes to apply a bijective transformation from this simplex space to  $\mathbb{R}^q$ . A metamodel without constraint could then be built between the input parameters and the transformed coefficients.

## 2. NUMERICAL TESTS ON TOY FUNCTIONS

In this section, we apply the methods presented on two toy examples respectively for  $d = 1$  in Section 2.1 and  $d = 5$  in Section 2.2. We compare the relative error of the estimation of different quantities of interest:  $L^1$ ,  $L^2$  and the Hellinger distances between two densities, mean, variance and the (1%, 25%, 75%, 99%)-quantiles. We recall that the Hellinger distance between  $f$  and  $g$  is the  $L^2$  distance between  $\sqrt{f}$  and  $\sqrt{g}$ . The relative error is defined below for the three norms and the other scalar quantities of interest:

$$\begin{aligned}
 & 100 \frac{\int_I |f(t) - \hat{f}(t)|^2 dt}{\int_I f(t)^2 dt} \\
 & 100 \frac{\int_I \left| \sqrt{f(t)} - \sqrt{\hat{f}(t)} \right|^2 dt}{\int_I \sqrt{f(t)}^2 dt} \\
 & 100 \frac{\int_I |f(t) - \hat{f}(t)| dt}{\int_I |f(t)| dt} \\
 & 100 \frac{|u - \hat{u}|}{|u|},
 \end{aligned}$$

where  $u$  is the mean, variance and the studied quantiles. This study is separated in two independent parts for the two toy examples.

First, we compare the error obtained from the two estimators  $\hat{f}_{L^2}$  and  $\hat{f}_{He}$  based on kernel regression. We show the relative error in function of different sizes  $N_1 < \dots < N_k$  of the design of experiments such that a design with  $N_i$  points is included in the designs with  $N_j$  points ( $i < j$ ). For each  $N_i$ , relative error for different quantities of interest are averaged on the same 1000 test points chosen uniformly in  $\mathcal{X}$ . These computations are repeated 25 times with different designs of experiments.

In the second part, we compare the relative error obtained from the construction of basis obtained in Sections 1.2.1, 1.2.2 and 1.2.3 to reconstruct the  $N$  density probability functions of the design of experiments. These two methods decompose the known probability density functions on a functional basis and approximate them. However, no metamodel has been designed between the coefficients of the probability density functions on the basis and the parameters of the computer code, so that no estimation of an unknown probability density function can be done. Therefore, both methods cannot be compared to the kernel regression.

Each probability density function is discretized on 512 points.

### 2.1. A one-dimensional example (Toy example 1)

The first toy example is defined as follows:

$$G(x, \xi_1, \xi_2, U) = (\sin(x(\xi_1 + \xi_2)) + U) \mathbb{1}_{\{\sin(x(\xi_1 + \xi_2)) + U \geq -1\}}$$

where

$$\begin{aligned}
 x & \in \mathcal{X} = [0, 1] \\
 \xi_1 & \sim \mathcal{N}(1, 1) \\
 \xi_2 & \sim \mathcal{N}(2, 1) \\
 U & \sim \mathcal{U}([0, 1]).
 \end{aligned}$$



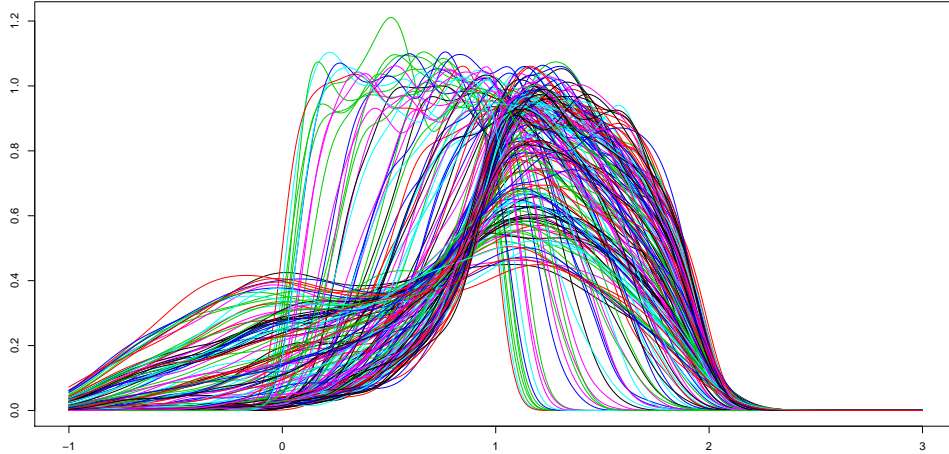


FIGURE 4. Toy example 1: representations of outputs for  $N = 100$  different parameters.

Let  $x_0 \in [0, 1]$ , we want to estimate the probability density function  $f_0$  of the random variable  $G(x_0, \xi_1, \xi_2, U)$ .

#### 2.1.1. Kernel regression

In this section, one compares the estimators based on Hellinger distance and  $L^2$  norm given respectively by equations (2) and (5). Figure 4 represents the output for  $N = 100$  points simulated uniformly and independently in  $\mathcal{X}$ . Figure 5 represents for different values of  $x_0$  the mean (in plain line) and the standard deviation (in dashed line) of the estimation obtained from the two estimators of the kernel regression in red for  $L^2$  norms and in blue for the Hellinger distance. For these four parameters, the two estimators give approximately the same estimations. Figure 6 shows the boxplot of the relative error for  $L^2$  norms and the Hellinger distance. In this first example the estimator based on the Hellinger distance seems slightly better than the one based on the  $L^2$  norm. The errors are very close for the norms and the mean but are more different for other quantities of interest. The variance of the error decreases very quickly for norms. Moreover, the errors are low for most quantities of interest except for 1% and 25% quantiles.

#### 2.1.2. Functional decomposition methods

In this section, the four functional decompositions presented in 1.2, MMP with  $L^2$  and Hellinger distances, AQM and CPCA are compared. The decompositions are applied to learning samples whose sizes are  $N = 50$  and 100 and for decomposition basis sizes ranging from 1 to 20. The relative errors between the learning sample functions and their approximations are computed for the 9 quantities of interest. Figure 7 represents, in logarithmic scale, these relative errors in function of the basis size for the four decompositions. First, the relative errors are low for all quantities of interest and especially for the norms, the mean and the higher quantiles. For the four decompositions, the relative error for norm decreases very evenly. The decrease is quite regular too for the mean, variance, 75% and 99% quantiles. The quality of approximation of the lower quantiles behaves more irregularly, as it sometimes increases with the basis size. CPCA method outperforms the three others for the three distances. This method seems better for all other quantities of interest except the 25% quantile, but the gap between the decompositions is smaller. The relative error for the learning sample of 100 densities is for most of the quantities higher than the relative error for the sample of 50 densities. For the same number of components  $q$ , it is indeed more difficult to approximate a higher number  $N$  of functions, so that the relative error increases with the size of the design of experiments for a constant number of components.

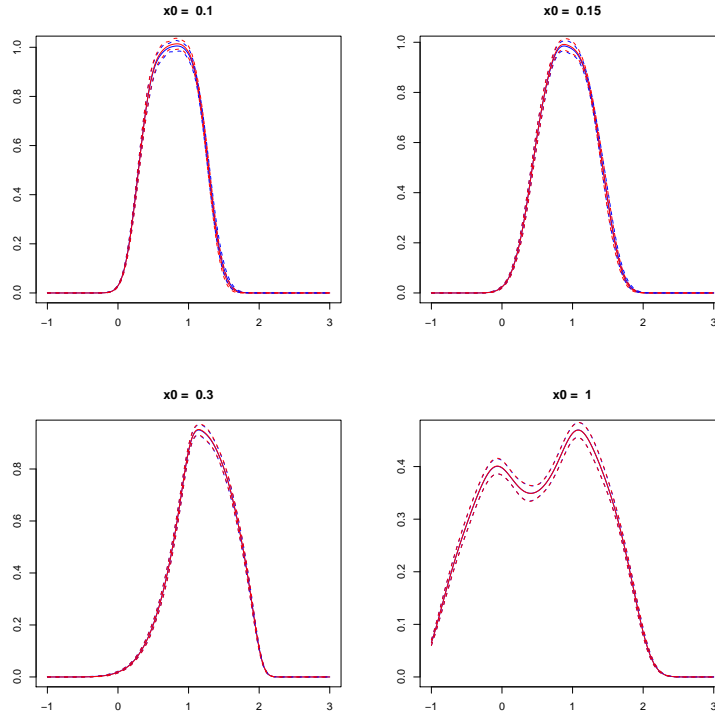


FIGURE 5. Toy example 1: curves in red (resp. blue) plain line represents the estimation of a density for a fixed  $x_0$  from the first (resp. second) method and dashed red (resp. blue) line represent the standard deviation of the estimation obtained from 25 independent design of experiments. The blue and the red lines are superposed.

Figure 8 represents the  $L^2$  relative error of MMP (in red) with  $N = 200$  and the relative error of twenty independent basis (in black) containing probability functions chosen randomly in the learning sample  $f_1, \dots, f_N$ . The MMP outperforms the random strategies. This result validates the way to choose a new curve among the sample of probability density functions, as this choice is better than randomly picking functions in the sample.

## 2.2. A five-dimensional example (Toy example 2)

Let  $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5) \in \mathcal{X} = [0, 1]^5$ . A second toy example is defined as follows:

$$G(\mathbf{x}, N, U_1, U_2, B) = (x_1 + 2x_2 + U_1) \sin(3x_3 - 4x_4 + N) + U_2 + 10x_5 B + \sum_{i=1}^5 ix_i$$

where

$$\begin{aligned} N &\sim \mathcal{N}(0, 1) \\ U_1 &\sim \mathcal{U}([0, 1]) \\ U_2 &\sim \mathcal{U}([1, 2]) \\ B &\sim \text{Bern}(1/2). \end{aligned}$$

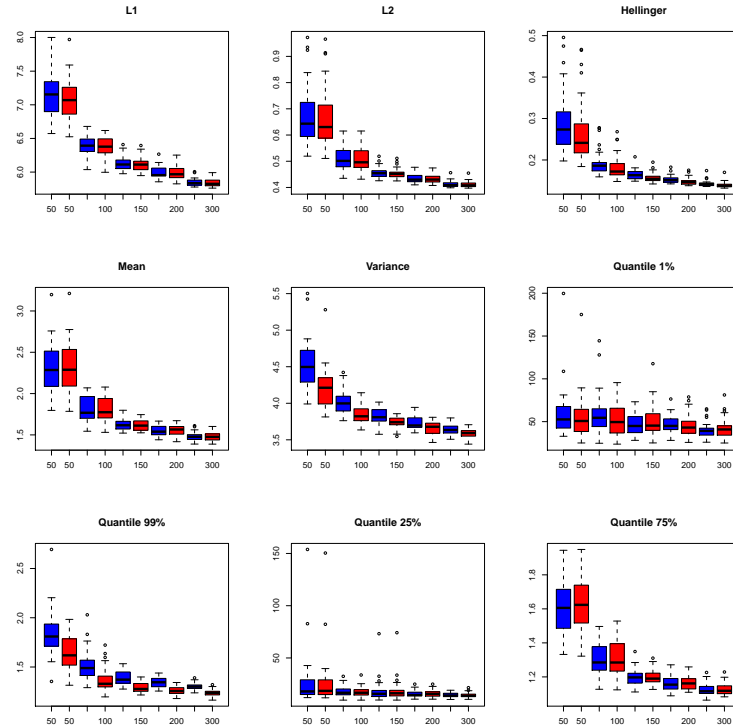


FIGURE 6. Toy example 1: boxplot of the errors for different sizes of  $N$ . Blue: estimator given by  $L^2$  norm. Red: estimator given by the Hellinger distance. Results have been averaged with 25 independent experiments. Blue: estimator given by  $L^2$  norm. Red: estimator given by the Hellinger distance.

Figure 9 represents the output for  $N = 100$  points simulated uniformly and independently in  $\mathcal{X}$ .

### 2.2.1. Kernel regression

In this section, the kernel regression is tested on this second toy example with isotropic (Figure 11) and anisotropic bandwidth (Figure 12). As in the first example, Figure 10 represents for different values of  $x_0$  the true probability density function (blue dashed line) and the estimation obtained from the two estimators of the kernel regression (red plain line and orange dots line). In this example, the two estimators give different results. In red is represented the relative error for the estimator based on Hellinger distance, in blue for the estimator based on  $L^2$  norm. The Hellinger estimator gives a better approximation of the quantities presented in Figure 11, except for the mean. Contrary to the case of the first toy example, the errors are quite high. The variance of the error on the norms, presented in Figure 11, is low but does not seem to decrease steadily. The use of an anisotropic bandwidth does not seem to improve the quality of the estimation, while it is much longer to compute. The errors are approximately the same with isotropic or anisotropic bandwidth. The greater precision brought by the anisotropic bandwidth is compensated by the difficulty to estimate it.

### 2.2.2. Functional decomposition methods

In this section, the four functional decompositions MMP with  $L^2$  and Hellinger distances, AQM and CPCA are compared. The decompositions are applied to learning samples whose sizes are  $N = 50$  and  $100$  and for decomposition basis sizes ranging from 1 to 20. The relative errors between the learning sample functions and

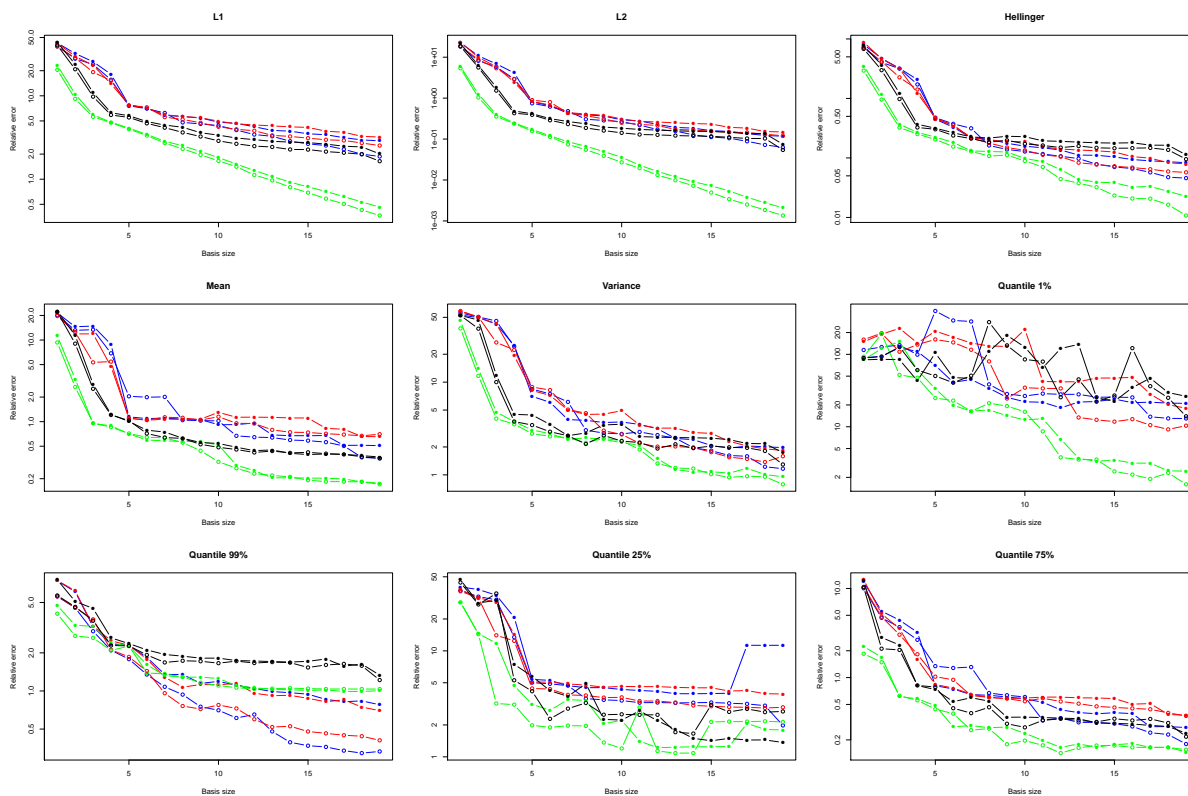


FIGURE 7. Toy example 1: comparison of the relative error for different quantities in function of the size of the basis  $q$  with a design of experiments of size  $N = 50$  (circles) and  $100$  (filled circles). Blue: MMP decomposition given by  $L^2$  norm. Red: MMP decomposition given by the Hellinger distance. Black: AQM decomposition. Green: CPCA decomposition.

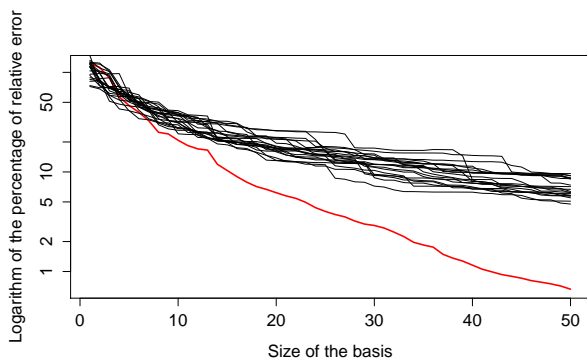


FIGURE 8. Toy example 1: comparison of the  $L^2$  relative error obtained with MMP (in red) and a random choice in the design of experiments (in black).

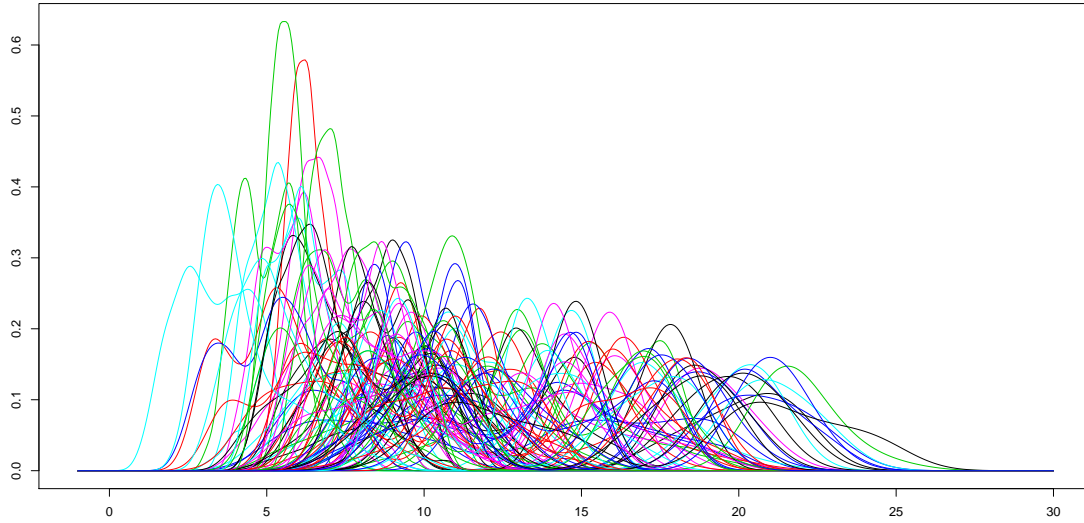


FIGURE 9. Toy example 2: representations of outputs for  $N = 100$  different parameters.

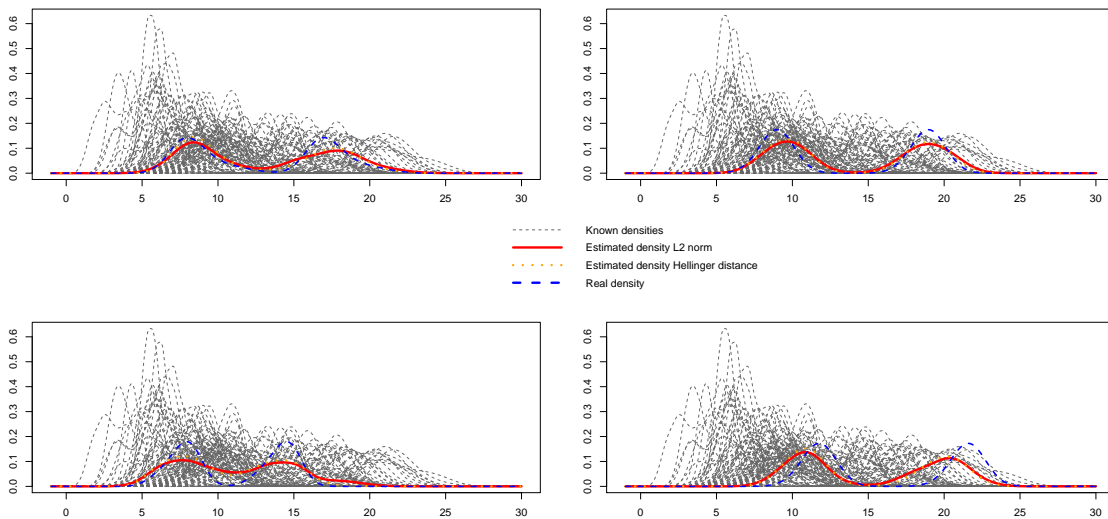


FIGURE 10. Toy example 2: curves in dashed gray lines are probability density functions, red plain line (resp. orange dots line) represents the estimation of a density for a fixed  $\mathbf{x}_0$  from the first (resp. second) method and dashed blue line is the true probability density function. These estimations were made with  $h$  isotropic. The orange dotted line and the red line are superposed.

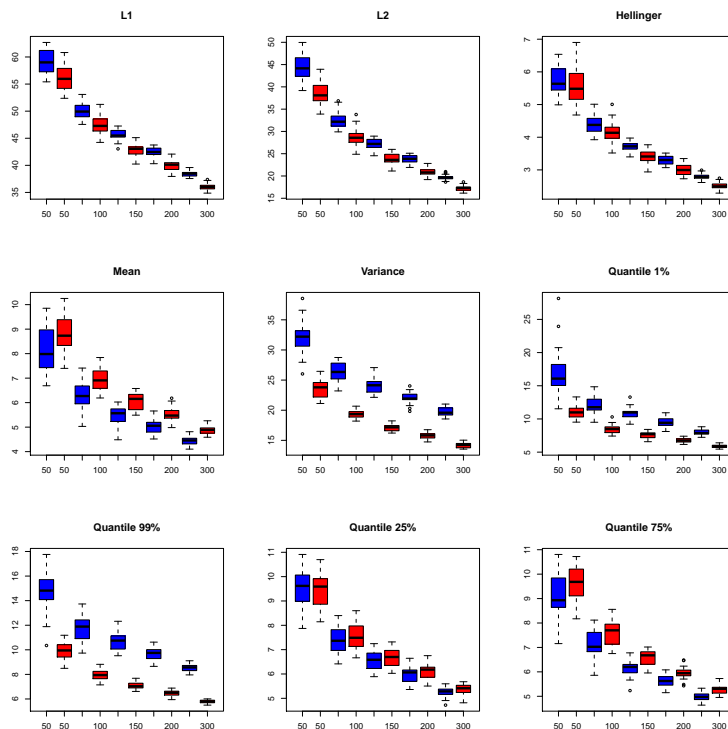


FIGURE 11. Toy example 2: boxplot of the errors for different sizes of  $N$ . Left: estimator given by  $L^2$  norm. Right: estimator given by the Hellinger distance. Results have been averaged with 25 independent experiments. These estimations were made with  $h$  isotropic. Blue: estimator given by  $L^2$  norm. Red: estimator given by the Hellinger distance.

their approximations are computed for the 9 quantities of interest. Figure 13 represents, in logarithmic scale, these relative errors in function of the basis size for the three decompositions. First, the relative errors are low for all quantities of interest, as in the previous example. The relative errors decrease less even than in the one-dimensional case. The behavior of the errors on 1% and 99% quantiles is particularly unsteady. CPCA method gives better results than the other methods for  $L^1$ ,  $L^2$  norms, 25% and 75% quantiles. However, for all studied quantities, the AQM and CPCA method seem to give quite equivalent results. MMP method errors are slightly higher than AQM errors for most of the quantities.

### 3. INDUSTRIAL APPLICATIONS

In this section we apply the methods we developed in Section 1 to three industrial numerical codes of CEA, EDF, and IFPEN. We compute the same relative errors as for the toy examples of Section 2.

#### 3.1. CEA application: CASTEM test case

##### 3.1.1. Code description

In the framework of nuclear plant risk assessment studies, the evaluation of component reliability during accidental conditions is a major issue required for the safety case. Thermal-hydraulic (T-H) and thermal-mechanic (T-M) system codes model the behaviour of the considered component subjected to highly hypothetical accidental conditions. In the study that we consider here, the T-M code CASTEM takes as input 13 uncertain

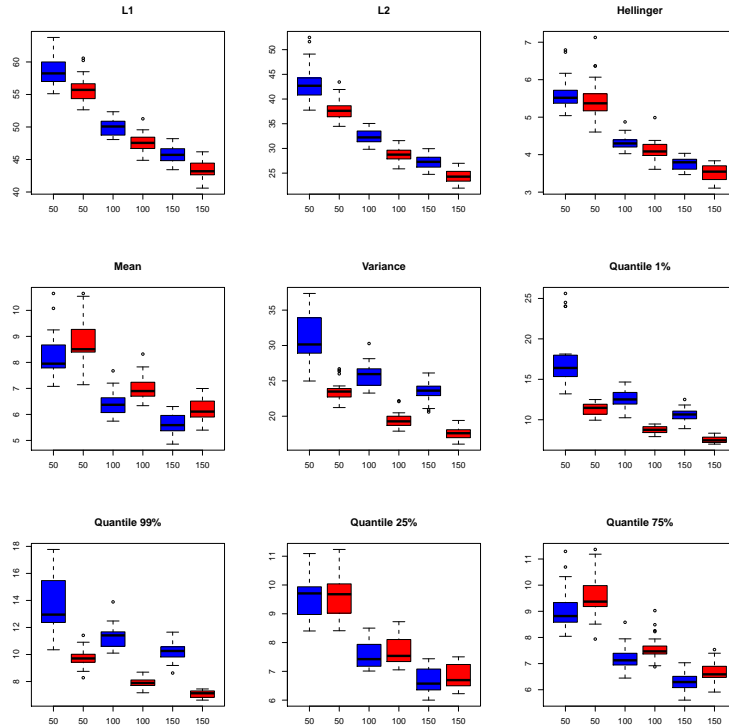


FIGURE 12. Toy example 2: boxplot of the errors for different sizes of  $N$ . Left: estimator given by  $L^2$  norm. Right: estimator given by the Hellinger distance. Results have been averaged with 25 independent experiments. These estimations were made with  $h$  anisotropic. Blue: estimator given by  $L^2$  norm. Red: estimator given by the Hellinger distance.

parameters, related to the initial plant conditions or to the safety system characteristics. Three of them are functional T-H parameters which depend on time: fluid temperature, flow rate and pressure. The other ten parameters are T-M scalar variables. For each set of parameters, CASTEM calculates the absolute mechanical strength of the component and the thermo-mechanical actual applied load. From these two elements, a safety margin (SM) is deduced.

The objective is to assess how these uncertain parameters can affect the code forecasts and more specifically the predicted safety margin. However, CASTEM code is too time expensive to be directly used to conduct uncertainty propagation studies or global sensitivity analysis based on sampling methods. To avoid the problem of huge calculation time, it can be useful to replace CASTEM code by a metamodel. One way to fit a metamodel on CASTEM could be to discretize the functional inputs and to consider the values of the discretization as scalar inputs of CASTEM code. Nevertheless, this solution is often intractable due to the high number of points in the discretization. To cope with this problem, in [13, 17] a method was proposed to treat implicitly these “uncontrollable” parameters functional parameters, while the other ten scalar parameters are considered as “controllable”. CASTEM output is then a random variable conditionally to “controllable” parameters.

A latin hypercube sampling method [18] is used to build a learning sample of 500 points in dimension 10. For each set of controllable parameters, CASTEM has been run 400 times with different uncontrollable functional parameters, randomly chosen in an available database. The probability density function  $f_i$  ( $i = 1, \dots, 500$ ) of the safety margin is computed by kernel estimation with the 400 outputs of CASTEM for each set of parameters. A few examples of the obtained probability density functions are represented on Figure 14. In the following,

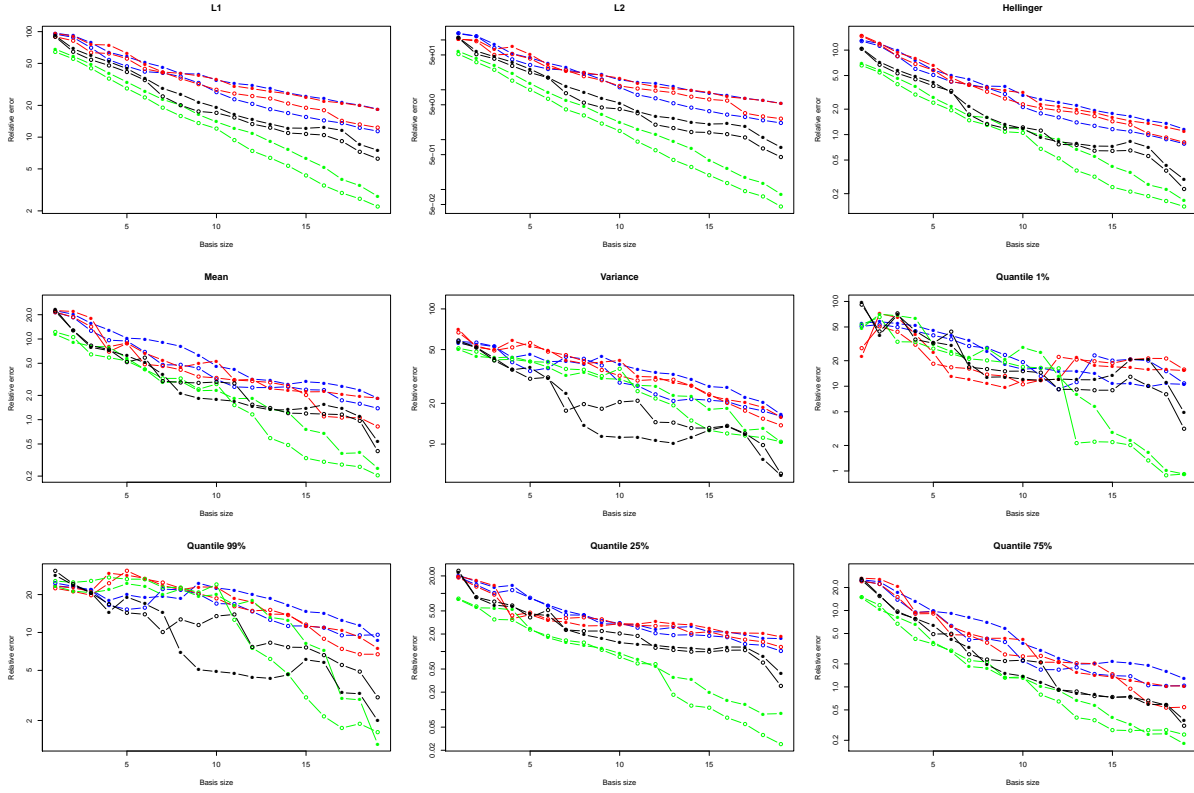


FIGURE 13. Toy example 2: comparison of the relative error for different quantities in function of the size of the basis  $q$  with a design of experiments of size  $N = 50$  (circles) and  $100$  (filled circles). Blue: MMP decomposition given by  $L^2$  norm. Red: MMP decomposition given by the Hellinger distance. Black: AQM decomposition. Green: CPCA decomposition.

the two kernel regression metamodells, then MMP, AQM and CPCA decomposition methods are applied on CASTEM test case.

### 3.1.2. Kernel regression

In this section, the kernel regression method is applied with isotropic and anisotropic bandwidths. Figure 15 represents for four different values of  $x_0$  the true probability density function (blue dashed line) and the estimation obtained from the two estimators of the kernel regression ( $L^2$  estimator in red plain line and Hellinger estimator in orange dotted line), with isotropic bandwidth. One can see that the estimation by the  $L^2$  estimator of the four probability density functions is very far from the real function. In particular, at the bottom left of the figure, the mean of the predicted probability function in red is very far from the mean of the real one in dashed blue. On these four probability density functions, the Hellinger estimator gives much better results than the  $L^2$  one.

To verify this first graphical analysis, the Leave-One-Out method has been used to assess the efficiency of the kernel regression-based metamodel. The bandwidth of the kernel regression is estimated thanks to all probability density functions except the  $i^{\text{th}}$  one. The function  $f_i$  is estimated with the corresponding metamodel and the relative errors on the quantities of interest are computed between  $f_i$  and its estimation. This process is repeated for each probability density function in the dataset and all the computed relative errors are averaged. The mean relative errors of the estimators based on Hellinger distance and  $L^2$  norm are given in Table 1 and



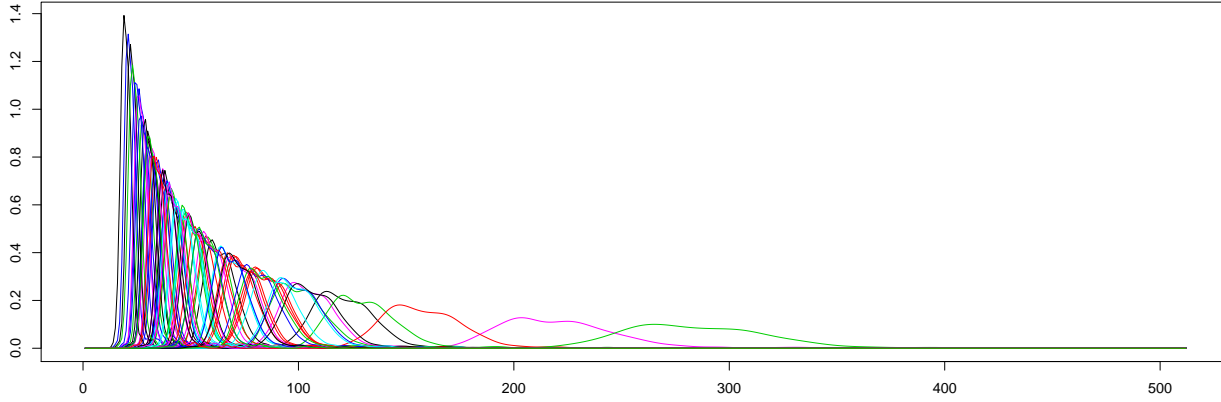


FIGURE 14. CASTEM test case: representations of outputs for  $N = 75$  different parameters.

	Hellinger distance estimator	$L^2$ norm estimator
$L^1$ norm	114.2%	114.9%
$L^2$ norm	167.7%	171.9%
Hellinger distance	17.0%	17.0%
Mean	22.2%	22.4%
Variance	86.1%	86.9%
1% quantile	72.0%	74.3%
99% quantile	39.7%	40.9%
25% quantile	30.9%	31.5%
75% quantile	23.7%	23.9%

TABLE 1. CASTEM test case: the mean relative errors on the quantities of interest computed by the Leave-One-Out method for the kernel regression estimators based on the Hellinger distance and  $L^2$  norm with an isotropic bandwidth.

Table 2 for the different quantities of interest respectively for an isotropic and anisotropic bandwidths. The two estimators perform poorly on this dataset. The errors for  $L^1$  and  $L^2$  norms are particularly high. The Leave-One-Out validation confirms that the kernel regression method is not adapted to CASTEM test case. This can be explained by the important influence of the controllable T-M parameters compared to the one of the uncontrollable T-H parameters.

### 3.1.3. Functional decomposition methods

In this section, the four functional decompositions MMP with  $L^2$  and Hellinger distances, AQM and CPCA are compared. Figure 16 represents, in logarithmic scale, the relative errors on the different norms, modes and quantiles versus the basis size (from 1 to 20). First, the relative errors are low for all quantities of interest, except for the variance. For the variance, the errors are over 10% with a decomposition basis with 20 functions. The errors for small bases are high but decrease very quickly. The decrease is especially quick and even for errors on  $L^1$ ,  $L^2$  and Hellinger distances. For modes and quantiles, the errors of the four methods do not decrease steadily for all quantities of interest, and especially for small basis sizes. CPCA clearly outperforms other methods for the  $L^1$ ,  $L^2$  and Hellinger distances. For other quantities of interest, it gives good results. For

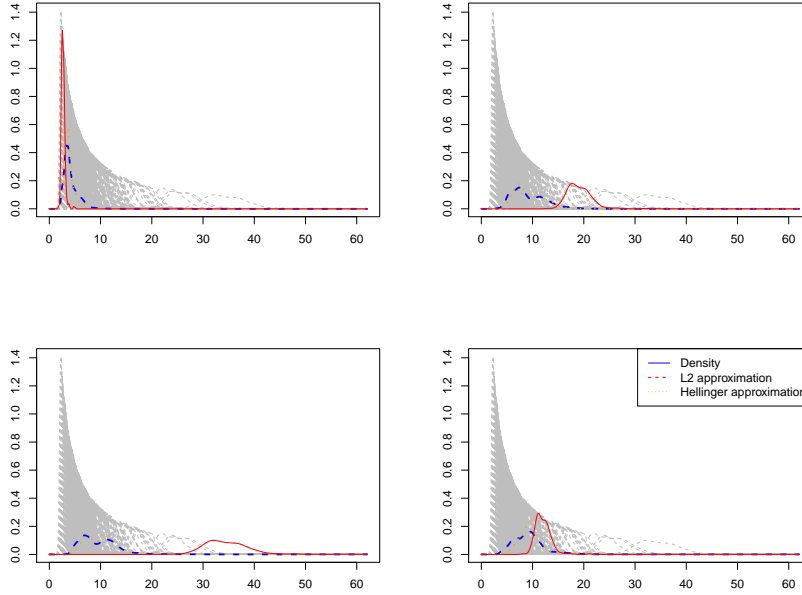


FIGURE 15. CASTEM test case: curves in dashed gray lines are probability density functions, red plain line represents the estimation of a density for a fixed  $x_0$  and dashed blue line is the true probability density function.

	Hellinger distance estimator	$L^2$ norm estimator
$L^1$ norm	94.3%	94.5%
$L^2$ norm	108.0%	108.6%
Hellinger distance	14.3%	14.3%
Mean	15.4%	15.6%
Variance	71.5%	71.9%
1% quantile	42.7%	43.3%
99% quantile	23.0%	23.8%
25% quantile	20.4%	20.4%
75% quantile	15.9%	16.3%

TABLE 2. CASTEM test case: the mean relative errors on the quantities of interest computed by the Leave-One-Out method for the kernel regression estimators based on the Hellinger distance and  $L^2$  norm with an anisotropic bandwidth.

the variance, 1% and 99% quantiles, AQM method performs poorly compared to the three others. The errors of AQM and CPCA methods seem more stable than for MMP method for most of the quantities of interest, especially for modes and quantiles. Overall, the results are quite good for the four methods.

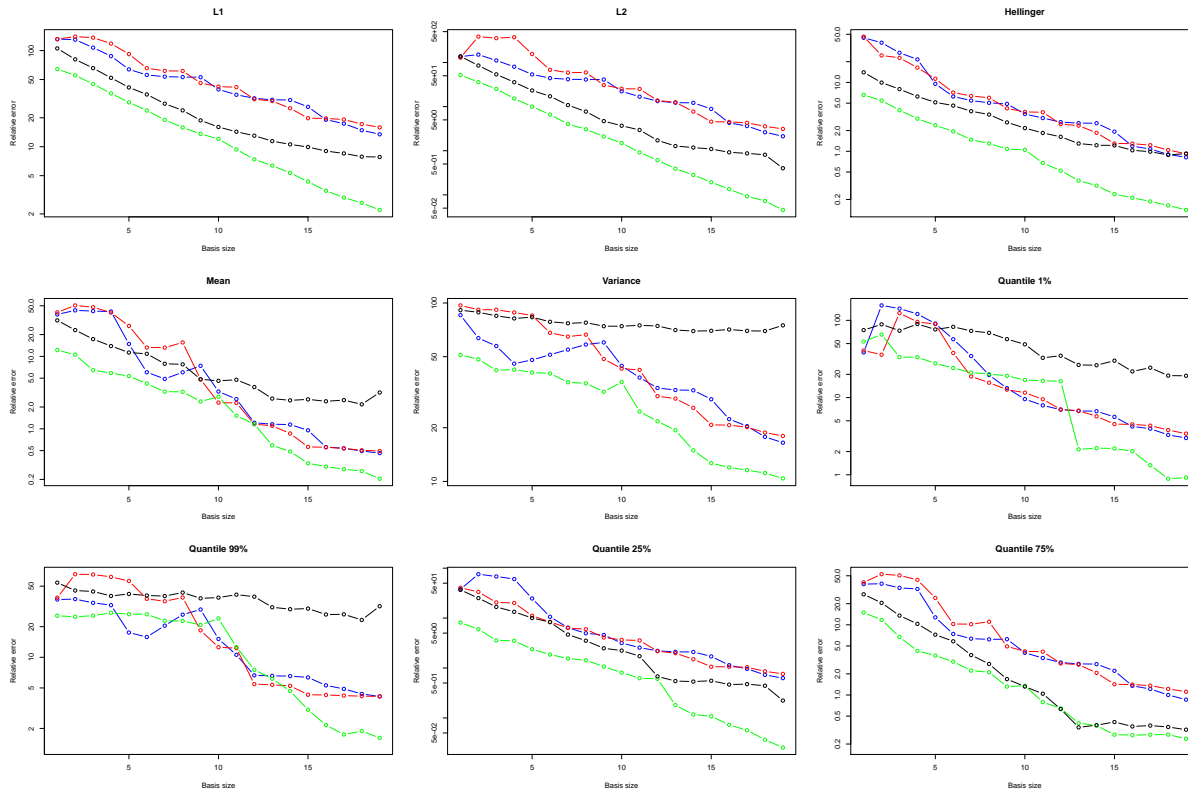


FIGURE 16. CASTEM test case: comparison of the relative error for the 9 quantities in function of the size of the basis  $q$ . Blue: MMP decomposition given by  $L^2$  norm. Red: MMP decomposition given by the Hellinger distance. Black: AQM decomposition. Green: CPCA decomposition.

### 3.2. EDF application: VME test case

#### 3.2.1. Code description

To optimize the whole life cost of its nuclear fleet, EDF has developed an asset management methodology [9]. A part of this methodology deals with exceptional maintenance tasks strategies. To help the decision maker to choose the best strategy (how many times do we need to carry out exceptional tasks, when,...?), EDF has developed a dedicated tool called VME (described in Figure 17) based on Monte-Carlo simulation to compute many technical economic indicators among which the density function of the Net Present Value (called VAN for “Valeur Actuelle Nette”) is the most relevant.

This tool leads to an important simulation time and requires an important amount of input data that are surrounded with uncertainties:

- Reliability data: generally there is not enough (or sometimes not any) feedback data to precisely evaluate reliability model parameters;
- Economic data: economic indicators and duration of maintenance tasks remain on several hypothesis that can be modified;
- Other data: uncertainty on operating times of power plants, maintenance tasks dates, etc.

In our work, we study a particular scenario simulation with VME concerning the life cycle management of four large transformers, adapted from the EPRI study made for CENG [15]. The uncertainty on the inputs are

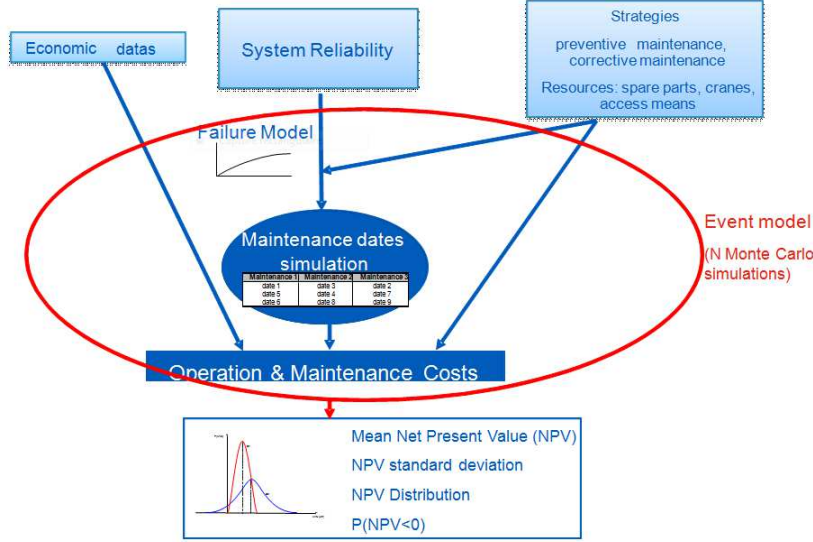


FIGURE 17. VME: EDF valorization of maintenance strategies tool description.

Uncertainty inputs	distribution	shape	scale	location
$W_1$	Weibull	1/0.021	2.6	-1
$W_2$	Weibull	100	3.8	-20
$W_3$	Weibull	100	4	-24
$W_4$	Weibull	100	3.8	-20

TABLE 3. VME: uncertainty inputs used in the model. Each distribution depends of three parameters: the shape, the scale and the location.

represented by four Weibull distributions which are summarized in Table 3. Define for  $i = 1, \dots, 4$ ,

$$D_i = \{\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 : \mathbb{E}[\text{Weibull}(x_1, x_2, x_3)] \in [\mathbb{E}[W_i] \pm 0.2\mathbb{E}[W_i]]\},$$

and denote  $sh_i, sc_i, loc_i$  respectively the shape, the scale and the location of  $W_i$ . For  $i = 1, \dots, 4$  let  $x_{i,1}, x_{i,2}, x_{i,3}, x_{i,4}$  be parameters simulated uniformly on  $D_i$ . Note that  $W_2$  and  $W_4$  follow the same law but  $sh_2, sc_2, loc_2$  and  $sh_4, sc_4, loc_4$  are considered as different parameters. That means there are three parameters for each Weibull distribution. Moreover, there are four other parameters which represent some costs and two other parameters which represent durations of services. All these 18 parameters are summarized in Table 4.

In this case, we are particularly interested by the VAN indicator. It is defined by

$$VAN = \mathbf{P}(G(\mathbf{x}, \mathbf{W}) \leq 0),$$

where  $\mathbf{W} = (W_1, \dots, W_4)$  and  $\mathbf{x}$  the eighteen fixed parameters.

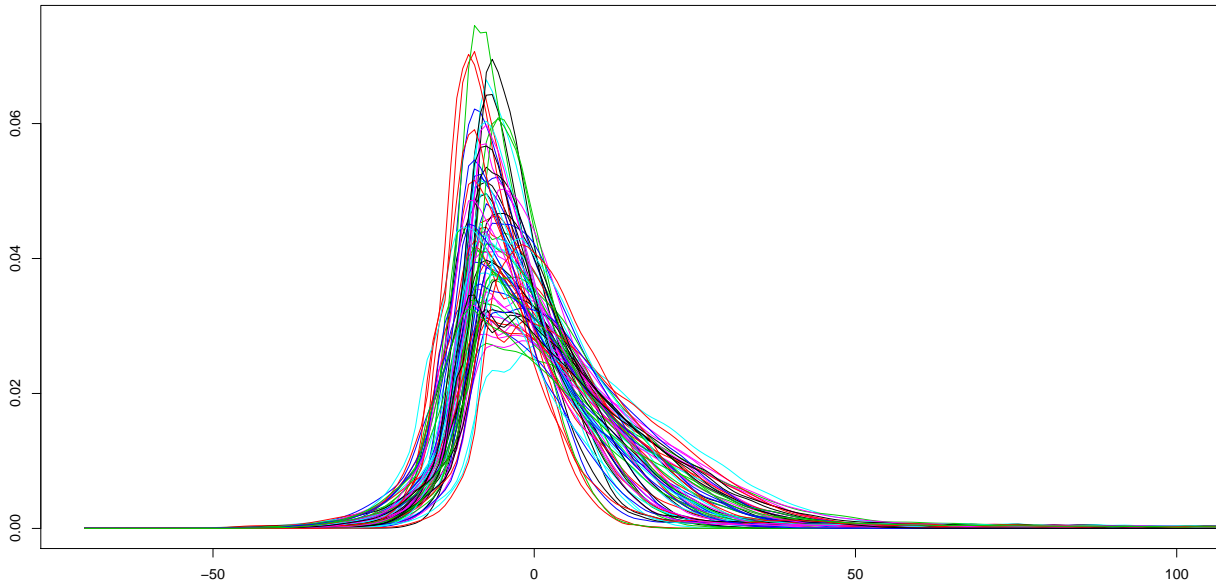
Figure 18 represents 75 outputs from the VME model. The kernel regression is applied to designs of experiments of increasing sizes such that smaller designs are included in greater ones. A test sample of 500 points is used to assess the efficiency of the regression.

### 3.2.2. Kernel regression

Figure 19 represents for different values of  $x_0$  the true probability density function (blue dashed line) and the estimation obtained from the two estimators of the kernel regression (red plain line and orange dots line).

Uncertainty inputs	distribution	support
$(sh_1, sc_1, loc_1)$	<i>Uniform</i>	$D_1$
$(sh_2, sc_2, loc_2)$	<i>Uniform</i>	$D_2$
$(sh_3, sc_3, loc_3)$	<i>Uniform</i>	$D_3$
$(sh_4, sc_4, loc_4)$	<i>Uniform</i>	$D_4$
Cost1	<i>Uniform</i>	[796 , 1194]
Cost2	<i>Uniform</i>	[49.6 , 74.4]
Cost3	<i>Uniform</i>	[5.6 , 8.4 ]
Cost4	<i>Uniform</i>	[4 , 8 ]
Duration1	<i>Uniform</i>	[0.8 , 1.2 ]
Duration2	<i>Uniform</i>	[0.8 , 1.2]

TABLE 4. VME: distribution of the eighteen parameters.

FIGURE 18. VME: representations of outputs for  $N = 75$  different parameters.

In this section, the kernel regression is tested on this real case study with isotropic and anisotropic bandwidth (Figure 20).

In red is represented the relative error for the estimator based on the Hellinger distance, in blue for the estimator based on the  $L^2$  norm. The estimation seems equivalent for the two estimators. The results are equivalent for anisotropic and isotropic bandwidth. For the anisotropic bandwidth, the estimation is less stable and the estimation of  $H$  becomes much more time-consuming.

The relative error in norm is smaller for VME (with isotropic and anisotropic bandwidth) than for the toy example 2, whereas the dimension is higher. That shows that the quality of estimation depends not only on the dimension but also on the regularity of the model, if the bandwidth is sufficiently well estimated.

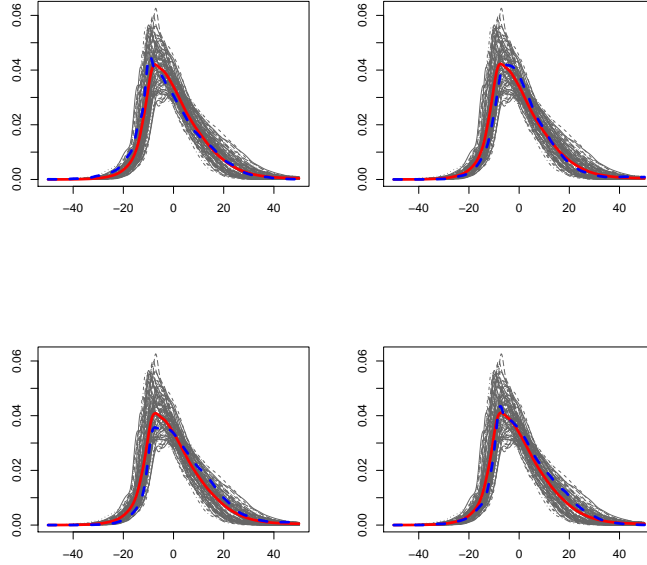


FIGURE 19. VME: curves in dashed gray lines are probability density functions, red plain line represents the estimation of a density for a fixed  $x_0$  and dashed blue line is the true probability density function for  $N = 100$ .

Bad results are found for the variance, 99% and 99.9% quantiles. The relative error for the mean is also very high. The real mean is quite close to zero, so that in the relative error computation, the numerator is divided by a quantity close to zero.

Figure 21 represents the estimation of the  $\mathbf{P}(VAN < 0)$  for isotropic and anisotropic bandwidths. In this case, choosing an anisotropic bandwidth does not bring information.

### 3.2.3. Functional decomposition methods

In this section, the four functional decompositions MMP with  $L^2$  and Hellinger distances, AQM and CPCA are compared. Figure 22 represents, in logarithmic scale, the relative errors on the different norms, modes and quantiles versus the basis size (from 1 to 20). First, the relative errors are low for all quantities of interest, except for the mean and variance. For the variance, the errors are over 20% with a decomposition basis with 20 functions. The very high errors on the mean can be explained by the fact that the mean to be estimated is close to zero for most of densities in the VME case. The errors on mean, variance have moreover a very chaotic behavior. The errors on distances decrease very quickly. For modes and quantiles, the errors of the four methods do not decrease steadily for all quantities of interest. CPCA outperforms other methods for the  $L^1$ ,  $L^2$ , Hellinger distances, the variance and the 99% and 25% quantiles. The performances of AQM and MMP are very close for modes and quantiles. CPCA seems more stable than MMP and AQM on most of the quantities of interest.

In Figure 23, the errors on the probability for the VAN to be negative  $P(VAN < 0)$  for the different decomposition methods are represented in function of the basis size. The errors are quite low and decrease quickly. The CPCA method outperforms the others for both learning sets sizes. The other three methods give similar results.

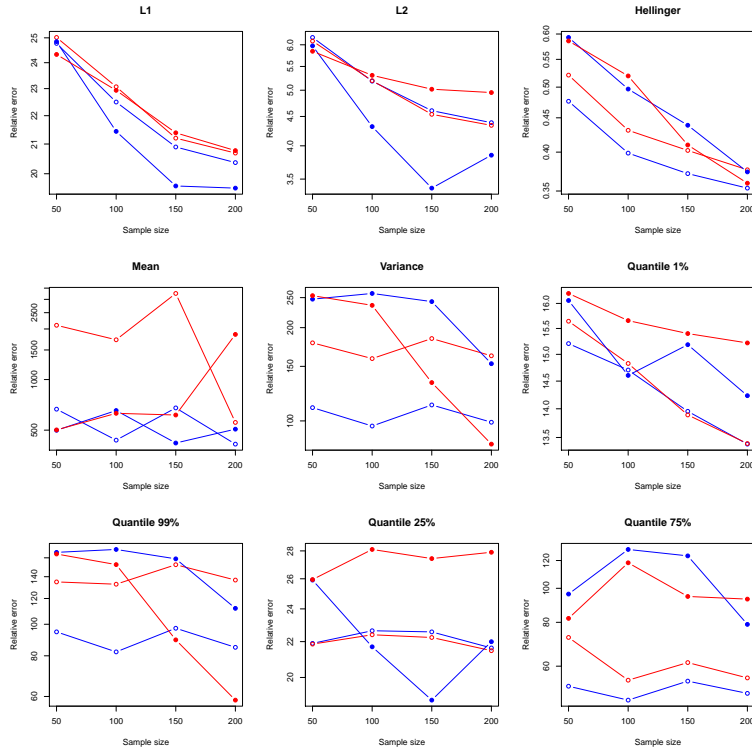


FIGURE 20. VME: comparison of the relative error for the two estimators with an isotropic and anisotropic (plain circle) bandwidth. Blue: estimators given by the  $L^2$  norms. Red : estimator given by the Hellinger distance

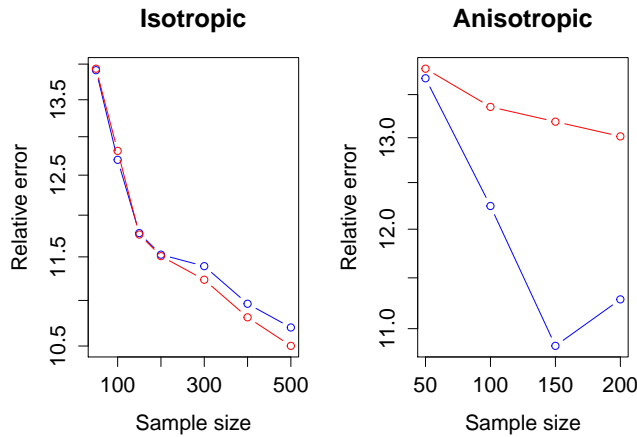


FIGURE 21. VME: comparison of the relative error for the  $\mathbf{P}(VAN < 0)$ . Left: isotropic bandwidth. Right: anisotropic bandwidth. Blue: estimators given by the  $L^2$  norms. Red : estimator given by the Hellinger distance

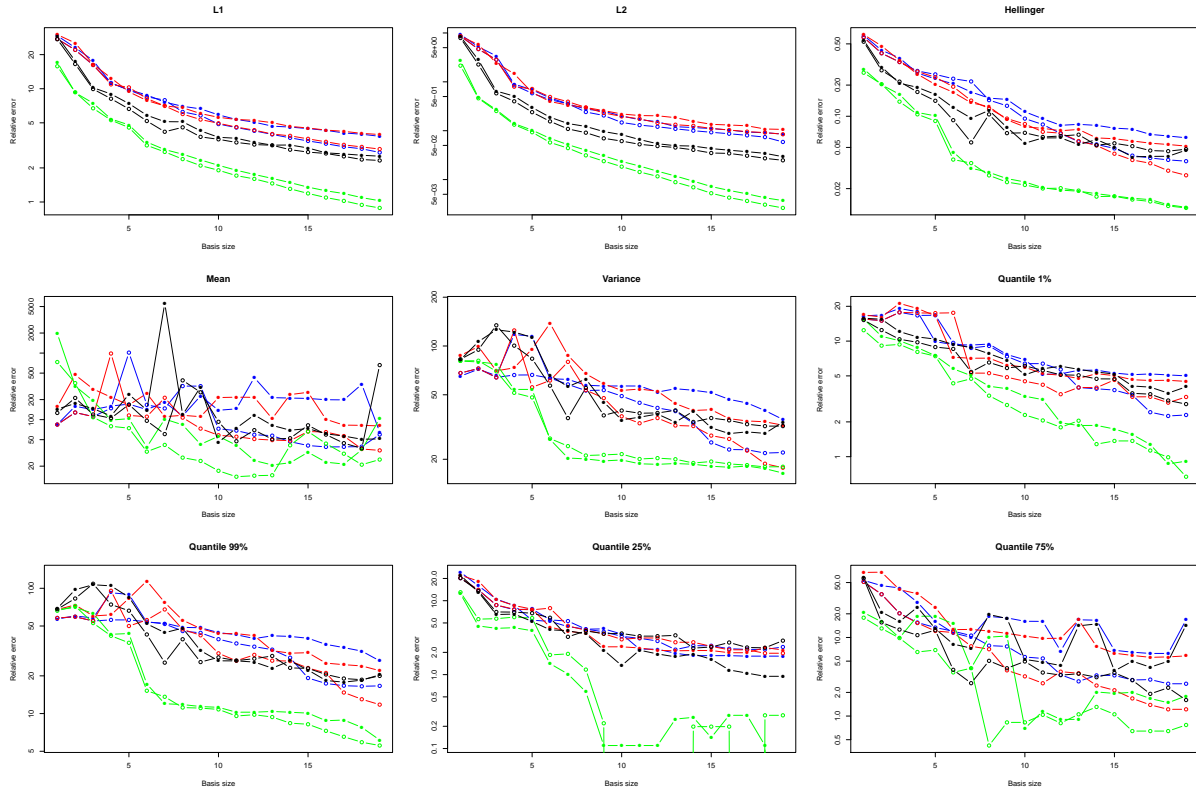


FIGURE 22. VME: comparison of the relative error for the 9 quantities in function of the size of the basis  $q$ . Blue: MMP decomposition given by  $L^2$  norm with a learning set size of  $N = 50$  (circles) and 100 (filled circles). Red: MMP decomposition given by the Hellinger distance. Black: AQM decomposition. Green: CPCA decomposition.

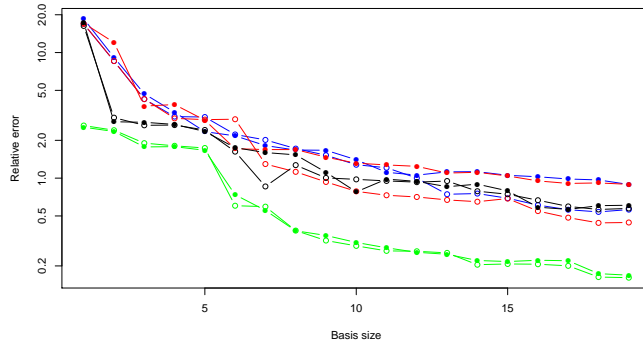


FIGURE 23. VME: comparison of the relative error for  $P(VAN < 0)$  in function of the size of the basis  $q$ . Blue: MMP decomposition given by  $L^2$  norm with a learning set size of  $N = 50$  (circles) and 100 (filled circles). Red: MMP decomposition given by the Hellinger distance. Black: AQM decomposition. Green: CPCA decomposition.



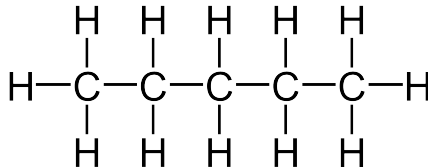


FIGURE 24. The structural formula of pentane. (Wikipedia)

### 3.3. IFPEN application: GIBBS test case

#### 3.3.1. Molecular Modelling

At IFP Energies Nouvelles, the researchers of the Department of Thermodynamics and Molecular Modelling elaborate models for the structure of molecules such as hydrocarbons and alcohols. In the following numerical study we consider a system of *pentane* molecules at a *reduced temperature* of 0.75. Pentane is an *alkane* with five carbon atoms ( $C_5H_{12}$ , cf. Figure 24.) Here we use the Anisotropic United Atoms model (AUA) [29] to represent the molecular structure of pentane. The carbon (C) and hydrogen (H) atoms in each of the two terminal *methyls* ( $CH_3$ ) and each of the three *methylene* bridges ( $CH_2$ ) are treated as a single *interaction center*. To each of these two types of interaction centers correspond three parameters of the model: an energy parameter ( $\varepsilon_{CH_3}$  and  $\varepsilon_{CH_2}$ ), a size parameter ( $\sigma_{CH_3}$  and  $\sigma_{CH_2}$ ), and a displacement parameter ( $\delta_{CH_3}$  and  $\delta_{CH_2}$ ). These six parameters will be the components of the set of input parameters of the numerical codes under consideration:

$$\mathbf{x} = (\varepsilon_{CH_3}, \varepsilon_{CH_2}, \sigma_{CH_3}, \sigma_{CH_2}, \delta_{CH_3}, \delta_{CH_2}) \in \mathbb{R}^6.$$

Here the matter of interest is the prediction of two macroscopic properties of a pentane system in *chemical equilibrium*: the *volumetric mass density*  $\rho_{liq}^{eq}$  of the liquid phase and the *vapor pressure*  $P_{gas}^{eq}$  (pressure of the gas phase). The algorithm carried out to numerically predict  $\rho_{liq}^{eq}$  and  $P_{gas}^{eq}$  – for a given set of parameters  $\mathbf{x}$  – relies on the *fundamental postulate of statistical mechanics*: an isolated system in equilibrium is found with equal probability in each of its accessible *microstates*. Thus,  $\rho_{liq}^{eq}$  (respectively  $P_{gas}^{eq}$ ) can be approximated by the average of the values of the volumetric mass density of the liquid phase  $\rho_{liq}$  (respectively the vapor pressure  $P_{gas}$ ) of a large number of accessible microstates of the system. Therefore the algorithm is essentially a loop, each step consisting of the following two instructions:

- (1) randomly generate an accessible microstate of the system (characterized by the positions of the molecules and the volume of the system),
- (2) compute the value of  $\rho_{liq}$  (respectively  $P_{gas}$ ) given by the model for this particular microstate.

Let us denote  $G_{\rho_{liq}}(\mathbf{x})$  (respectively  $G_{P_{gas}}(\mathbf{x})$ ) the probability density function of the random value  $\rho_{liq}$  (respectively  $P_{gas}$ ) computed at each loop step. The two instructions above are iterated a hundred million times. After the execution of the loop, the arithmetical mean of these millions of values is computed – it is a Monte Carlo approximation of the integral of  $G_{\rho_{liq}}(\mathbf{x})$  (respectively  $G_{P_{gas}}(\mathbf{x})$ ) – and is returned by the algorithm as the prediction of  $\rho_{liq}$  (respectively  $P_{gas}$ .) The whole algorithm is carried out in about twenty-four hours on a supercomputer (for each set  $\mathbf{x}$  of input parameters.)

In order to circumvent this time-consuming process, we want to build a metamodel of the numerical codes mapping the input parameters  $\mathbf{x}$  to the probability density functions  $G_{P_{gas}}(\mathbf{x})$  and  $G_{\rho_{liq}}(\mathbf{x})$  of the random outputs. We had at our disposal a sample of size  $N = 50$ . Figure 25 and Figure 26 show the corresponding 50 probability functions for volumetric mass density and vapor pressure respectively.

#### 3.3.2. Kernel regression

We applied the kernel regression method on the numerical codes providing the volumetric mass density of the liquid phase and the vapor pressure. The results provided by the  $L^2$  estimator with isotropic bandwidth are provided in Table 5 for the volumetric mass density and in Table 6 for the vapor pressure. The errors have

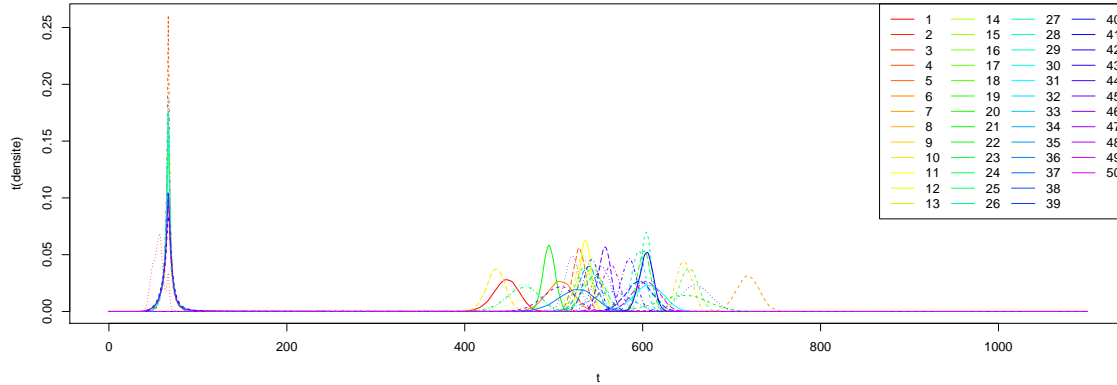


FIGURE 25. GIBBS test case: distributions of the volumetric mass density for different sets of input parameters.

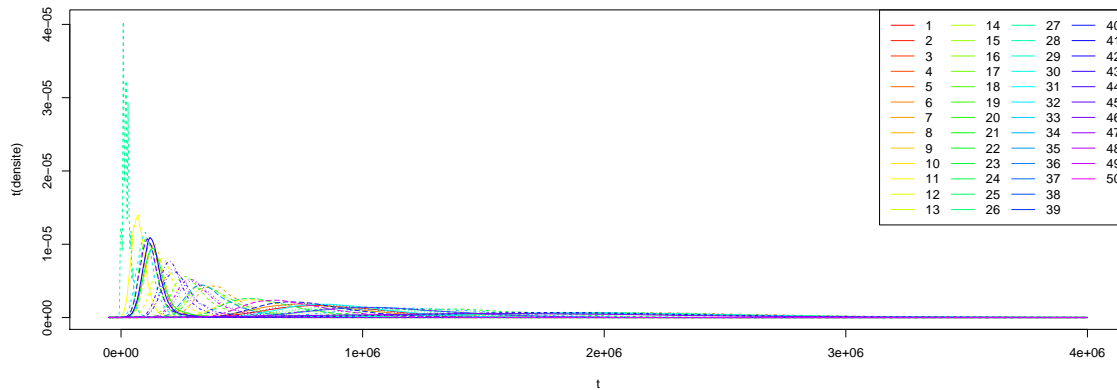


FIGURE 26. GIBBS test case: distributions of the vapor pressure for different sets of input parameters.

been computed by the Leave-One-Out methodology as in Section 3.1.2. We can see that most of the results are disappointing: all relative errors are higher than 30%, excepted – surprisingly – the relative Hellinger error of the approximation. The latter amounts to nearly 7% for the volumetric mass density, and only nearly 0.05% for the vapor pressure. The errors for quantile 1% and quantile 25% are particularly high. The Hellinger estimator give almost the same results: the difference between the error for the  $L^2$  estimator and the error for the Hellinger estimator is lower than  $10^{-5}$  in every case. That is why they are not charted here.

In Figure 27 and Figure 28 we plot the probability density functions of the volumetric mass density and the vapor pressure with their kernel regression approximations, for arbitrarily chosen sets of input parameters. We can see how poorly the densities are approximated.

L1	L2	Hellinger
151.1%	375.9%	7.089%
Mean	Variance	Quantile 1%
50.77%	98.07%	513.8%
Quantile 99%	Quantile 25%	Quantile 75%
30.44%	198.4%	41.52%

TABLE 5. GIBBS test case: kernel regression relative errors of the  $L^2$  estimator for the volumetric mass density.

L1	L2	Hellinger
123.7%	254.1%	0.04878%
Mean	Variance	Quantile 1%
69.86%	80.72%	667.63%
Quantile 99%	Quantile 25%	Quantile 75%
74.82%	156.1%	73.5%

TABLE 6. GIBBS test case: KR relative errors of the  $L^2$  estimator for the vapor pressure.

### 3.3.3. Functional decomposition methods

In this section, the four functional decompositions MMP with  $L^2$  and Hellinger distances, AQM and CPCA are compared. Figure 29 represents, in logarithmic scale, the relative errors on the different norms, modes and quantiles versus the basis size (from 1 to 20) for the volumetric mass density of the liquid phase. The error curves are globally decreasing excepted the one corresponding to the variance relative error for AQM. Moreover, the relative error on variance is high even for basis with 20 functions and especially for AQM. The behavior of the AQM approximation is more erratic for the 99% quantile. Both MMP methods perform worse on the 25% and 75% quantiles than other methods. Other relative errors are low. CPCA gives better results for  $L^1$ ,  $L^2$ , Hellinger distances and the 25% quantile. For 25% and 75% quantiles, errors on both MMP bases are higher than for others methods.

The results for the vapor pressure are presented in Figure 30. The errors on all quantities of interest are more steady than for the volumetric mass. In this case, CPCA and AQM have very different behavior compared to both MMP methods. For all quantities of interest except  $L^1$ , CPCA and AQM errors have a plateau for higher basis sizes. For higher basis sizes, MMP method performs better than AQM and CPCA for all studied quantities, except the 25% quantile. Thus, contrary to what happens in the other presented numerical examples, CPCA method does not perform better than AQM and MMP for distances. For MMP decompositions, the errors on the mean and quantiles display a jump for basis sizes of 12 or 13.

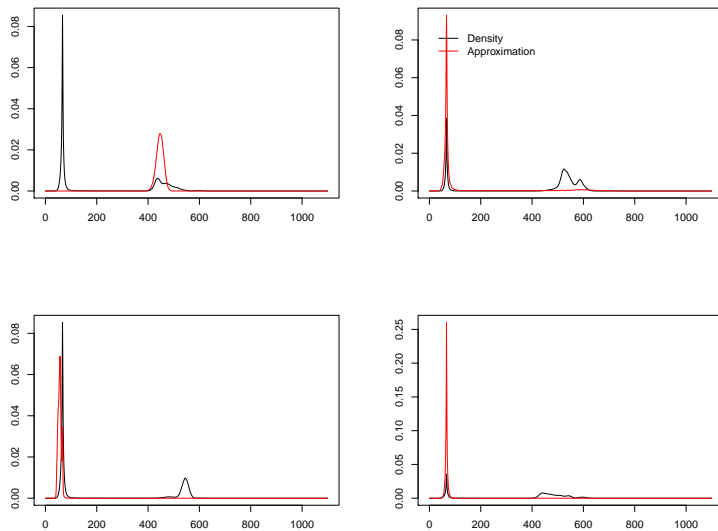


FIGURE 27. GIBBS test case: distribution of the volumetric mass density and its KR approximation for the sets of input parameters 1, 2, 3 and 4.

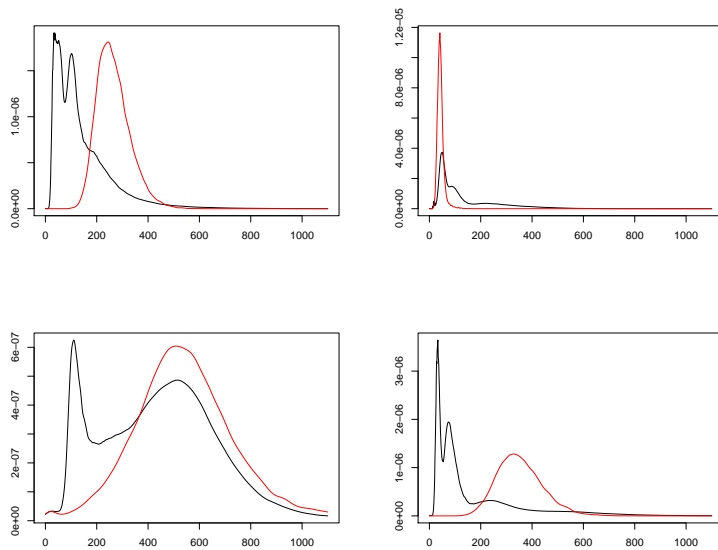


FIGURE 28. GIBBS test case: distribution of the vapor pressure and its KR approximation for the sets of input parameters 31, 32, 33 and 34.

#### 4. CONCLUSION

The aim of this work was to build a metamodel for code outputs which are probability density functions. A first idea to design such a model was to build a convex combination of the sample probability density functions.

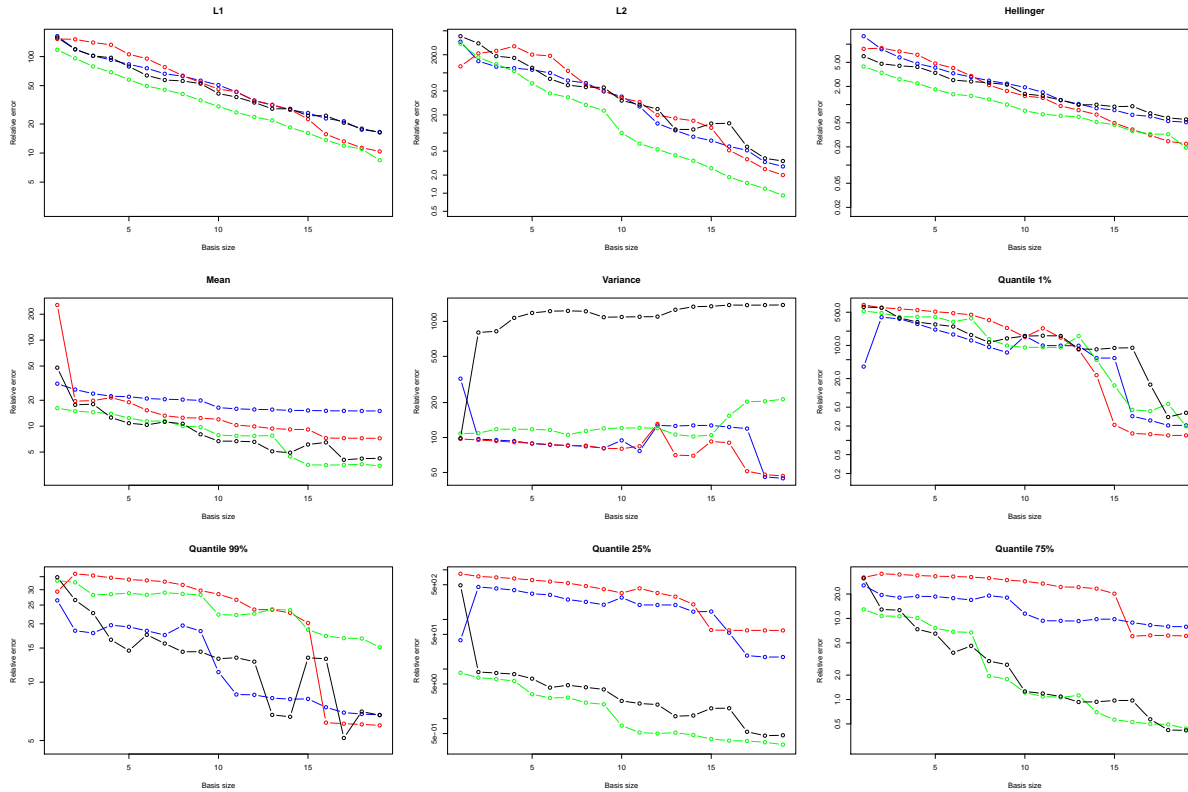


FIGURE 29. GIBBS test case: comparison of the relative errors on different quantities for the volumetric mass density in function of the size of the basis  $q$ . Blue: MMP decomposition given by  $L^2$  norm. Red: MMP decomposition given by the Hellinger distance. Black: AQM decomposition. Green: CPCA decomposition.

To this mean, we first adapted the well-known kernel regression developed by [20] and [30]. Second, we proposed to approximate the sample probability density functions on a functional basis in order to reduce the problem dimension. In this way, the probability density functions are characterized by their coefficients on the basis, so that the problem becomes finite-dimensional. Two methods have been proposed to build a functional basis. The first method is adapted from Magic Points Method [16] and builds iteratively the basis by adding the sample function that maximizes the approximation error on the previous basis. The second one, called Alternate Quadratic Minimization (AQM), aims at minimizing the  $L^2$  approximation error under the defined constraints. Both methods require the resolution of constrained optimization problems. Then, in future work, a metamodel could be adjusted on the coefficients, to link them with the uncertain inputs, and provide a global metamodel for probability density function outputs.

Our methods have been tested on two analytical test cases and three industrial applications proposed by CEA, EDF and IFPEN. The kernel regression-based metamodel performs well on the analytical test cases, but shows its limits on the three numerical codes. Indeed, kernel regression is efficient in low dimension, but the difficulty to estimate the parameters increases with the dimension and the size of the sample. Moreover, this method is not adapted for numerical codes for which the influence of the controllable parameters is much more important than one of the uncontrollable parameters. Both Modified Magic Points (MMP) and AQM methods give good results on all the test cases. However, in most cases, the method proposed by Kneip and Utikal [14] performs better than the proposed ones.

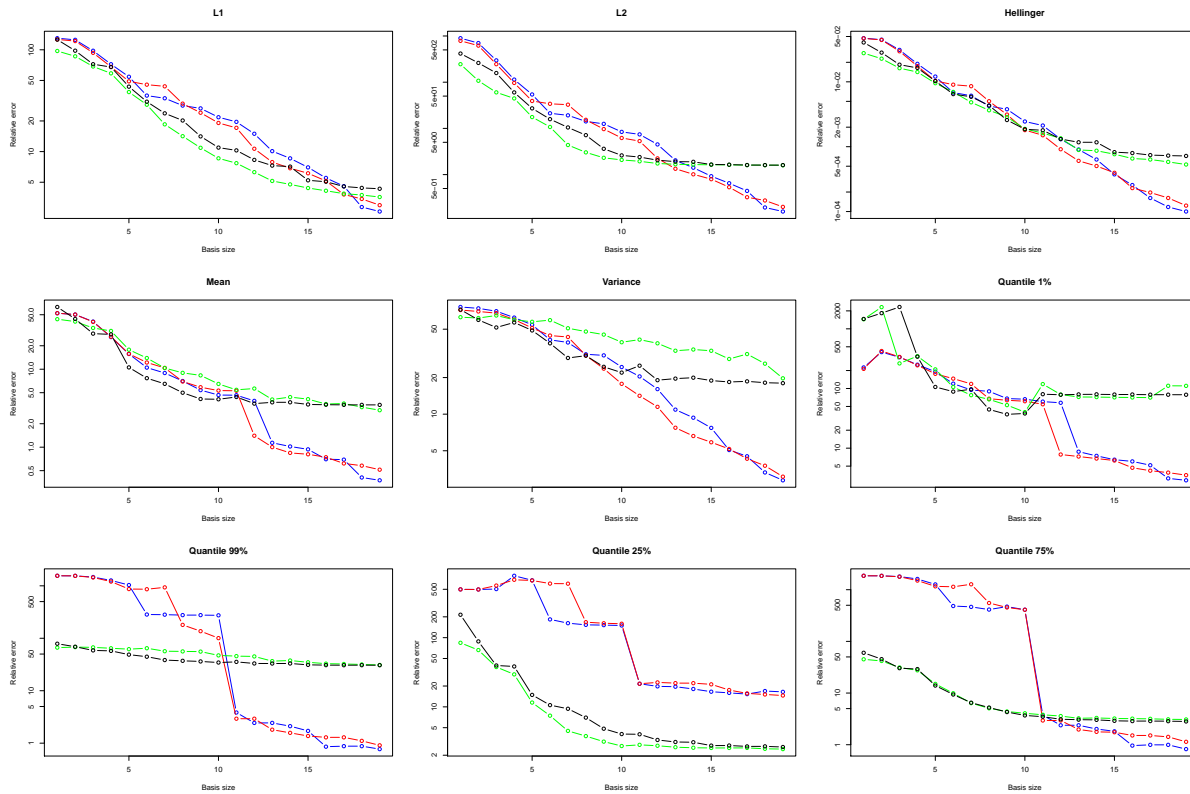


FIGURE 30. GIBBS test case: comparison of the relative errors on different quantities for the vapor pressure in function of the size of the basis  $q$ . Blue: MMP decomposition given by  $L^2$  norm. Red: MMP decomposition given by the Hellinger distance. Black: AQM decomposition. Green: CPCA decomposition.

Metamodels linking the uncertain inputs of the code and the functional basis coefficients remains to be defined. The main difficulty to build this metamodel lies in ensuring that the sum of the functional basis coefficients is equal to one. One of the limits of our methods is that all quantities of interest are not as well approximated. In particular, quantiles are often poorly approximated. Indeed, all the proposed methods are based on norm minimization, and the norms used are not well suited for the study of extreme values. It could be interesting to adapt the proposed methods to norms giving more weights on the tails of the densities. The use of the Wasserstein metric between probability measures will be also studied.

We would like to thank our supervisors Nicolas Bousquet (EDF), Frédéric Delbos (IFPEN), Bertrand Iooss (EDF), Jérôme Lonchamp (EDF), Rafael Lugo (IFPEN), Amandine Marrel (CEA) and Nadia Pérot (CEA). We would also like to thank Anthony Nouy (École Centrale de Nantes) and Sébastien Da Veiga (IFPEN) for their guidance. Furthermore, we would like to thank Nicolas Champagnat (Institut Élie Cartan de Lorraine) and Tony Lelièvre (École des Ponts ParisTech) who organized the CEMRACS 2013 in collaboration with Anthony Nouy. Finally, we would like to thank CEA, EDF and IFPEN for funding this research project.

## REFERENCES

- [1] John Affleck-Graves, A.H. Money, and C.G. Troskie. A principal component index subject to constraints. *Investment Analysts Journal*, November 1979.
- [2] John Aitchison. The statistical analysis of compositional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44(2):139–177, 1982.
- [3] Bruce Ankenman, Barry L Nelson, and Jeremy Staum. Stochastic kriging for simulation metamodeling. *Operations research*, 58(2):371–382, 2010.
- [4] Mario Bebendorf, Yvon Maday, and Benjamin Stamm. Comparison of some reduced representation approximations. *arXiv preprint arXiv:1305.5066*, 2013.
- [5] Richard H Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208, 1995.
- [6] Nguyen Ngoc Cuong. Reduced-basis approximations and a posteriori error bound for nonaffine and nonlinear partial differential equations: Application to inverse analysis, 2005.
- [7] Sébastien Da Veiga and Amandine Marrel. Gaussian process modeling with inequality constraints. In *Annales de la faculté des sciences de Toulouse Mathématiques*, volume 21, pages 529–555. Université Paul Sabatier, Toulouse, 2012.
- [8] Pedro Delicado. Dimensionality reduction when data are density functions. *Computational Statistics & Data Analysis*, 55(1):401–420, 2011.
- [9] Karine Fessart and Jérôme Lonchamp. Optimisation d’un programme d’investissements : méthode IPOP. *Conférence lambda-mu, La Rochelle, France*, 2010.
- [10] Donald Goldfarb and Ashok Idrani. A numerically stable dual method for solving strictly convex quadratic programs. *Mathematical programming*, 27(1):1–33, 1983.
- [11] Wolfgang Hardle, James Stephen Marron, et al. Optimal bandwidth selection in nonparametric regression function estimation. *The Annals of Statistics*, 13(4):1465–1481, 1985.
- [12] Trevor J Hastie and Robert J Tibshirani. *Generalized Additive Models*. Chapman & Hall, London, 1990.
- [13] Bertrand Iooss and Mathieu Ribatet. Global sensitivity analysis of computer models with functional inputs. *Reliability Engineering & System Safety*, 94(7):1194 – 1204, 2009.
- [14] Alois Kneip and Klaus J Utikal. Inference for density families using functional principal component analysis. *Journal of the American Statistical Association*, 96(454):519–542, 2001.
- [15] J. Lonchamp and K. Fessart. An optimization approach for life cycle management applied to large power transformers. *1023033 EPRI Report*.
- [16] Yvon Maday, Ngoc Cuong Nguyen, Anthony T Patera, and George SH Pau. A general, multipurpose interpolation procedure: the magic points. In *Proceedings of the 2nd International Conference on Scientific Computing and Partial Differential Equations*, 2007.
- [17] Amandine Marrel, Bertrand Iooss, Sébastien Da Veiga, and Mathieu Ribatet. Global sensitivity analysis of stochastic computer models with joint metamodels. *Statistics and Computing*, 22(3):833–847, 2012.
- [18] Michael D McKay, Richard J Beckman, and William J Conover. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.
- [19] Hans-Georg Muller and Ulrich Stadtmuller. Variable bandwidth kernel estimators of regression curves. *The Annals of Statistics*, 15(1):182–201, 1987.
- [20] Elizbar A Nadaraya. On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142, 1964.
- [21] Emanuel Parzen. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33(3):1065–1076, 1962.
- [22] J. O. Ramsay and B. W. Silverman. *Functional Data Analysis*. Springer Series in Statistics. Springer, 2nd edition edition, 2005.
- [23] Brian J Reich, Eric Kalendra, Curtis B Storlie, Howard D Bondell, and Montserrat Fuentes. Variable selection for high dimensional bayesian density estimation: application to human exposure simulation. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 61(1):47–66, 2012.
- [24] RA Rigby and DM Stasinopoulos. Construction of reference centiles using mean and dispersion additive models. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 49(1):41–50, 2000.
- [25] Murray Rosenblatt. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, pages 832–837, 1956.
- [26] Gianluigi Rozza. Reduced basis methods for stokes equations in domains with non-affine parameter dependence. *Computing and Visualization in Science*, 12(1):23–35, 2009.
- [27] George R Terrell and David W Scott. Variable kernel density estimation. *The Annals of Statistics*, pages 1236–1265, 1992.
- [28] Timo Tonn and Karsten Urban. A reduced-basis method for solving parameter-dependent convection-diffusion problems around rigid bodies. In *ECCOMAS CFD*, 2006.

- [29] Philippe Ungerer, Christèle Beauvais, Jérôme Delhommelle, Anne Boutin, Bernard Rousseau, and Alain H. Fuchs. Optimization of the anisotropic united atoms intermolecular potential for n-alkanes. *Journal of Chemical Physics*, 112(12):5499–5510, March 2000.
- [30] Geoffrey S Watson. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 359–372, 1964.
- [31] Isabelle Zabalza-Mezghani, Emmanuel Manceau, Mathieu Feraille, and Astrid Jourdan. Uncertainty management: From geological scenarios to production scheme optimization. *Journal of Petroleum Science and Engineering*, 44(1):11–25, 2004.

## A. APPENDIX

### A.1. $L^2$ distance-based kernel estimator

We prove here the equivalence between the classical kernel estimator and the  $L^2$  distance-based kernel estimator introduced in Sections 1.1.1 and 1.1.2 respectively. Indeed the classical kernel estimator (1) can be retrieved by writing the optimality condition of the minimization problem (3) defining the  $L^2$  distance-based kernel estimator:

$$\begin{aligned}
 & \sum_{i=1}^N -2K_H(\mathbf{x}_i, \mathbf{x}_0) \int_I (f_i - \hat{f}_0) = 0 \\
 \iff & \int_I \left( \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) (f_i - \hat{f}_0) \right) = 0 \\
 \iff & \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) (f_i - \hat{f}_0) = 0 \\
 \iff & \hat{f}_0 = \sum_{i=1}^N \frac{K_H(\mathbf{x}_i, \mathbf{x}_0)}{\sum_{j=1}^N K_H(\mathbf{x}_j, \mathbf{x}_0)} f_i.
 \end{aligned}$$

### A.2. Hellinger distance-based kernel estimator

We prove here the equivalence between expressions (4) and (5) of the Hellinger distance-based kernel estimator given in Section 1.1.2. The constraint that the integral of  $\hat{f}_0$  is 1 in the optimization problem (4) is handled using the associated Lagrangian function. The problem becomes

$$\hat{f}_0 = \arg \min_{f, \lambda} L(f, \lambda),$$

with

$$L(f, \lambda) = \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) \int_I (\sqrt{f_i} - \sqrt{f})^2 - \lambda \left( \int_I f - 1 \right).$$

The first order optimality conditions are

$$\begin{cases} 0 = \frac{\partial L}{\partial f} = \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) \int_I \left( 1 - \frac{\sqrt{f_i}}{\sqrt{f}} \right) - \lambda \int_I 1, \\ 0 = \frac{\partial L}{\partial \lambda} = 1 - \int_I f. \end{cases} \quad (15)$$



The first condition in (15) gives

$$\begin{aligned}
& \int_I \left( \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) \left( 1 - \frac{\sqrt{f_i}}{\sqrt{\hat{f}_0}} \right) - \lambda \right) = 0 \\
\iff & \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) \left( 1 - \frac{\sqrt{f_i}}{\sqrt{\hat{f}_0}} \right) - \lambda = 0 \\
\iff & \hat{f}_0 = \left( \sum_{i=1}^N \frac{K_H(\mathbf{x}_i, \mathbf{x}_0) \sqrt{f_i}}{\sum_{j=1}^N K_H(\mathbf{x}_j, \mathbf{x}_0) - \lambda} \right)^2 = \frac{\left( \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) \sqrt{f_i} \right)^2}{\left( \sum_{j=1}^N K_H(\mathbf{x}_j, \mathbf{x}_0) - \lambda \right)^2}. \tag{16}
\end{aligned}$$

Then we replace  $\hat{f}_0$  in the second condition of (15) by the expression given in (16):

$$\left( \sum_{j=1}^N K_H(\mathbf{x}_j, \mathbf{x}_0) - \lambda \right)^2 = \int_I \left( \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) \sqrt{f_i} \right)^2. \tag{17}$$

The combined equations (16) and (17) entail the formula given by (5) for the kernel estimator associated to the Hellinger distance:

$$\hat{f}_0 = \frac{\left( \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) \sqrt{f_i} \right)^2}{\int_I \left( \sum_{i=1}^N K_H(\mathbf{x}_i, \mathbf{x}_0) \sqrt{f_i} \right)^2}.$$