A moment-matching method to study the variability of phenomena described by partial differential equations

Jean-Frédéric Gerbeau, Damiano Lombardi, Eliott Tixier

A MOMENT-MATCHING METHOD TO STUDY THE VARIABILITY OF PHENOMENA DESCRIBED BY PARTIAL DIFFERENTIAL EQUATIONS

Jean-Frédéric Gerbeau*†, Damiano Lombardi*†, Eliott Tixier†*

November 7, 2016

Abstract

Many phenomena are modeled by deterministic differential equations, whereas the observation of these phenomena, in particular in life sciences, exhibits an important variability. This paper addresses the following question: how can the model be adapted to reflect the observed variability?

Given an adequate model, it is possible to account for this variability by allowing some parameters to adopt a stochastic behavior. Finding the parameters probability density function that explains the observed variability is a difficult stochastic inverse problem, especially when the computational cost of the forward problem is high. In this paper, a non-parametric and non-intrusive procedure based on offline computations of the forward model is proposed. It infers the probability density function of the uncertain parameters from the matching of the statistical moments of observable degrees of freedom (DOFs) of the model. This inverse procedure is improved by incorporating an algorithm that selects a subset of the model DOFs that both reduces its computational cost and increases its robustness. This algorithm uses the pre-computed model outputs to build an approximation of the local sensitivities. The DOFs are selected so that the maximum information on the sensitivities is conserved. The proposed approach is illustrated with elliptic and parabolic PDEs. In the Appendix, an nonlinear ODE is considered and the strategy is compared with two existing ones.

Keywords:
stochastic inverse problem, maximum entropy, moment matching, backward uncertainty quantification

*Inria, Centre Inria de Paris, 2 rue Simone Iff, 75012 Paris, France.
†Sorbonne Universités UPMC, Laboratoire J-L. Lions, 75005 Paris, France.

1 Introduction

The context of this work is the following: a collection of experimental measurements is available, which exhibit variability, caused for instance by an heterogeneity in the physical settings [1, 2]. We assume that the observable quantities correspond to the degrees of freedom (DOFs) of a model that depends on fixed and uncertain parameters. The model is typically a system of ordinary differential equations (ODE) or partial differential equations (PDE).

The aim of this paper is twofold. First, we propose a non-parametric and non-intrusive method to estimate the uncertain parameters probability density function (PDF) by exploiting the observable variability. Second, we propose a method to make this estimation "parsimonious", *i.e.* requiring as few model evaluations as possible and as few observables (or DOFs) as possible.

To tackle the first problem, two different strategies may be envisioned. First, one could estimate the model parameters associated with each experimental sample using classical inverse problem tools such as Bayesian approaches [3, 4] or genetic algorithms [5]. These strategies would yield a collection of parameters values from which the PDF would be computed by using histograms or more sophisticated PDF estimation techniques [6]. As straightforward as this approach is, it becomes computationally intensive as the number of experimental samples grows larger. Second, one may see the experimental data set as a whole, which has the advantage of being both computationally cheaper and more robust to noise and low-quality measurements. In this paper, we focus on the second strategy and present an adaptation of the well-known problem of moments [7]. The problem of moments consists in finding the PDF of the parameters such that its statistical moments have a prescribed set of values. It has been used as an inverse problem tool with success in various contexts [8, 2, 9]. A popular regularization of the problem of moments is the maximum entropy principle, which is rooted in information theory and is justified by practical mathematical considerations [10, 11]. In most cases however, parameters of a model are not directly observable. Therefore, one needs a technique that takes into account the observable variability. In this context, we introduce an "observable moment matching" method which consists in maximizing the PDF entropy under the constraints of matching the moments of the observable itself (not of the parameters). This is a two-step method. First, the model is evaluated for a fixed number of parameters samples and the corresponding outputs, i.e. the simulated observables, are stored. Second, the PDF is found by an iterative process that maximizes its entropy under the constraints of matching the moments of the experimental and simulated observables.

To address the second problem, we propose an algorithm that selects the DOFs in the physical domain where the moments are to be matched in order

to alleviate the cost of the inverse problem – which is crucial for complex models such as PDEs – and to improve its conditioning. This algorithm exploits the sensitivity information provided by the pre-computed model evaluations. The sensitivity Gram matrix, computed for every DOF, reveals active subspaces [12, 13] of the parameter space. The DOFs are selected by clustering the active subspaces and choosing their best representatives. This strategy allows for a reduction of the number of DOFs by several orders of magnitude and therefore proves to drastically reduce the computational cost of the inverse problem without requiring any additional evaluation of the model.

This paper is organized as follows. The whole methodology is detailed in Section 2. First, we introduce the observable moment matching algorithm and we formulate the associated inverse problem in terms of an optimization problem. Then, the clustered sensitivities algorithm is introduced and the reduction of the number of DOFs is explained. In Section A, our approach is illustrated with a set of ODEs modeling the transient action potential of a heart cell. We compare its performance with two existing statistical inverse problem techniques: one proposed by N. Zabaras and B. Ganapathysubramanian [14], the other one proposed by E. Kuhn and M. Lavielle [15]. In Section 3, our algorithm is applied to the Darcy equations. The PDF of five coefficients that parametrize an inner field is recovered using measurements on the domain boundaries. Then, we consider a nonlinear parabolic PDE model, namely the FKPP equation. Under certain conditions, this model exhibits a wave propagation whose shape depends on the location of the source term and on certain parameters. The PDFs of the source term and the reaction parameters are recovered using measurements at different times and locations.

Finally, we present some concluding remarks in Section 4.

2 Methodology

2.1 Notation

Let us consider a data set that exhibits variability and a physical model assumed to accurately depict the observations. Let $\mathcal{D} \subseteq \mathbb{R}^d$ be an open subset, the physical domain (space, time or space-time), in which the governing equations are written. Let $(\Theta, \mathcal{A}, \mathcal{P})$ be a complete probability space, Θ being the set of outcomes, $\mathcal{A}$ a σ–algebra and $\mathcal{P}$ a probability measure. The model can be written in a compact notation as:

$$\mathcal{L}(u(\mathbf{x}, \boldsymbol{\theta})) = \mathbf{0}, \tag{1}$$

where $\mathcal{L}$ denotes a generic nonlinear differential operator.
The vector $\boldsymbol{\theta} = \left(\theta_1, \ldots, \theta_{n_p}\right) \in \Theta$ denotes the uncertain parameters of the

model and Θ is a bounded subset of $\mathbb{R}^{n_p}$, sometimes referred to as the stochastic domain [14]. A set of measurements $\{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$ is available. Each measurement $\mathbf{y}_i$ is assumed to take the following form:

$$\mathbf{y} = g\left(u(\mathbf{x}, \boldsymbol{\theta})\right) + \epsilon, \tag{2}$$

where g is a function describing the measurement process and ϵ is the noise, assumed to be additive and independent. For practical reasons, g is normalized to take values in $[0, 1]$. Let $\mathbb{E}$ be the expectation operator. We make the hypothesis that the random fields associated with the observables are p-integrable, that is: $\int_{\mathcal{D}} |\mathbb{E}(y^p)| \, \mathrm{d}\mathbf{x} < M$, where the exponent p is the highest available moment. The variability in the observations is due to two main contributions: the variability in the parameters and the noise in the measurement process. In a classical forward Uncertainty Quantification (UQ) context, given the probability density function (PDF) of the parameters ρ, the moments of the observables are computed. In the present work, an inverse problem is solved which consists in finding the PDF of the parameters that generates the observed variability in a set of available data. Let us introduce the m^{th} order empirical moment of the measurements:

$$\mu_m(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^{N} y_i(\mathbf{x})^m \approx \mathbb{E}((g + \epsilon)^m), \tag{3}$$

and the m^{th} order moment of the simulations:

$$\mu_m^{\rho}(\mathbf{x}) = \int_{\boldsymbol{\theta} \in \Theta} \left(y_{sim}(\boldsymbol{\theta})\right)^m \rho(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} = \mathbb{E}(y_{sim}^m), \tag{4}$$

where y_{sim} are the observations of the simulated system.

2.2 Handling the noise

Under the assumption that the noise is additive, independent and with a known structure, it is straightforward to account for its influence on the measurements moments. Using the linearity of the expectation operator and the independence of the noise, it follows from definition (2) that:

$$\mathbb{E}\left[y^m\right] = \sum_{k=0}^{m} \binom{m}{k} \mathbb{E}\left[g^m\right] \mathbb{E}\left[\epsilon^{m-k}\right].$$

As an example, consider the case where the noise follows a zero-mean normal distribution with a known variance τ^2: $\epsilon \sim \mathcal{N}(0, \tau^2)$. Then, the following corrections may be applied to the first three empirical moments defined in Eq. (3):

$$\begin{aligned}
\tilde{\mu}_1(\mathbf{x}) &= \mu_1(\mathbf{x}), \\
\tilde{\mu}_2(\mathbf{x}) &= \mu_2(\mathbf{x}) - \tau^2, \\
\tilde{\mu}_3(\mathbf{x}) &= \mu_3(\mathbf{x}) - 3\tau^2 \mu_1(\mathbf{x}).
\end{aligned}$$

In the numerical experiments, the noise is assumed to be gaussian and its level is defined as the ratio $4\tau/A$ where A is the signal amplitude. In Section A, the effect of τ^2 on the PDF estimation is investigated.

Only Gaussian noises are considered here. However, the same procedure may be applied to any noise whose power moments are known. If the noise structure is completely unknown, a strategy can be set up to estimate it but it is not investigated in the present work.

2.3 Overview of the strategy

The overall algorithm aims at estimating the PDF ρ of the uncertain parameters $\boldsymbol{\theta}$, given the empirical moments of the observables. The Jaynes principle of maximum entropy is applied (see [10]): the PDF is sought so that it has the maximum entropy under the constraints that the experimental and simulated moments be equal. Two additional constraints correspond to the positivity and the PDF normalization. This leads to the following optimization problem:

$$\begin{cases} \textbf{Minimize:} & \int_\Theta \rho \log(\rho) & \\ \textbf{Subject to:} & \tilde{\mu}_m(\mathbf{x}) - \mu_m^\rho(\mathbf{x}) = 0, & \forall \mathbf{x} \in \mathcal{D}, 1 \leq m \leq N_m, \\ & \rho(\boldsymbol{\theta}) \geq 0, & \forall \boldsymbol{\theta} \in \Theta, \\ & \int_\Theta \rho = 1. & \end{cases} \tag{5}$$

In what follows, this is referred to as the Observable Moment Matching (OMM) problem. In Section 2.4 the optimality conditions for the OMM problem are derived and a dual formulation is introduced. The latter leads to a nonlinear problem which is, in general, ill-conditioned. Moreover, its computational cost is prohibitive when models described by PDEs are at hand. To overcome these difficulties a reduction approach is introduced, based on a sensitivity analysis. As a consequence, the OMM procedure is only applied to a subset $\mathcal{S}$ of the DOFs of the model variables discretized in the physical domain $\mathcal{D}$. More precisely, the eigendecomposition of an approximation of the following matrix is computed:

$$\mathbf{C}(\mathbf{x}) = \int_\Theta \left[\nabla_{\boldsymbol{\theta}} g(\mathbf{x}, \boldsymbol{\theta})\right] \left[\nabla_{\boldsymbol{\theta}} g(\mathbf{x}, \boldsymbol{\theta})\right]^T \rho(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta}, \tag{6}$$

referred to as the exact sensitivity Gram matrix (SGM).The study of the SGM eigenvalues allows us to identify active subspaces [12] in the parameter space associated with each DOF. The subspaces are clustered based on a similarity function and the "best" DOFs are then picked based on a criterion defined in Section 2.5 to form the selected subset $\mathcal{S}$. This selection method will be later referred to as the Clustered Sensitivities (CS) procedure.

2.4 Inverse problem: observable moment matching (OMM)

The classical problem of moments consists in finding a PDF ρ of the parameters θ_k from the knowledge of a finite number N_m of its power moments $\mu_{m,k}$, $m = 1, \dots, N_m$, $k = 1, \dots, n_p$:

$$\mathbb{E}_\rho\left[\theta_k^m\right] = \mu_{m,k}, \qquad m = 1, \dots, N_m,\ k = 1, \dots, n_p,$$

where $\mathbb{E}_\rho(\cdot)$ denotes the expectation operator given a density function ρ. This problem has been extensively discussed in the literature and has been addressed by adopting a wide range of strategies. When only a finite number of moments are known, which is often the case in practice, the problem becomes under-determined. Indeed, there exists an infinite number of densities that have the same N_m moments. Therefore, one needs to introduce a regularization in order to obtain a unique distribution function among all the feasible solutions. Several approaches exist, such as minimizing the mean squared error $\epsilon(\rho) = \sum_{m,k}\left(\mathbb{E}_\rho\left[\theta_k^m\right] - \mu_{m,k}\right)^2$ with the constraint that ρ be a finite expansion of polynomials [16] or Padé approximants [17].

This problem has been successfully used in situations where the moments of the model parameters are directly measurable, for instance in the context of microstructure reconstruction [8, 2, 9]. In general however, the moments of the model parameters are not observable. Therefore, we propose to apply the moment matching constraints not on the parameters but on the observable itself.

To regularize the problem, the maximum entropy principle is used: find the PDF that maximizes the entropy under the constraint of matching the first N_m moments, where the Shannon definition [18] of the PDF entropy reads: $S(\rho) = -\int_\Theta \rho \log(\rho)$. There are three main reasons why this choice of regularization is well suited to the present case. First, from an information theory point of view, the maximum entropy PDF is considered the best choice when a limited amount of information is available (here, only a finite number of moments are known). This principle was first introduced by Jaynes [10] and was successfully applied to numerous practical cases [11, 2, 19, 20]. Second, $-S(\rho)$ is a convex cost function which enables the use of efficient optimization tools. Last, ρ can be written as an exponential term (see below), which dispenses the addition of an inequality constraint ensuring its positivity.

A set of constraint functions is introduced, expressing the mismatch between the moments of the measured observable and the moment of the simulated observable. They read:

$$c_m(\mathbf{x}) = \mu_m^\rho(\mathbf{x}) - \tilde{\mu}_m(\mathbf{x}) = \int_\Theta g^m(\mathbf{x}, \boldsymbol{\theta})\rho(\boldsymbol{\theta})\ \mathrm{d}\boldsymbol{\theta} - \tilde{\mu}_m(\mathbf{x}), \quad m = 1, \dots, N_m.$$

Introducing the Lagrange multipliers $\lambda(\mathbf{x}) = (\lambda_m(\mathbf{x}))_{m=1\dots N_m}$, λ_0 and

$\nu(\boldsymbol{\theta})$, the initial optimization problem (5) is recast in the following saddle-point problem:

$$\inf_{\rho} \sup_{\lambda,\lambda_0,\nu\geq 0} \mathcal{L}\left(\rho,\lambda,\lambda_0,\nu\right), \tag{7a}$$

with

$$\mathcal{L}\left(\rho,\lambda,\nu\right) = \int_\Theta \rho\log(\rho) - \sum_{m=1}^{N_m}\int_{\mathcal{D}} \lambda_m(\mathbf{x})c_m(\mathbf{x})\ \mathrm{d}\mathbf{x} - \lambda_0\left(\int_\Theta \rho - 1\right) - \int_\Theta \rho\nu. \tag{7b}$$

The necessary conditions for optimality give:

$$\rho = \exp(\lambda_0 - 1)\exp\left(\textstyle\sum_{m=1}^{N_m}\int_{\mathcal{D}} g^m\lambda_m\ \mathrm{d}\mathbf{x}\right), \tag{8}$$

$$\textstyle\int_\Theta g^m \exp(\lambda_0 - 1)\exp\left(\sum_{h=1}^{N_m}\int_D g^h\lambda_h\ dx\right)\ \mathrm{d}\boldsymbol{\theta} - \tilde{\mu}_m = 0,\ m = 1,\ldots,N_m \tag{9}$$

In the present case, by virtue of the entropy regularization, the primal variable ρ can be expressed in an analytic form as a function of the dual variable and the positivity constraint is automatically satisfied (Eq.(8)). Hence, the solution of the system can be reduced to the solution of a nonlinear problem for the dual variable (Eq.(9)).

The error in the density can be evaluated by means of the Kullback-Leibler divergence. In the following Lemma, it is shown that it is bounded by the error in the dual variable. Let the distribution that maximizes the entropy under the moment constraints be denoted by ρ^* and let its actual approximation be ρ. The true dual variable being λ^* and its approximation being λ, the error in the dual variable is defined as: $\delta\lambda := \lambda^* - \lambda$. The following result holds.

Lemma 1. *The Kullback-Leibler divergence between ρ^* and ρ is bounded by the error in the dual variable as follows:*

$$|KL(\rho^*|\rho)| \leq |\delta\lambda_0| + \mathrm{meas}(\mathcal{D})^{1/2}\sum_{m=1}^{N_m}\|\delta\lambda_m\|_{\mathbf{x},2}. \tag{10}$$

Proof. The Kullback-Leibler (KL) divergence reads:

$$KL(\rho^*|\rho) := \int_\Theta \log\left(\frac{\rho^*}{\rho}\right)\ \rho^*\mathrm{d}\boldsymbol{\theta}. \tag{11}$$

We deduce from the optimality conditions:

$$\frac{\rho^*}{\rho} = \exp(\delta\lambda_0)\exp\left(\sum_{m=1}^{N_m}\langle g^m, \delta\lambda_m\rangle_{\mathbf{x}}\right), \tag{12}$$

so that the expression of the KL divergence can be rewritten as:

$$KL(\rho^*|\rho) = \delta\lambda_0 + \sum_{m=1}^{N_m} \int_\Theta \langle g^m, \delta\lambda_m \rangle_{\mathbf{x}} \ \rho^* \mathrm{d}\boldsymbol{\theta}. \tag{13}$$

The Cauchy-Schwarz inequality is applied to the scalar product in the physical space:

$$|KL(\rho^*|\rho)| \leq |\delta\lambda_0| + \sum_{m=1}^{N_m} \|\delta\lambda_m\|_{\mathbf{x},2} \int_\Theta \|g^m\|_{\mathbf{x},2}(\boldsymbol{\theta}) \ \rho^* \mathrm{d}\boldsymbol{\theta}. \tag{14}$$

Hence the result since the observable is bounded by 1. □

2.4.1 Discretization of the inverse problem

The discretization of the nonlinear system Eq.(9) is addressed in this section. The observable, as well as the Lagrange multipliers λ_m, are discretized in space (or space-time) by means of standard methods and the total number of DOFs is denoted by $N_{\mathbf{x}}$. The integrals in the stochastic space are approximated by a quasi-Monte Carlo method. The stochastic domain Θ is discretized using the Sobol sequence [21]. These quasi-random samples have a low-discrepancy, and are competitive compared to random uniform samples [22]. When integrating functions featuring a certain regularity, sparse grid methods, which are often used in uncertainty propagation (see [23]), can outperform quasi-Monte Carlo ones [24]. In the present context, however, a reason to prefer a quasi-Monte Carlo discretization of the stochastic domain is that the probability density distribution is the unknown of the problem, and it is not known in advance. Roughly speaking, since sparse grids have strong preferential directions, the risk of "missing" the area of interest in the stochastic domain is non-negligible, making evenly distributed points a more suitable discretization. Figure 1 shows how a two-dimensional domain is discretized using each of the three options described above. It illustrates how the Sobol sequence both performs a more even coverage of the domain than uniform pseudo-random samples and does not favor specific directions such as in sparse grids. Let us denote by N_c the number of sample points in Θ and $|\Theta|$ the stochastic domain volume.

To compute the integrals approximations in (9), the model is evaluated for each sample $\boldsymbol{\theta}_i$. The corresponding set $\{y_{sim}(\boldsymbol{\theta}_i, \mathbf{x}_j), i = 1, \ldots, N_c, \ j = 1, \ldots, N_{\mathbf{x}}\} \in \mathbb{R}^{N_c \times N_{\mathbf{x}}}$ will later be referred to as the *simulation set*. Assuming a subset $\mathcal{S}$ of $\mathcal{D}$ has been selected, the number of DOFs in $\mathcal{S}$ is denoted by N_k. For the sake of clarity, the following notation is now used:

$$\rho_i = \rho(\boldsymbol{\theta}_i), \quad g_{i,j} = y_{sim}(\boldsymbol{\theta}_i, \mathbf{x}_j), \quad \lambda_{j,m} = \lambda_m(\mathbf{x}_j), \quad \mu_{j,m} = \mu_m(\mathbf{x}_j), \quad \beta = \frac{|\Theta|}{N_c}$$

$$\text{for} \quad i = 1, \ldots, N_c, \quad j = 1, \ldots, N_k, \quad m = 1, \ldots, N_m.$$

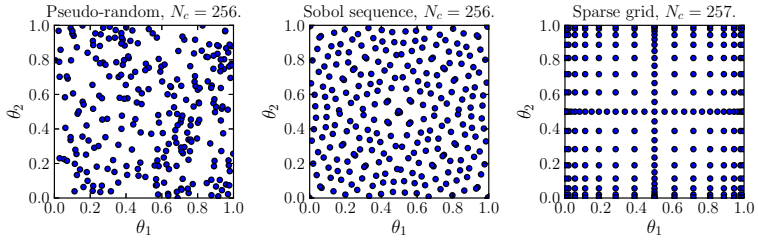

Figure 1: Different discretizations of the parameter space: random uniform (left), Sobol sequence (center), sparse grid (right).

The discretization of Eq.(8) reads:

$$\rho_i = \exp\left(\lambda_0 - 1\right) \exp\left(\sum_{j=1}^{N_k} \sum_{m=1}^{N_m} \omega_j \lambda_{j,m} g_{i,j}^m\right), \quad i = 1, \ldots, N_c, \tag{15}$$

where the w_j are quadrature weights for the physical domain discretization.

Before discretizing Eq.(9), a vector form is introduced, for the sake of compactness. Let $\boldsymbol{\omega} = \begin{bmatrix}\omega_1 & \ldots & \omega_{N_k}\end{bmatrix}$, $\boldsymbol{\lambda} = \begin{bmatrix}\lambda_{1,1} & \ldots & \lambda_{N_k,N_m} & \lambda_0 - 1\end{bmatrix}^T$, $\boldsymbol{\mu} = \begin{bmatrix}\tilde{\mu}_{1,1} & \ldots & \tilde{\mu}_{N_k,N_m} & 1\end{bmatrix}^T$ and $\mathbf{G} = \begin{bmatrix}\mathbf{G}^{(1)} & \ldots & \mathbf{G}^{(N_m)} & \mathbf{1}\end{bmatrix}^T$ with $G_{i,j}^{(k)} = g_{i,j}^k$, $k = 1, \ldots, N_m$, $i = 1, \ldots, N_c$, $j = 1, \ldots, N_k$. Note that $\mathbf{G} \in \mathbb{R}^{N_G \times N_c}$ where $N_G = N_k N_m + 1$. It has an extra column of ones to take into account the normalization constraint. Finally, let $\boldsymbol{\Delta} = \mathrm{diag}(\underbrace{\boldsymbol{\omega}, \boldsymbol{\omega}, \ldots, \boldsymbol{\omega}}_{N_m \text{ times}}, 1)$.

The density can be written as: $\boldsymbol{\rho} = \exp\left(\mathbf{G}^T \boldsymbol{\Delta} \boldsymbol{\lambda}\right)$. The discretization of Eq.(9) reads:

$$\beta \mathbf{G} \boldsymbol{\rho} - \boldsymbol{\mu} = \beta \mathbf{G} \exp\left(\mathbf{G}^T \boldsymbol{\Delta} \boldsymbol{\lambda}\right) - \boldsymbol{\mu} = 0. \tag{16}$$

A Newton method is used to solve this step. However, since the Hessian is ill-conditioned in practical cases, a regularization is proposed. Let $\mathbf{U}, \mathbf{S}, \mathbf{V}$ be the SVD decomposition of $\mathbf{G}$, done with respect to the scalar product induced by $\boldsymbol{\Delta}$, *i.e.* $\mathbf{U}^T \boldsymbol{\Delta} \mathbf{U} = I$. The residual now reads:

$$\mathbf{r} = \beta \mathbf{U} \mathbf{S} \mathbf{V}^T \exp\left[\mathbf{V} \mathbf{S} \mathbf{U}^T \boldsymbol{\Delta} \boldsymbol{\lambda}\right] - \boldsymbol{\mu}. \tag{17}$$

Instead of making $\mathbf{r}$ vanish, we propose to solve for $\hat{\mathbf{r}} = \hat{\mathbf{U}}^T \boldsymbol{\Delta} \mathbf{r} = 0$. This is equivalent to taking a low-rank approximation $\hat{\mathbf{G}}$ of $\mathbf{G}$ by replacing the matrix of singular values $\mathbf{S}$ with its truncation $\hat{\mathbf{S}}$. $\hat{\mathbf{S}}$ is defined so that it shares the first n_σ singular values with $\hat{\mathbf{S}}$ and the following are set to zero. Replacing $\mathbf{G}$ by $\hat{\mathbf{G}}$ in (17) and left-multiplying by $\hat{\mathbf{U}}^T \boldsymbol{\Delta}$, one obtains:

$$\hat{\mathbf{r}} = \beta \hat{\mathbf{S}} \hat{\mathbf{V}}^T \exp\left[\hat{\mathbf{V}} \hat{\mathbf{S}} \hat{\mathbf{U}}^T \boldsymbol{\Delta} \boldsymbol{\lambda}\right] - \hat{\mathbf{U}}^T \boldsymbol{\Delta} \boldsymbol{\mu}. \tag{18}$$

Proceeding to the change of variables $\boldsymbol{\lambda} = \hat{\mathbf{U}}\boldsymbol{\phi}$, the residual now reads:

$$\hat{\mathbf{r}} = \beta\hat{\mathbf{S}}\hat{\mathbf{V}}^T\boldsymbol{\rho} - \hat{\mathbf{U}}^T\boldsymbol{\Delta\mu}, \tag{19a}$$

$$\text{where} \quad \boldsymbol{\rho} = \exp\left[\hat{\mathbf{V}}\hat{\mathbf{S}}\boldsymbol{\phi}\right]. \tag{19b}$$

Note that the residual is no longer a function of the vector of experimental moments $\boldsymbol{\mu}$ but rather its projection $\hat{\mathbf{U}}^T\boldsymbol{\Delta\mu}$. Therefore, the number of non-truncated singular values n_σ is chosen so that the representation error $\| \left(\mathbf{I} - \hat{\mathbf{U}}\hat{\mathbf{U}}^T\boldsymbol{\Delta}\right)\boldsymbol{\mu}\|$ is smaller than a user-defined tolerance parameter α.

The Hessian matrix of the problem now reads:

$$\mathbf{H} = \frac{\partial\hat{\mathbf{r}}}{\partial\boldsymbol{\phi}} = \beta\hat{\mathbf{S}}\hat{\mathbf{V}}\text{diag}(\boldsymbol{\rho})\hat{\mathbf{V}}^T\hat{\mathbf{S}}, \tag{20}$$

which is symmetric, positive semi-definite of rank n_σ. Its Moore-Penrose pseudo-inverse $\mathbf{P}$ is computed and the Newton actualization step reads:

$$\boldsymbol{\phi}^{(n+1)} = \boldsymbol{\phi}^{(n)} - \mathbf{P}\hat{\mathbf{r}}. \tag{21}$$

The components of $\boldsymbol{\phi}$ are initialized to zero, which is equivalent to taking a uniform PDF as the initial guess for $\boldsymbol{\rho}$ or, more precisely, a uniform mass on the discrete ρ_i.

The overall OMM inverse procedure in summarized in Algorithm 1.

Algorithm 1: Observable moment matching algorithm.

Input:

- $\mathcal{S} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{N_k}\}$ subset of $\mathcal{D}$ selected using CS algorithm.
- $\tilde{\mu}_{j,m}$, $j = 1, \ldots, N_k$, $m = 1, \ldots, N_m$: corrected experimental moments.
- $g_{i,j}$, $i = 1, \ldots, N_c$, $j = 1, \ldots, N_k$: simulation subset.
- A tolerance $\alpha > 0$.
- A stopping criterion for Newton iterations ϵ_{Newton}

Initialization:

- Assemble $\mathbf{G} = \begin{bmatrix}((g_{i,j})) & \ldots & ((g_{i,j}))^{N_m} & \mathbf{1}\end{bmatrix}$ and $\boldsymbol{\mu} = \begin{bmatrix}((\tilde{\mu}_{j,k})) & 1\end{bmatrix}$.
- Compute SVD decomposition of $\mathbf{G}$: $\mathbf{U}, \mathbf{S}, \mathbf{V}$
- Number of singular values $n_\sigma = \text{Card}\left\{\sigma \mid \| \left(\mathbf{I} - \hat{\mathbf{U}}\hat{\mathbf{U}}^T\right) \boldsymbol{\mu}\| \leq \alpha\right\}$.
- $\boldsymbol{\phi}^{(0)} = \mathbf{0}$ (*i.e.* $\boldsymbol{\rho}^{(0)}$ uniform over Θ).

$n = 1$;
while $\|\hat{\mathbf{r}}^{(n-1)}\| > \epsilon_{\text{Newton}}$ **do**
  Compute $\boldsymbol{\rho}^{(n)}$ using (19b);
  Compute residual $\hat{\mathbf{r}}^{(n)}$ = using (19a);
  Assemble Hessian matrix $\mathbf{H}^{(n)}$ using (20);
  Update Lagrange multipliers using (21);
  $n \leftarrow n + 1$;
end

Output: ρ_i, $i = 1, \ldots, N_c$: the PDF estimate.

Remark that the problem of computing the PDF value at each collocation point $\rho_i, i = 1, \ldots, N_c$ has been transformed into a problem of computing the unknown Lagrange multipliers λ_0, $\lambda_{1,1}, \ldots, \lambda_{N_k,N_m}$. In other words, the size of the problem is now that of the physical domain subset ($\mathcal{S}$) times the number of moments instead of that of the stochastic domain. This is therefore computationally cheaper as long as the physical subset size remains sufficiently small, an issue that is addressed in the next section.

2.4.2 Analysis of the regularization error

In this section, we propose to justify some aspects of the proposed strategy. The true measure on the stochastic domain is $\mathcal{P}_e$, absolutely continuous with respect to Lebesgue measure. The associated probability density is ρ_e. The density which maximizes the entropy under the moment constraints is denoted by ρ^* and the actual approximation is ρ. There are two main contributions to the error: the first one is related to the entropic regularization, and the second one is due to the approximation of the constrained optimization problem. The latter is controlled by the norm of the error in

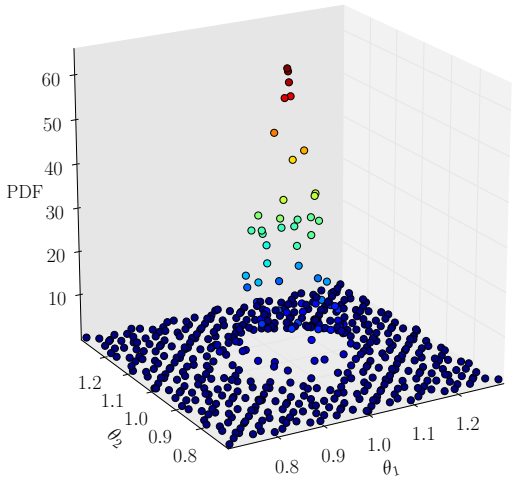

Figure 2: Solution of the moment-matching method: joint PDF of two parameters.

the dual variable approximation, as shown in Lemma 1. In what follows, the regularization error is studied.

The hypotheses under which this analysis is performed are the following: the observable is $g(\mathbf{x}, \boldsymbol{\theta}) \in H^1(\mathcal{D} \times \Theta) \cap L^\infty(\mathcal{D} \times \Theta)$. We remind that we assume that the g takes values in $[0, 1]$. The standard $L^2(\mathcal{D} \times \Theta)$ scalar product is denoted by $\langle u, v \rangle$, and the norms are defined accordingly. The scalar product in the physical and in the stochastic space will be denoted by $\langle u, v \rangle_{\mathbf{x}}$ and $\langle u, v \rangle_{\boldsymbol{\theta}}$ respectively.

The regularization error is studied in the case where an infinite number of moments exists. A first Lemma is presented to prove under which condition the total residual on the moments is L^2 summable, and then an identifiability condition for the inverse problem is derived.

Lemma 2. *Let* $\|v\|_{L^p(\mathcal{D}\times\Theta),\rho_e} = \left(\int_{\mathcal{D}} \int_{\Theta} v^p \mathrm{d}\mathbf{x}\rho_e \mathrm{d}\boldsymbol{\theta}\right)^{1/p}$ *be the* L^p *norm. If there exist* $C, \delta > 0$ *such that* $\|g\|_{L^p(\mathcal{D}\times\Theta),\rho_e} \leq \left(\frac{C}{p^{1+\delta}}\right)^{1/p}$, *then* $\sum_{m=1}^{\infty} \|\mu_m^{\rho_e}\|^2_{L^2(\mathcal{D})} < +\infty$.

Proof. The Jensen inequality gives:

$$\|\mu_m^{\rho_e}\|^2_{L^2(\mathcal{D})} = \int_{\mathcal{D}} \left(\int_{\Theta} g^m \rho_e \, d\boldsymbol{\theta} \right)^2 d\mathbf{x} \leq \|g^m\|^2_{L^2(\mathcal{D}\times\Theta),\rho_e}. \tag{22}$$

The norm can be rewritten as follows:

$$\|g^m\|^2_{L^2(\mathcal{D}\times\Theta),\rho_e} = \int_{\mathcal{D}} \int_{\Theta} g^{2m} \, d\mathbf{x} \, \rho_e \mathrm{d}\boldsymbol{\theta} = \|g\|^{2m}_{L^{2m}(\mathcal{D}\times\Theta),\rho_e} \leq \frac{C}{(2m)^{1+\delta}}, \tag{23}$$

and thus:

$$\sum_{m=1}^{\infty} \|\mu_m^{\rho_e}\|^2_{L^2(\mathcal{D})} \leq \sum_{m=1}^{\infty} \frac{C}{(2m)^{1+\delta}} < +\infty. \tag{24}$$

□

Let us assume that both the exact density ρ_e and the entropic regularization ρ^* satisfy the hypotheses of Lemma 2. Upper and lower bounds for the L^2 error $\varepsilon := \rho_e - \rho$ can be found. Consider that, by linearity: $\int_{\mathcal{D}} g^m \varepsilon = \mu_m^{\rho_e} - \mu_m^{\rho^*} = \delta\mu_m$. The result is summarized in the following proposition.

Proposition 1. *Let ρ_e, ρ^* satisfy the hypotheses of Lemma 2.*

Let $\gamma = \inf_{\|v\|_{L^2(\mathcal{D}\times\Theta)}=1} \left[\sum_{m=1}^{\infty} \langle g^m, v\rangle^2\right]$; let $\beta = \sum_{m=1}^{\infty} \|g^m\|^2_{L^2(\mathcal{D}\times\Theta)}$. Then:

$$\frac{\sum_{m=1}^{\infty} \|\delta\mu_m\|^2_{L^2(\mathcal{D})}}{\beta} \leq \|\varepsilon\|^2_{L^2(\Theta)} \leq \frac{\sum_{m=1}^{\infty} \|\delta\mu_m\|^2_{L^2(\mathcal{D})}}{\gamma}. \tag{25}$$

Proof. The Cauchy-Schwarz inequality implies:

$$\sum_{m=1}^{\infty} \|\delta\mu_m\|^2_{L^2(\mathcal{D})} = \sum_{m=1}^{\infty} \int_{\mathcal{D}} \left(\int_{\Theta} g^m \varepsilon \, d\boldsymbol{\theta}\right)^2 \mathrm{d}\mathbf{x} \leq \sum_{m=1}^{\infty} \int_{\mathcal{D}} \int_{\Theta} \|g^m\|^2_{L^2(\Theta)} \|\varepsilon\|^2_{L^2(\Theta)} \mathrm{d}\mathbf{x}, \tag{26}$$

The error norm does not depend on the physical space coordinates and thus:

$$\sum_{m=1}^{\infty} \int_{\mathcal{D}} \int_{\Theta} \|g^m\|^2_{L^2(\Theta)} \|\varepsilon\|^2_{L^2(\Theta)} \mathrm{d}\mathbf{x} \leq \left(\sum_{m=1}^{\infty} \|g^m\|^2_{L^2(\mathcal{D}\times\Theta)}\right) \|\varepsilon\|^2_{L^2(\Theta)} = \beta \|\varepsilon\|^2_{L^2(\Theta)}. \tag{27}$$

Analogously, the upper bound for the error is proved:

$$\sum_{m=1}^{\infty} \|\delta\mu_m\|^2_{L^2(\mathcal{D})} = \sum_{m=1}^{\infty} \int_{\mathcal{D}} \left(\int_{\Theta} g^m \varepsilon \mathrm{d}\boldsymbol{\theta}\right)^2 \mathrm{d}\mathbf{x}, \tag{28}$$

$$\geq \sum_{m=1}^{\infty} \inf_{\|v\|_{L^2(\mathcal{D}\times\Theta)}=1} \left[\langle g^m, v\rangle^2\right] \|\varepsilon\|^2_{L^2(\theta)}, \tag{29}$$

that can be deduced by considering that g^m can be expressed on a dense tensorized complete orthonormal basis of $L^2(\mathcal{D}) \otimes L^2(\Theta)$. □

The condition for the error to be bounded, namely $\gamma > 0$, can be seen also as an identifiability condition for the problem and it is verified when the set of function g^m is a complete basis of the space. The result of the following Lemma shows a meaningful case in which the density is not identifiable and the error is unbounded.

Lemma 3. *Let the stochastic domain be the box* $\Theta = \Theta_1 \times \dots \Theta_d$. *Let* $\mathcal{D}_1 \subseteq \mathcal{D}$ *an open subset of the physical domain where the observable does not depend on* θ_i, *i.e. for which* $\partial_{\theta_i} g = 0$. *Then* $\gamma = 0$.

Proof. The proof is done in a constructive way, by building a function v which is of unitary norm, making the scalar product with all the g^m vanish. Let $v = f_1(\theta_i) f_2(\theta_{j\neq i}) f_3(\mathbf{x})$ such that $\int_\Theta f_1 \mathrm{d}\boldsymbol{\theta} = 0$ and $f_3(\mathbf{x}) = 0$ *on* $\mathcal{D}/\mathcal{D}_1$. For all h,

$$\int_{\mathcal{D}} \int_\Theta g^m v \ \mathrm{d}\boldsymbol{\theta} \ \mathrm{d}\mathbf{x} = \int_{\mathcal{D}_1} \int_\Theta g^m f_1 f_2 \mathrm{d}\boldsymbol{\theta} f_3(\mathbf{x}) \ \mathrm{d}\mathbf{x}, \tag{30}$$

since f_3 vanishes outside $\mathcal{D}_1$. Then, since the observable g does not depend on θ_i,

$$\int_{\mathcal{D}_1} \int_\Theta g^m f_1 f_2 \mathrm{d}\boldsymbol{\theta} f_3(\mathbf{x}) \mathrm{d}\mathbf{x} = \int_{\mathcal{D}_1} \left(\int_{\Theta_i} f_1 \mathrm{d}\theta_i \right) \left(\int_{\Theta/\Theta_i} g^m f_2(\theta_{j\neq i}) \mathrm{d}\theta_j \right) f_3(\mathbf{x}) \mathrm{d}\mathbf{x} = 0. \tag{31}$$

□

The result of this Lemma sheds some light onto the identifiability of the inverse problem. In particular, the problem is ill-posed whenever there are regions in the physical space in which the observable does not depend on one or more parameters. A way to overcome this is to reduce the physical domain by excluding the regions (*i.e.* the DOFs) where the observable is not sensitive to the parameters.

2.5 Physical DOFs reduction: clustered sensitivities (CS) algorithm

As explained before, the dual variable formulation of the optimization problem transfers the resolution effort onto the solution of a system whose size is the number of DOFs in the physical domain times the number of moments. However, in many practical applications, as for instance when models are described by PDEs, the number of DOFs used to discretize the solution in the physical domain is large, making the Hessian matrix inversion computationally intensive. Aside from the sheer computational cost of linear algebra operations, dealing with many large simulations – say thousands of simulations counting millions of DOFs – poses undeniable issues in terms of storage capacity and Input/Output computer operations. The main idea to reduce the computational cost is to retain only the subsets of the physical domain in which the observable conveys more information about the variability of the parameters. Consider for instance a region in which the observable does not vary, or its variation amplitude is lower than the noise level: then, matching the moments in this region will certainly not convey any meaningful information about the parameters. Even worse, it may increase the Hessian condition number and degrade the overall accuracy of the

method. It may also happen that part of the data is redundant, meaning that the observable exhibits the same variations with respect to the parameters in two different DOFs. In this section, we propose an algorithm that selects a subset $\mathcal{S}$ of the full set of DOFs $\mathcal{D}$. This subset is then used in the OMM inverse procedure described before. Notice that we are not interested in building a low-dimensional surrogate model with fewer outputs. On the contrary, we aim at developing a non-intrusive approach where we only choose to discard some outputs of the high fidelity model. To do that, we propose the following gradient-based algorithm which is rooted in the global sensitivity analysis of the model.

2.5.1 The SGM matrix

For each $\mathbf{x}_j$, we consider an approximation of the exact SGM matrix (defined in (6)) as follows:

$$\mathbf{C}^j \simeq \beta \sum_{i=1}^{N_c} \left[\nabla_{\boldsymbol{\theta}} g(\mathbf{x}_j, \boldsymbol{\theta}_i)\right] \left[\nabla_{\boldsymbol{\theta}} g(\mathbf{x}_j, \boldsymbol{\theta}_i)\right]^T \rho_i,$$

where $\nabla_{\boldsymbol{\theta}} g(\mathbf{x}_j, \boldsymbol{\theta}_i)$ is a vector of size n_p whose components are the derivatives of g with respect to each parameter at a given $\mathbf{x}_j$ and a given parameter sample $\boldsymbol{\theta}_i$. $\mathbf{C}^j$ is a n_p-by-n_p matrix containing the sensitivity information of the observable with respect to the input parameters at $\mathbf{x}_j$. It may also be seen as the uncentered covariance matrix of the gradient of the observable with respect to the uncertain parameters.
In this work, the gradient $\nabla_{\boldsymbol{\theta}} g$ is approximated by using local polynomial approximations. Other well-known methods exist, such as adjoint equations [25] or automatic differentiation [26], but they will not be discussed here. For each sample $\boldsymbol{\theta}_i$ in the stochastic space, its K nearest neighbors are found and their indices are denoted by i_k, $k = 1, \ldots, K$. An implementation of the `k-NN` algorithm (using k-d trees) from the Scikit-learn library [27] was used for an efficient search of the nearest neighbors. The method consists in fitting a polynomial model to the K values of the observable $g_{i_k,j}$, $k = 1, \ldots, K$. Given a set of linearly independent polynomials $\{P_l(\boldsymbol{\theta})\}_{l=1,\ldots,n_l}$, the collocation matrix Φ_i reads:

$$\Phi_i = \begin{pmatrix} P_1(\boldsymbol{\theta}_{i_1}) & \cdots & P_{n_l}(\boldsymbol{\theta}_{i_1}) \\ \vdots & \ddots & \vdots \\ P_1(\boldsymbol{\theta}_{i_K}) & \cdots & P_{n_l}(\boldsymbol{\theta}_{i_K}) \end{pmatrix}.$$

The local polynomial model is obtained by solving the following linear system:

$$\boldsymbol{\Phi}_i \mathbf{q} = \mathbf{y}_{i,j},$$

where $\mathbf{y}_{i,j} = \begin{pmatrix} g_{i_1,j} & \cdots & g_{i_K,j} \end{pmatrix}^T$ and $\mathbf{q}$ is the vector of unknowns of size n_l. For stability reasons, K must be greater than n_l and therefore the system is solved in the least-squares sense. In practice, we used a basis of local multivariate quadratic monomials so that $n_l = \frac{n_p^2+3n_p+2}{2}$. The number of nearest neighbors is set to $K = n_l + 2$. Once $\mathbf{q}$ is computed, one obtains the following approximation of the gradient:

$$\nabla_{\boldsymbol{\theta}} g(\mathbf{x}_j, \boldsymbol{\theta}_i) \simeq \sum_{l=1}^{n_l} q_l \nabla_{\boldsymbol{\theta}} P_l(\boldsymbol{\theta}_i). \tag{32}$$

In what follows, this approximation of $\nabla_{\boldsymbol{\theta}} g(\mathbf{x}_j, \boldsymbol{\theta}_i)$ is denoted by $\mathbf{d}_{i,j}$. We now have an easily computable approximation $\hat{\mathbf{C}}^j$ of the SGM:

$$\hat{\mathbf{C}}^j = \beta \sum_{i=1}^{N_c} \mathbf{d}_{i,j} \mathbf{d}_{i,j}^T \rho_i, \tag{33}$$

which is symmetric and positive semidefinite so its eigenvalues are real and non-negative. Note that the approximation in (33) is computed using the Sobol sequence quadrature rule and the same simulation set $\{g_{i,j}\}$ as previously computed. This means that no additional model evaluation is required.

2.5.2 Parameter space dominant directions

The eigenvalues of the SGM play an important role in the classification of the DOFs. For a given $\mathbf{x}_j$ the eigenvalues are denoted by $\eta_1^j, \ldots, \eta_{n_p}^j$, in descending order. The corresponding eigenvectors, denoted by $\mathbf{e}_1^j, \ldots, \mathbf{e}_{n_p}^j$, form an orthonormal basis of the parameter space. The vector $\mathbf{e}_1^j$ corresponds to the direction (in the parameter space) of maximum variation, on average, of g at $\mathbf{x}_j$. Its associated eigenvalue η_1^j corresponds to the mean-squared directional derivative of the observable along the direction $\mathbf{e}_1^j$ [12, Lemma 3.1]. For instance, if there are two input parameters θ_1 and θ_2, then $\mathbf{e}_1^j = (1, 0)$ means that the observable variation of g at $\mathbf{x}_j$ is mostly due, on average, to variations of θ_1. Each $\mathbf{x}_j$ is therefore associated with a dominant direction in the parameter space $\mathbf{e}_1^j$ and its corresponding eigenvalue η_1^j. We are now able to address the initial problem: on the one hand, the DOFs where the variation of the observable is not significant are characterized by a low first eigenvalue. A threshold on η_1^j may be applied to remove the DOFs where the observable variation amplitude is lower than the noise level. On the other hand, the DOFs that are redundant from the observable point of view are characterized by "similar" dominant directions. This notion of similarity will be introduced hereafter. Knowing this, we propose to divide the set of $N_{\mathbf{x}}$ dominant directions into N_k clusters using an agglomerative hierarchical clustering algorithm. This algorithm consists in clustering vectors according to a given similarity function. First, each vector is associated

with its own cluster and pairs of similar clusters are iteratively merged. We refer to [28] for an overview of such algorithms. In the present work, we used the Scikit-learn library [27] which provides a Python implementation of an agglomerative hierarchical algorithm that accepts user-defined similarity functions. The similarity function between two (unit-norm) vectors is defined as follows:

$$s(\mathbf{u}, \mathbf{v}) = |\mathbf{u} \cdot \mathbf{v}| \,,$$

i.e. the cosine of the angle between $\mathbf{u}$ and $\mathbf{v}$. Once the $N_{\mathbf{x}}$ DOFs of the full physical set are divided into N_k clusters, the ones with maximum trace of $\hat{\mathbf{C}}^j$ are chosen as their cluster representatives. The subset $\mathcal{S}$ is then formed by the N_k representatives.

Remark 1. *The agglomerative clustering guarantees that the sequence of selected subsets is nested. This means that if $\mathcal{S}^{(n)}$ and $\mathcal{S}^{(n+1)}$ respectively count n and $n+1$ elements, then they have n elements in common. From a practical viewpoint, the full sequence of clusters can be computed once so that there is no additional cost linked to the clustering when N_k increases. Furthermore, in our simulations, we noticed that the residual had a smoother behavior as N_k increases compared to other clustering techniques.*

The output of the CS algorithm is a nested sequence of subsets $\mathcal{S}^{(1)} \subset \ldots \subset \mathcal{S}^{(N_{\mathbf{x}})}$ and we denote by N_k the cardinality of a given subset $\mathcal{S}$.

Remark 2. *In the works by Constantine [29] and Russi [30], where the term "active subspace" was introduced, the matrix $\mathbf{C}$ is used to reduce the parameter space dimension. It is particularly efficient when dealing with complex models counting a very large number of parameters while only a few directions in the parameter space are responsible for the observed variability [13]. In our case it is used to reduce the number of DOFs in the discretized physical domain. However, the interpretation of the SGM eigenvalues and eigenvectors in terms of the sensitivity of the model is the same. In the papers by Streif et al. [31] and Himpe & Ohlberger [32], a similar Gramian matrix is used to assess the observability and controllability of linear and nonlinear systems. Though quite different from the CS analysis, their approach is another illustration of the interpretation of Gramian matrices in terms of sensitivity analysis.*

Algorithm 2: Clustered Sensitivities algorithm.

Input:

- $g_{i,j}$, $i = 1, \ldots, N_c$, $j = 1, \ldots, N_{\mathbf{x}}$: simulation subset.
- ρ: PDF estimate.

for $j = 1$ **to** $N_{\mathbf{x}}$ **do**
  for $i = 1$ **to** N_c **do**
    Compute $\mathbf{d}_{ij}$ using (32)
  end
  Compute $\hat{C}^{(j)}$ using (33);
  Compute first eigenvector $\mathbf{e}_1^j$ and eigenvalues trace $t(\mathbf{x}_j) = \sum_k \eta_k^j$;
end
Compute sequence of clusters for $j = 1, \ldots, N_{\mathbf{x}}$ using similarity function s;
for $N_k = 1$ **to** $N_{\mathbf{x}}$ **do**
  $\mathcal{S}^{(N_k)} = \{\}$;
  for $k = 1$ **to** N_k **do**
    Select representative $\mathbf{x}_k$ of cluster C_k as: $\arg\max_{\mathbf{x} \in C_k} \{t(\mathbf{x})\}$;
    Append $\mathbf{x}_k$ to $\mathcal{S}^{(N_k)}$;
  end
end

Output: Subset sequence: $S^{(1)} \subset \ldots \subset \mathcal{S}^{(N_{\mathbf{x}})}$.

2.6 Visualization and interpretation of the results

The output of the proposed algorithm is the estimated PDF values at the collocation points: $\rho(\boldsymbol{\theta}_i), i = 1, \ldots, N_c$. Although a direct visualization of the PDF is possible (see Fig. 2), it becomes irrelevant if the number of parameters is greater than two. Therefore it may be convenient to consider the marginal density of the k^{th} parameter, defined as follows:

$$z_k(x) = \int_{\theta_l, l \neq k} \cdots \int \rho(\theta_1, \ldots, x, \ldots, \theta_{n_p}) \prod_{l \neq k} \mathrm{d}\theta_l. \tag{34}$$

An approximation of Eq.(34) may be computed using the discrete PDF values and the corresponding quadrature rule.
In the numerical tests we illustrate the proposed algorithm with synthetic data, meaning the true PDF ρ^* of the parameters is known. There are several ways to compare the estimated and true PDFs such as the 2–norm of their difference. However, for density functions, it is more natural to consider the Kullback-Leibler (KL) divergence, introduced in Lemma 1. Here we use a

discrete approximation of the symmetric KL divergence, defined as follows:

$$KL(\rho|\rho^*) = \frac{1}{2}\left(\varphi(\rho,\rho^*) + \varphi(\rho^*,\rho)\right),$$
$$\text{where} \quad \varphi(u,v) = \beta \sum_{i=1}^{N_c} u_i \log(u_i/v_i).$$

It is also possible to compute the parameters moments with the estimated PDF and compare them with their true values.

2.7 Main algorithm

The proposed inverse procedure consists in combining the OMM and CS algorithms (see Alg. (3)). To assess the convergence of the procedure, we use the *global* moment residual $\mathbf{R} \in \mathbb{R}^{N_\mathbf{x} \times N_m}$ defined as follows:

$$R_{j,m} = \beta \sum_{i=1}^{N_c} g_{i,j}^m \rho_i - \tilde{\mu}_{j,m}, \ j = 1, \ldots, N_\mathbf{x}, \ m = 1, \ldots, N_m.$$

Note that while the OMM algorithm is designed to cancel out the residual $\mathbf{r}$ defined on a subset $\mathcal{S}$ of $\mathcal{D}$, $\mathbf{R}$ is defined on the full DOFs set $\mathcal{D}$.
One iteration of the main algorithm consists in progressively adding DOFs to the subset $\mathcal{S}$ using the CS algorithm and applying the OMM algorithm for each $\mathcal{S}$ until stagnation of the 2-norm of the residual, $\|\mathbf{R}\|_2$. Then, the SGM is updated with the new PDF estimate and another iteration is done. The main algorithm stops when no improvement of $\|\mathbf{R}\|_2$ is observed. The total number of iterations is later referred to as n_{iter}. Figure 13 shows an example of the dependence between the global residual norm $\|\mathbf{R}\|_2$ and the cardinality of the subset $\mathcal{S}$.

Algorithm 3: Main algorithm.

Input:

- Corrected experimental moments: $\tilde{\mu}_{j,m}$, $j = 1, \ldots, N_{\mathbf{x}}$, $m = 1, \ldots, N_m$;
- Number of stochastic samples: N_c;
- Tolerance parameters : α, ϵ_{Newton};

Step 1:

- Build the simulation set $\{g_{i,j}\}$.

Initial guess $\rho^{(0,0)}$: uniform distribution over Θ ;
$j = 1$, $N_k = 0$;
while $\underline{\|\mathbf{R}^{(j-1,N_k)}\|_2 \text{ not converged}}$ **do**
  Apply CS procedure with $\rho^{(j-1,0)}$ (**Step 2**);
    $\rightarrow$ nested subsets sequence $\mathcal{S}^{(j,1)} \subset \ldots \subset \mathcal{S}^{(j,N_{\mathbf{x}})}$;
  $n = 1$;
  while $\underline{\|\mathbf{R}^{(j-1,n)}\|_2 \text{ not converged}}$ **do**
    Apply OMM procedure with $\mathcal{S}^{(j,n)}$ (**Step 3**);
      $\rightarrow \rho^{(j,n+1)}$;
    $n \leftarrow n + 1$;
  end
  $j \leftarrow j + 1$;
  $N_k \leftarrow n$;
end
$n_{iter} = j$

Table 2.7 presents an overview of the computational cost of the whole procedure. In practice, this cost is strongly dominated by the construction of the simulation set (step one), each model evaluation having a high cost C_{forward}. However, this step is embarrassingly parallelizable with respect to N_c. Step two is embarrassingly parallelizable both with respect to N_c and $N_{\mathbf{x}}$. In our implementation, the SGM computations are only parallelized with respect to $N_{\mathbf{x}}$. Step three is dominated by the cost of the Hessian pseudo-inverse computation. As it scales with $(N_m N_k)^3$, the need to reduce the number of DOFs in the physical space becomes obvious. The pseudo-inverse computation could also be parallelized but this was not done in our implementation. The cost of the pseudo-inverse is multiplied by n_{Newton}, the number of Newton iterations.

3 Numerical illustrations

We apply our strategy to the PDF estimation of parameters for two PDEs. Comparisons with existing methods are presented in the Appendix for a nonlinear ODE.

Table 1: Complexity of the inverse procedure. C_{forward} denotes the cost of one model evaluation.

Step	Complexity	Parallelizable
1. Simulation Set	$\mathcal{O}\left(N_c \times C_{\text{forward}}\right)$	massively w.r.t. N_c
2. Clustured Sensitivities	$\mathcal{O}\left(N_{\mathbf{x}} \times N_c \times n_p^3\right)$	massively w.r.t. N_c and $N_{\mathbf{x}}$
3. Observable Moment Matching	$\mathcal{O}\left(n_{\text{Newton}} \times (N_k \times N_m)^3\right)$	possible

3.1 Application to an elliptic PDE: the Darcy equations

In this section, we focus on the following two-dimensional PDE posed in the bounded domain $\mathcal{D} = [0,1] \times [0,1]$:

$$\begin{aligned} -\nabla \cdot (K \boldsymbol{\nabla} p) &= 0, & \mathbf{x} &\in \mathcal{D}, \\ p &= f, & \mathbf{x} &\in \Gamma_D, \\ K \boldsymbol{\nabla} p \cdot \mathbf{n} &= 0, & \mathbf{x} &\in \Gamma_N, \end{aligned}$$

where p is the fluid pressure, f a deterministic function defined on the boundary Γ_D and $\{\Gamma_D, \Gamma_N\}$ is a partition of $\partial\mathcal{D}$. In what follows, f will be set to 1 at the inlet and to 0 at the outlet (see Fig. 3). The Darcy model states that the fluid velocity is linked to the pressure as follows by $\mathbf{u} = -K\boldsymbol{\nabla} p$. We assume that the source of variability comes from the heterogeneous per-

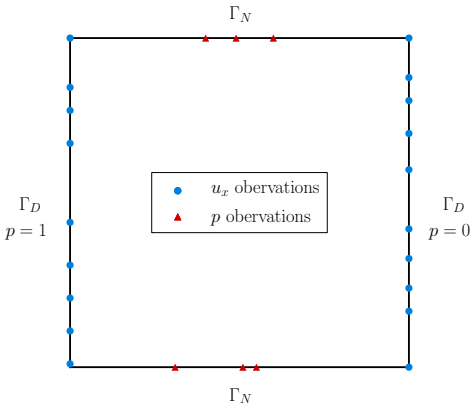

Figure 3: Schematic of the problem geometry and location of 25 sensors automatically selected by the CS procedure (out of 400 available sensors).

meability field $K(\mathbf{x})$. Using a similar example found in [14], we assume that the spatial variation in the permeability field follows an exponential correlation: $c(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{|\mathbf{x}_i - \mathbf{x}_j|}{b}\right)$, where b is the correlation length, set to

$b = 0.2$ in our case. From a physical viewpoint, this means that the porous medium is relatively smooth. Then, we choose to represent the random field K as a linear combination of the first 5 eigenmodes $\hat{K}_k$ of the correlation kernel $K(\mathbf{x}) = 1 + \sum_{k=1}^{5} \theta_k \hat{K}_k(\mathbf{x})$, where the θ_k are the random parameters. Figure 4 shows the eigenmodes of the correlation kernel and Figure 5 shows one realization of the random permeability field, along with the outputs of the model, namely the pressure field p and the horizontal velocity u_x.
The objective is to apply the proposed approach to recover the PDF of the permeability field expansion coefficients θ_k from observations of u_x and p on the boundaries. Retrieving the permeability in the domain by exploiting only boundary measurements is a particular case of the Calderón problem, which is a difficult and generally ill-posed inverse problem.

Numerical settings The observable is defined as follows: 200 sensors for u_x (resp. p) are uniformly distributed over the boundary Γ_D (resp. Γ_N) so that $N_{\mathbf{x}} = 400$. The synthetic data set is generated by evaluating the model for $N = 10^4$ samples of $\boldsymbol{\theta} = (\theta_1, \dots, \theta_5)$. The samples are drawn from an uncorrelated multivariate normal distribution of mean $\boldsymbol{\mu} = 2.5 \times 10^{-2} \times [1, 1, 1, 1, 1]$ and covariance matrix $\boldsymbol{\Sigma} = 3.3 \times 10^{-2} \times \mathbf{I}_5$. $N_c = 2^{14}$ collocation points are generated using the Sobol sequence over the domain $\Theta = [-0.2, 0.2]^5$, the number of moments to be matched is set to $N_m = 3$ and the tolerance parameter is set to $\alpha = 1 \times 10^{-3}$. The PDE model is solved using the `FreeFem++` [33] finite element software. A different discretization is used for both sets. For the synthetic dataset, the model is solved on a fine grid of 23550 triangles. For the simulation set, the model is solved on a coarse mesh of 944 triangles. In addition, a Gaussian zero-mean noise of amplitude 5% is added to the sensors measurements.

Results The proposed inverse procedure is applied and convergence is reached at $n_{iter} = 2$ and $N_k = 25$. Figure 3 shows the position of final selected DOFs. Note that points were automatically selected on each boundary even though this was not imposed in the CS procedure. Figure 6 shows the estimated marginals of the five parameters along with their exact distributions. Table 3.1 summarizes the estimated parameters statistics to be compared to their exact values. The means are in good agreement, with an error of the order of 1%. The standard deviations feature a higher error, especially for the fifth mode parameter. The sources of error are diverse. The mesh used to generate the simulation dataset is coarser than the one used for the synthetic dataset. This induces a higher numerical diffusion. Moreover, the added noise may also contribute to the error, especially for the higher order modes coefficients.

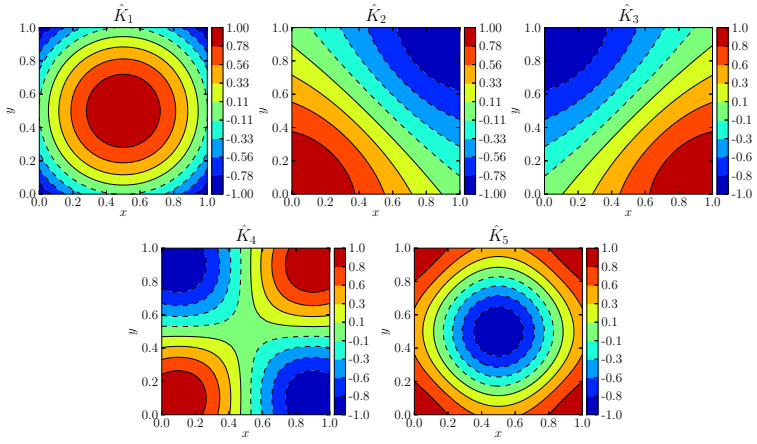

Figure 4: Contours of the first 5 eigenmodes of the correlation kernel.

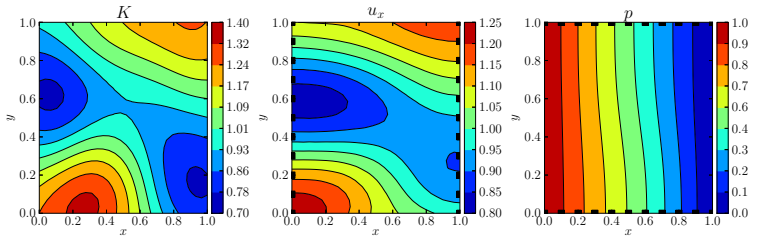

Figure 5: Contours of one realization of the Darcy model. Left: permeability, center:horizontal velocity, right: pressure. Black squares indicate where the velocity and pressure fields are observed.

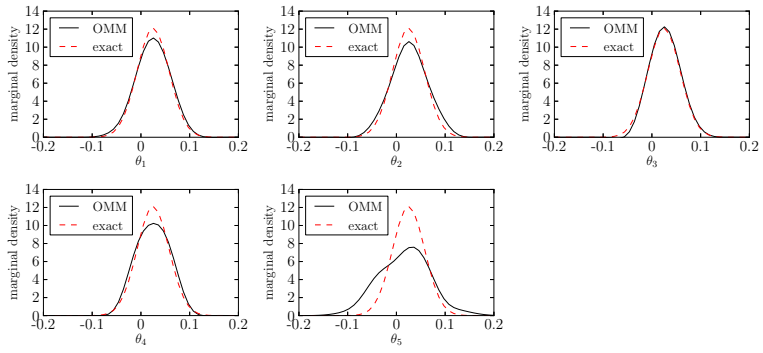

Figure 6: Marginal densities of five parameters estimated using the OMM method.

Table 2: Darcy model

Statistics	mean			std.		
Parameter	exact	OMM	rel. err.(%)	exact	OMM	rel. err.(%)
θ_1	**2.48e-02**	2.49e-02	0.5	**3.33e-02**	3.08e-02	7.4
θ_2	**2.46e-02**	2.49e-02	1.2	**3.35e-02**	3.39e-02	1.3
θ_3	**2.56e-02**	2.53e-02	0.9	**3.36e-02**	2.82e-02	16
θ_4	**2.55e-02**	2.58e-02	1.1	**3.33e-02**	2.90e-02	13
θ_5	**2.49e-02**	2.52e-02	1.4	**3.31e-02**	5.00e-02	51

3.2 Application to a parabolic PDE: the FKPP equation

In this section, we illustrate our strategy with the FKPP equation, originally introduced by Fisher [34], later revisited by Kolmogorov, Petrovskii and Piskunov. It is a nonlinear reaction-diffusion equation defined by:

$$\begin{aligned}
\frac{\partial u}{\partial t} - \nu \Delta u &= Ru(1-u) + f(\mathbf{x},t), \qquad \mathbf{x} \in [0,1]^2, t \in [0,T], \\
\boldsymbol{\nabla} u \cdot \mathbf{n} &= 0, \qquad \mathbf{x} \in \partial[0,1]^2, t \in [0,T], \\
u(\mathbf{x}, t=0) &= 0, \qquad \mathbf{x} \in [0,1]^2,
\end{aligned}$$

where u is a time and space dependent variable, R is the reaction parameter and ν the diffusion field, here considered uniform and constant equal to 10^{-3}. Provided that $R/\nu \gg 1$ and given an *ad hoc* source term f, the FKPP equation admits travelling waves solutions. In practice, u exhibits a propagation front across which u switches from 0 to 1. It is often considered as the simplest PDE model presenting this feature. This has motivated the use of FKPP for a large variety of applications (examples include population dynamics, tumor growth and fire propagation). Here f, later referred to as the stimulation, was designed so that such a propagation would be observable: if $(x-x_0)^2 + (y-x_0)^2 \leq r_0^2, \quad t \in [t_0, t_0+\delta_0]$ then $f(\mathbf{x},t) = I_0$, otherwise $f(\mathbf{x},t) = 0$, where (x_0, y_0) are the coordinates of the stimulation, $I_0 = 1.0$ its amplitude, $r_0 = 3 \times 10^{-2}$ its radius and $\delta_0 = 5$ its duration. The total duration of the simulation is set to $T = 20$. Figure 7 shows an instance of the FKPP model output. The contour plots of u exhibit the propagating front (left and right) while the time dependence of u at a given location exhibits a logistic shape. The source of variability is assumed to come from the reaction parameter R and from the stimulation coordinates x_0 and y_0: $R = \bar{R}\theta_1$, $x_0 = \theta_2$, $y_0 = \theta_3$, where $\bar{R} = 10$.

Numerical settings The observations are the values of u at $N_t = 200$ time steps times $N_h = 81$ sensors locations, uniformly distributed over

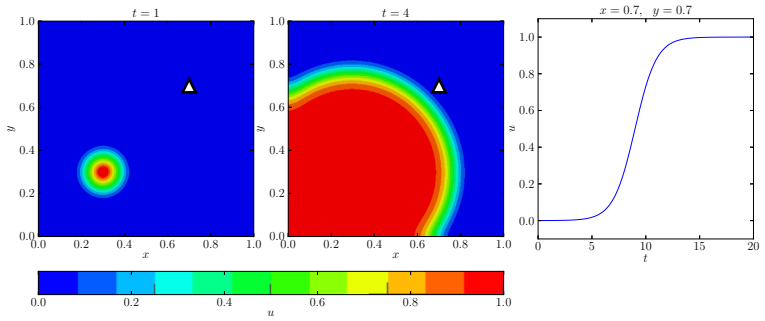

Figure 7: Solution of the FKPP model at different times (left, center) and time-dependent solution at a given point (right).

$[0,T]\times[0,1]^2$ so that $N_{\mathbf{x}} = 16200$. The synthetic dataset is generated by evaluating the model for $N = 10^3$ samples of $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$. The samples are drawn from a multivariate normal distribution of mean $\boldsymbol{\mu} = [0.55, 0.55, 0.50]$ and covariance matrix $\boldsymbol{\Sigma} = \sigma^2 \times \mathbf{I}_3$, where $\sigma = 0.1$. $N_c = 2048$ collocation points are generated using the Sobol sequence over the domain $\Theta = [0.1, 1.0]^3$, the number of moments to be matched is set to $N_m = 3$ and the tolerance parameter is set to $\alpha = 5 \times 10^{-3}$. The PDE model is solved using an in-house software implementing the finite element method. Time integration is performed using the Strang [35] splitting scheme with fixed time step. Its application to a similar reaction diffusion model is detailed in [36]. Again, a different discretization is used for both simulation sets. The simulations used to generate the synthetic data are run on a mesh counting 40328 elements whereas the simulations used to solve the inverse problem are run on a coarse mesh counting 11478 elements. In addition, a Gaussian zero-mean noise of amplitude 5% is added to the sensors measurements.

Physical domain reduction This test case where the observable depends on time and space is a good illustration of the crucial need for a DOF selection procedure. Indeed, in this setting, $N_{\mathbf{x}} \simeq 10^4$ which makes the inverse problem both ill-conditioned and computationally intensive. In this example, it is particularly interesting to interpret the results of the CS procedure. Figure 8 shows the contours of the components of the SGM first eigenvector $\mathbf{e}_1^j$ (dominant direction) multiplied by its associated eigenvalue η_1^j over the physical domain $\mathcal{D}$. Each column corresponds to one component of $\mathbf{e}_j$, *i.e.* to one parameter, and each row to a different time. The space-time areas of interest now appear clearly. For small times, the parameters are the most identifiable in the vicinity of the domain center. As the front propagates outwards, the important areas are located near the domain boundaries.

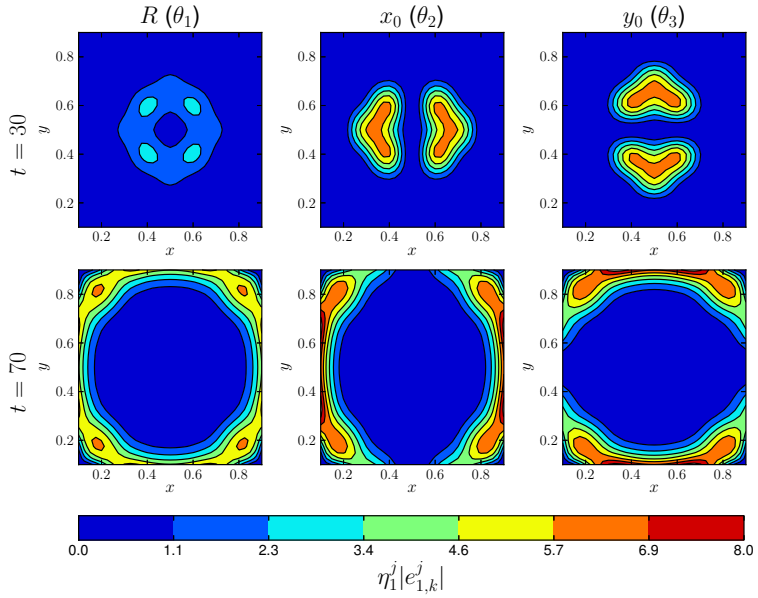

Figure 8: Measure of the sensitivity ($\eta_j |e_{j,k}|$) over the spatial domain at different times and for each parameter θ_k.

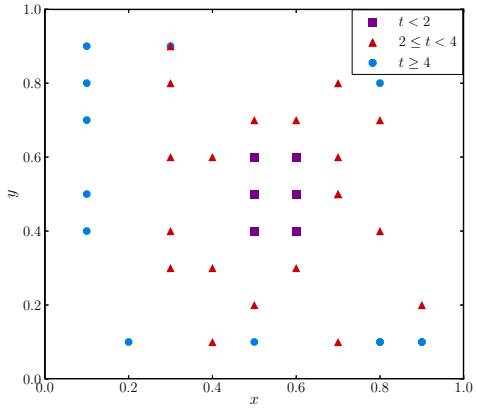

Figure 9: Representation of the 48 selected sensors locations and times.

Results The proposed inverse procedure is applied and convergence is reached at $n_{iter} = 3$ and $N_k = 48$ DOFs are selected. Figure 9 shows the location and time of the selected sensors. Again, note that they are concentrated around the center of the domain for small times (the stimulation occurs, in average, near the center of the domain) and that they gradually spread outwards as time increases. Figure 10 shows the estimated marginals of the three parameters of interest and Table 3.2 summarizes the parameters estimated statistics. Again, the method yields reasonably accurate results considering the low number of model evaluations and the difficulty of the inverse problem. As explained in the previous test case, the errors in the standard deviations estimates stem from the noise and the mesh differences. Note however that there is also a positive bias in the estimation of the reaction parameter R. This is due to the fact that the Sobol simulations mesh is coarser than the synthetic simulations one, inducing a higher numerical diffusion. The higher value obtained for R is therefore the result of a compensation. This explanation was confirmed by using identical meshes for both Sobol and synthetic simulations.

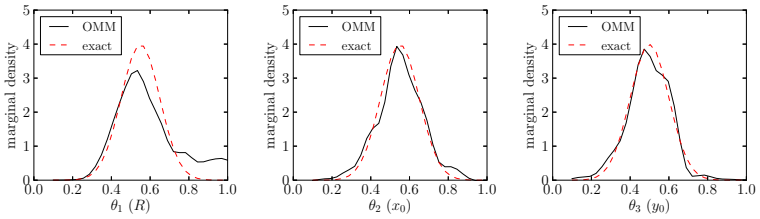

Figure 10: Marginal densities of three parameters of the FKPP model estimated using our strategy.

Table 3: Results for the FKPP equation.

Statistics	mean			std.		
Parameter	exact	OMM	rel. err.(%)	exact	OMM	rel. err.(%)
θ_1	**0.55**	0.59	5.8	**0.098**	0.16	66
θ_2	**0.55**	0.55	0.4	**0.105**	0.12	9.3
θ_3	**0.50**	0.50	0.5	**0.103**	0.11	8.0

4 Concluding remarks

We have developed a procedure to estimate the PDF of uncertain parameters from the knowledge of experimental moments of an observable. This

iterative procedure is based on two combined algorithms. The first one, the Observable Moment Matching (OMM) algorithm, computes an estimate of the parameters PDF using a given subset of the available model DOFs. It maximizes the PDF entropy under the constraints of matching the moments of the observable in the subset DOFs. The second one, the Clustered Sensitivies (CS) algorithm, selects a subset of the available model DOFs. The DOFs are clustered using a similarity measure and a representative for each cluster is chosen to maximize the sensitivity with respect to the parameters. Selecting a subset of N_k DOFs among the $N_{\mathbf{x}}$ available ones ensures a better-conditioned and less computationally expensive inverse problem solved in the OMM algorithm.
This approach has been compared to existing techniques on an ODE test case. While requiring much less model evaluations, our method has a similar accuracy. Then, it has been tested on more sophisticated cases involving an elliptic (resp. parabolic) PDE model with 5 (resp. 3) uncertain parameters. To conclude, we comment on details that have not been thoroughly investigated in this paper but still are worth mentioning. First, the choice of parameter box (or stochastic domain) is very important and conditions the overall success of the procedure. In our tests, we used a large box with respect to the exact PDF support and not centered on the exact mean to avoid any favorable bias. In the case of real experimental data, a reasonable strategy would be to first try a very large box and use the PDF estimate to recenter and rescale the box for a second run. Another strategy would be to locally refine the stochastic grid to capture the regions of interest. Applying different weights in the moment-matching constraints depending on the moment order has also not been investigated but could impact the precision of the method. One could use higher weights for the higher moment components or for certain DOFs. Finally, one possible use of the proposed approach could be to produce a cheap PDF estimation used as a prior for more expensive methods such as Bayesian inference.

References

[1] M. Christie, V. Demyanov, D. Erbas, Uncertainty quantification for porous media flows, Journal of Computational Physics 217 (1) (2006) 143–158.

[2] S. Sankaran, N. Zabaras, A maximum entropy approach for property prediction of random microstructures, Acta Materialia 54 (8) (2006) 2265–2276.

[3] J. Wang, N. Zabaras, A bayesian inference approach to the inverse heat conduction problem, International Journal of Heat and Mass Transfer 47 (17) (2004) 3927–3941.

[4] P.-S. Koutsourelakis, A multi-resolution, non-parametric, bayesian framework for identification of spatially-varying model parameters, Journal of computational physics 228 (17) (2009) 6184–6211.

[5] N. Hansen, A. S. Niederberger, L. Guzzella, P. Koumoutsakos, A method for handling uncertainty in evolutionary optimization with an application to feedback control of combustion, Evolutionary Computation, IEEE Transactions on 13 (1) (2009) 180–197.

[6] A. Alwan, N. R. Aluru, Improved statistical models for limited datasets in uncertainty quantification using stochastic collocation, Journal of Computational Physics 255 (2013) 521–539.

[7] J. A. Shohat, J. D. Tamarkin, The problem of moments, American Mathematical Society, 1943.

[8] J. Guilleminot, A. Noshadravan, C. Soize, R. Ghanem, A probabilistic model for bounded elasticity tensor random fields with application to polycrystalline microstructures, Computer Methods in Applied Mechanics and Engineering 200 (17) (2011) 1637–1648.

[9] E. Patelli, G. Schuëller, On optimization techniques to reconstruct microstructures of random heterogeneous media, Computational Materials Science 45 (2) (2009) 536–549.

[10] E. T. Jaynes, Information theory and statistical mechanics, Physical review 106 (4) (1957) 620.

[11] L. R. Mead, N. Papanicolaou, Maximum entropy in the problem of moments, Journal of Mathematical Physics 25 (8) (1984) 2404–2417.

[12] P. G. Constantine, Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies, SIAM, 2015.

[13] P. G. Constantine, M. Emory, J. Larsson, G. Iaccarino, Exploiting active subspaces to quantify uncertainty in the numerical simulation of the hyshot ii scramjet, Journal of Computational Physics 302 (2015) 1–20.

[14] N. Zabaras, B. Ganapathysubramanian, A scalable framework for the solution of stochastic inverse problems using a sparse grid collocation approach, Journal of Computational Physics 227 (9) (2008) 4697–4735.

[15] E. Kuhn, M. Lavielle, Maximum likelihood estimation in nonlinear mixed effects models, Computational Statistics & Data Analysis 49 (4) (2005) 1020–1038.

[16] D. Henrion, J. B. Lasserre, M. Mevissen, Mean squared error minimization for inverse moment problems, Applied Mathematics & Optimization 70 (1) (2014) 83–110.

[17] J. C. Wheeler, R. Gordon, Rigorous bounds for thermodynamic properties of harmonic solids, The Journal of Chemical Physics 51 (12) (1969) 5566–5583.

[18] C. E. Shannon, A mathematical theory of distribution, Bell System Technical Journal 27 (1948) 623.

[19] M. Massot, F. Laurent, D. Kah, S. De Chaisemartin, A robust moment method for evaluation of the disappearance rate of evaporating sprays, SIAM Journal on Applied Mathematics 70 (8) (2010) 3203–3234.

[20] E. Van der Straeten, C. Beck, Superstatistical distributions from a maximum entropy principle, Physical Review E 78 (5) (2008) 051101.

[21] I. M. Sobol, Uniformly distributed sequences with an additional uniform property, USSR Computational Mathematics and Mathematical Physics 16 (5) (1976) 236–242.

[22] C. Lemieux, Monte carlo and quasi-monte carlo sampling, Springer Science & Business Media, 2009.

[23] B. Ganapathysubramanian, N. Zabaras, Sparse grid collocation schemes for stochastic natural convection problems, Journal of Computational Physics 225 (1) (2007) 652–685.

[24] H.-J. Bungartz, M. Griebel, Sparse grids, Acta numerica 13 (2004) 147–269.

[25] Y. Cao, S. Li, L. Petzold, R. Serban, Adjoint sensitivity analysis for differential-algebraic equations: The adjoint dae system and its numerical solution, SIAM Journal on Scientific Computing 24 (3) (2003) 1076–1089.

[26] A. Griewank, A. Walther, Evaluating derivatives: principles and techniques of algorithmic differentiation, Siam, 2008.

[27] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, Journal of Machine Learning Research 12 (2011) 2825–2830.

[28] G. Lance, W. Williams, A general theory of classificatory sorting strategies. 1 hierarchical systems, Computer Journal 9 (1967) 373–380.

[29] P. G. Constantine, E. Dow, Q. Wang, Active subspace methods in theory and practice: Applications to kriging surfaces, SIAM Journal on Scientific Computing 36 (4) (2014) A1500–A1524.

[30] T. M. Russi, Uncertainty quantification with experimental data and complex system models, Ph.D. Thesis.

[31] S. Streif, R. Findeisen, E. Bullinger, Relating cross gramians and sensitivity analysis in systems biology, in: Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems, Kyoto, Japan, 2006, pp. 437–441.

[32] C. Himpe, M. Ohlberger, Cross-gramian-based combined state and parameter reduction for large-scale control systems, Mathematical Problems in Engineering 2014.

[33] F. Hecht, New development in freefem++, J. Numer. Math. 20 (3-4) (2012) 251–265.

[34] R. A. Fisher, The genetical theory of natural selection: a complete variorum edition, Oxford University Press, 1930.

[35] G. Strang, On the construction and comparison of difference schemes, SIAM Journal on Numerical Analysis 5 (3) (1968) 506–517.

[36] J. Sundnes, G. T. Lines, A. Tveito, An operator splitting method for solving the bidomain equations coupled to a volume conductor model for the torso, Mathematical biosciences 194 (2) (2005) 233–248.

[37] A. Bueno-Orovio, E. M. Cherry, F. H. Fenton, Minimal model for human ventricular action potentials in tissue, Journal of theoretical biology 253 (3) (2008) 544–560.

[38] M. R. Davies, H. B. Mistry, L. Hussein, C. E. Pollard, J.-P. Valentin, J. Swinton, N. Abi-Gerges, An in silico canine cardiac midmyocardial action potential duration model as a tool for early drug safety assessment, American Journal of Physiology-Heart and Circulatory Physiology.

[39] N. Hansen, The cma evolution strategy: a comparing review, in: Towards a new evolutionary computation, Springer, 2006, pp. 75–102.

[40] A. P. Dempster, N. M. Laird, D. B. Rubin, Maximum likelihood from incomplete data via the em algorithm, Journal of the royal statistical society. Series B (methodological) (1977) 1–38.

[41] Lixoft, Monolix Software, Version 4.3.2, Orsay, France (2014).
URL `http://http://www.lixoft.eu`

[42] E. Grenier, V. Louvet, P. Vigneaux, Parameter estimation in non-linear mixed effects models with saem algorithm: extension from ode to pde., ESAIM: Mathematical Modelling and Numerical Analysis 48 (5) (2014) 1303–1329.

[43] F. Heiss, V. Winschel, Likelihood approximation by numerical integration on sparse grids, Journal of Econometrics 144 (1) (2008) 62–80.

Acknowledgements

This research was supported by a French Ministry of Higher Education and Research grant.

A Comparison with existing techniques for an ODE model

In this appendix, a nonlinear ODE model is introduced. It serves as a simple reference test case to both illustrate the method and compare its accuracy and cost with different existing techniques.

A.1 The MV model

The proposed numerical method is applied to an ODE model counting four state variables g, u, v, w which satisfy

$$\begin{cases} \partial_t g &= -J_1(g,u) - J_2(g,\theta_1,\theta_2) - J_3(g,v,w) \\ \partial_t u &= f_1(g,u) \\ \partial_t v &= f_2(g,v) \\ \partial_t w &= f_3(g,w) \end{cases} \tag{35a}$$

along with the initial conditions

$$g(0) = 0, \quad u(0) = 0, \quad v(0) = 1, \quad w(0) = 1. \tag{35b}$$

The J_i and f_i are nonlinear functions of the variables and of the input parameters. The proposed model was designed to replicate the electrical activity of a heart muscle cell. It is known as the Minimum Ventricular model and will be referred to as the MV model in what follows. For the sake of simplicity, it is not fully transcribed here but we refer the reader to the original paper by Bueno-Orovio *et al.* [37] for the detailed equations. Out of the numerous input parameters of the MV model, θ_1 and θ_2 were picked for the illustration of the method. In the original paper, these two parameters are respectively denoted by k_{so} and τ_{so1}. All remaining parameters are fixed to reference values found in [37]. Our observable is the state variable $g(t)$ which corresponds to the cell membrane potential. Note that the relationship between the observable and the input parameters is nonlinear.

A.2 Reference test case

Numerical settings The ODE is solved using a BDF3 scheme with adaptive time steps. The number of DOFs $N_{\mathbf{x}} = 334$ corresponds in this case to the number of steps used in the time integration. The synthetic data set is generated by evaluating the model in (35) for $N = 10^3$ samples of $\boldsymbol{\theta} = (\theta_1, \theta_2)$. The samples are drawn from an uncorrelated bivariate normal distribution of mean $\boldsymbol{\mu} = [1.1, 1.1]$ and covariance matrix $\boldsymbol{\Sigma} = 0.1^2 \times \mathbf{I}_2$. First, the noise level is set to 5% for the comparison study but its influence is investigated later in this section. The first N_m order moments are computed using (3) and stored for the inverse problem. Our strategy is applied to the

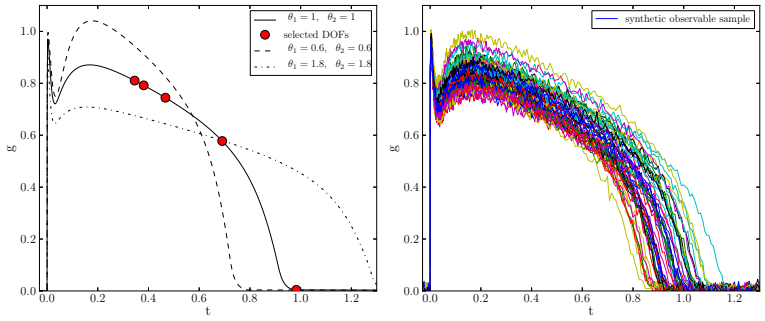

Figure 11: Solution of the ODE model for different values of the input parameters (left) and synthetic measurements (right). 5 time steps selected by the CS algorithm are indicated as red circles.

joint PDF estimation of the synthetic population θ_1 and θ_2. The stochastic domain $\Theta = [0.6, 1.8]^2$ is discretized using $N_c = 1024$ quadrature points from the Sobol sequence. It should be noted that the width of the stochastic domain is equal to 12σ and is not centered on μ. Taking a domain which is wide enough with respect to the exact PDF support, and not centered on the exact mean, is important if one wants to assess the accuracy of the method without any "favorable bias" induced by the choice of the stochastic domain bounds. Indeed, in practical cases, one does not have a precise knowledge on the exact means and standard deviations of the parameters distributions.

To investigate the effect of several hyper-parameters of the procedure, the number of global iterations in temporarily set to $n_{iter} = 1$. The CS procedure is applied with the initial guess $\rho^{(0,0)}$ being a uniform distribution over Θ. Figure 12 shows the SGM eigenvectors $\mathbf{e}_j = (e_1, e_2)_j$, $j = 1, \ldots, N_{\mathbf{x}}$. The size of the markers is proportional to the logarithm of the associated eigenvalues η_j. Since the eigenvectors are normalized, the points are scattered over the unitary circle. Each cluster is featured with a different color (here $N_k = 5$ so the points are divided into 5 clusters).

Influence of N_k Here we investigate the effect of N_k. The other hyper-parameters are fixed: $N_m = 3$ and $N_c = 512$. The CS procedure is applied for N_k varying from 2 (n_p) to 334 ($N_{\mathbf{x}}$). Figure 13 shows the evolution of the KL divergence $KL(\rho|\rho^*)$ and the residual norm $\|\mathbf{R}\|_2$ with respect to the number of selected DOFs N_k. The KL divergence and the global residual norm $\|\mathbf{R}\|_2$ are not monotonic with respect to N_k but they both follow the same decreasing trend. From $N_k = 50$, there is no significant change in the KL divergence. Both observations confirm the relevance of the CS procedure and of the *a priori* error analysis. Table A.2 summarizes the parameters estimated statistics (mean and standard deviation) with respect

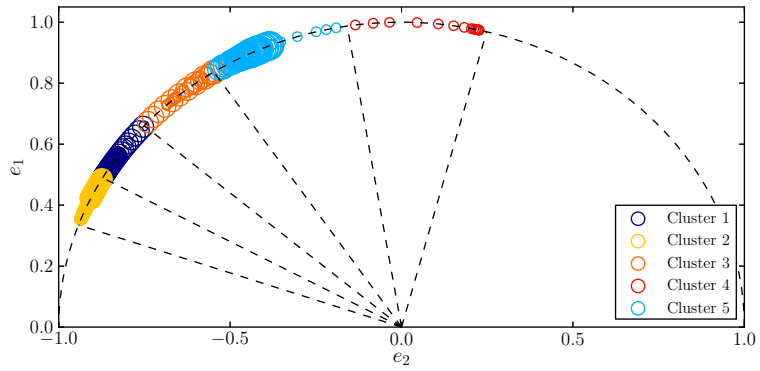

Figure 12: Scatter plot of the SGM first eigenvector for each DOF $\mathbf{x}_j$. $N_k = 5$.

to N_k. It is clear that a certain convergence is reached as N_k increases, both in the KL divergence and in the parameters statistics themselves.

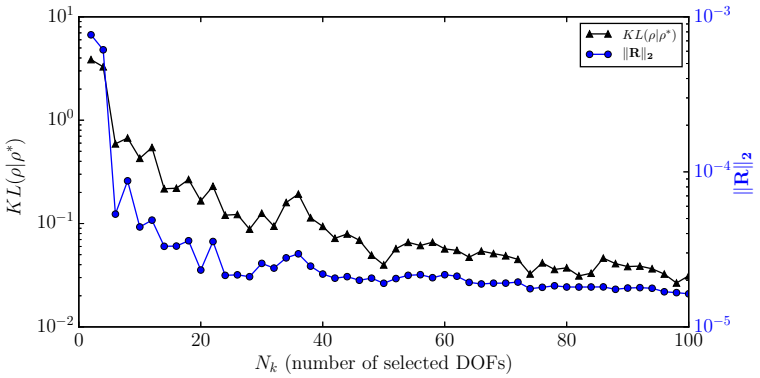

Figure 13: Convergence of the KL error and residual norm as the number of selected DOFs N_k increases.

Influence of N_c The effect of the number of stochastic collocation points N_c is investigated with $N_m = 3$ and $N_k = 50$. N_c varies from 2 to 2^{10} and $N_m = 3$ and $N_k = 50$ are fixed. Table A.2 shows the means and standard deviations of θ_1 and θ_2 estimated by the Observable Moment Matching algorithm as well as their empirical values. The empirical moments correspond to the moments computed directly from the synthetic parameter samples using the following formula:

$$\mu_{m,k} = \frac{1}{N}\sum_{i=1}^{N} \theta_{i,k}^{m}, \qquad m = 1, \ldots, N_m, \;\; k = 1, \ldots, n_p. \tag{36}$$

Table 4: Observable Moment Matching results for different values of N_k ($N_c = 512$ and $N_m = 3$).

Statistics	$KL(\rho\|\rho*)$	mean		std	
Number of DOFs		θ_1	θ_2	θ_1	θ_2
$N_k = 2$	3.84	1.2124	1.1458	0.343	0.231
$N_k = 5$	6.45×10^{-1}	1.1080	1.1071	0.150	0.127
$N_k = 10$	4.18×10^{-1}	1.1078	1.1118	0.140	0.135
$N_k = 20$	1.65×10^{-1}	1.1006	1.1068	0.119	0.121
$N_k = 50$	4.10×10^{-2}	1.0978	1.1037	0.108	0.103
$N_k = 100$	2.49×10^{-2}	1.0974	1.1037	0.106	0.103
$N_k = 200$	2.94×10^{-2}	1.0971	1.0383	0.105	0.103
$N_k = 334$	3.01×10^{-2}	1.0970	1.1039	0.105	0.103
Empirical	**0**	**1.0972**	**1.1042**	**0.104**	**0.102**

As expected, the estimation is more accurate when N_c increases. Note that the computational cost of increasing N_c is limited owing to the deterministic and nested nature of the Sobol sequence. If one already has evaluated the model for N_{c1} sample points and wants N_{c2} model evaluations, one only has to perform $N_{c2} - N_{c1}$ forward runs to complete the simulation set.

Influence of the noise level The synthetic measurements are corrupted by adding some noise to the numerical results. Table A.2 shows the estimated means and standard deviations of θ_1 and θ_2 for different noise levels. As expected, the accuracy of the method decreases as the noise increases.

Non-normal distributions In order to assess the robustness of the method, a similar but more complex heart cell model [38] is used. It consists of a set of 29 nonlinear coupled ODEs and we aim at estimating the PDF of two parameters of this model. The synthetic dataset is generated by sampling the parameters of interest from two known distributions: a bivariate log-normal distribution $\text{Log} - \mathcal{N}(0, \sigma_1^2)$ and a bivariate Gaussian mixture $\mathcal{N}(1, \sigma_2^2) + \mathcal{N}(2, \sigma_2^2)$ with $\sigma_1 = 0.7$ and $\sigma_2 = 0.2$. In both cases, the synthetic dataset is corrupted by a zero-mean Gaussian noise of amplitude 5%. The inverse procedure is applied to the log-normal case with the following numerical settings: $N_c = 2048$, $N_m = 3$ and convergence is reached at $n_{iter} = 1$ and $N_k = 21$. The PDF values are shown in Figure 14 and the marginal densities in Figure 15. Note that the strong skewness of the true distribution is well captured by the proposed inverse procedure.

The inverse procedure is then applied to the Gaussian mixture case with the following numerical settings: $N_c = 2048$, $N_m = 3$ and convergence is

Table 5: Observable Moment Matching results for different values of N_c ($N_k = 50$ and $N_m = 3$).

Statistics	$KL(\rho\|\rho*)$	mean		std	
Number of stochastic points		θ_1	θ_2	θ_1	θ_2
$N_c = 4$	1.62	1.1328	1.1378	0.187	0.189
$N_c = 8$	1.89	1.1000	1.1241	0.116	0.151
$N_c = 16$	8.48×10^{-1}	1.1006	1.1002	0.123	0.105
$N_c = 32$	2.69×10^{-1}	1.0995	1.1109	0.119	0.134
$N_c = 64$	1.02×10^{-1}	1.0968	1.1049	0.107	0.110
$N_c = 128$	5.04×10^{-2}	1.0965	1.1038	0.106	0.104
$N_c = 256$	4.99×10^{-2}	1.0979	1.1039	0.109	0.105
$N_c = 512$	4.10×10^{-2}	1.0978	1.1037	0.108	0.103
$N_c = 1024$	4.20×10^{-2}	1.0978	1.1037	0.108	0.104
Empirical	**0**	**1.0972**	**1.1042**	**0.104**	**0.102**

Table 6: Observable Moment Matching results for different noise levels ($N_c = 512$, $N_k = 50$ and $N_m = 3$).

Statistics	$KL(\rho\|\rho*)$	mean		std	
Noise level		θ_1	θ_2	θ_1	θ_2
80%	1.55	1.1027	1.1543	0.120	0.251
20%	1.23×10^{-1}	1.0967	1.1019	0.105	0.094
10%	6.97×10^{-1}	1.0994	1.1161	0.117	0.162
5%	4.10×10^{-2}	1.0978	1.1037	0.108	0.103
2%	3.92×10^{-2}	1.0978	1.1051	0.107	0.108
1%	3.79×10^{-2}	1.0977	1.1050	0.107	0.108
0%	3.70×10^{-2}	1.0977	1.1048	0.107	0.107
Empirical	**0**	**1.0972**	**1.1042**	**0.104**	**0.102**

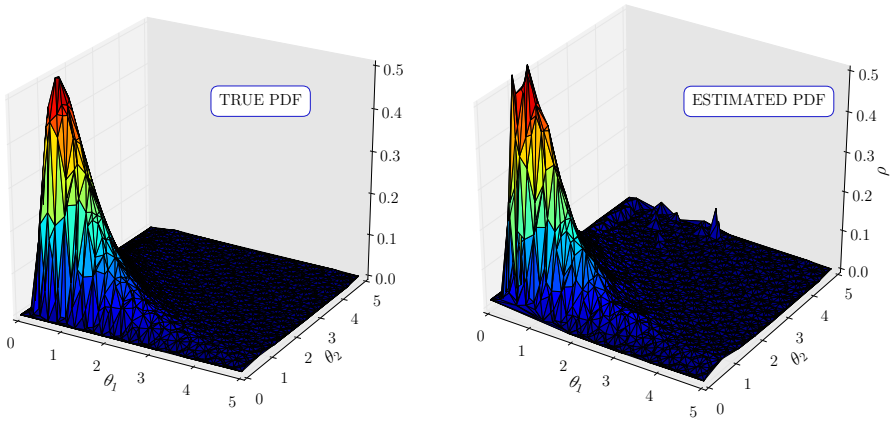

Figure 14: PDF estimation of a bivariate log-normal distribution: direct visualization.

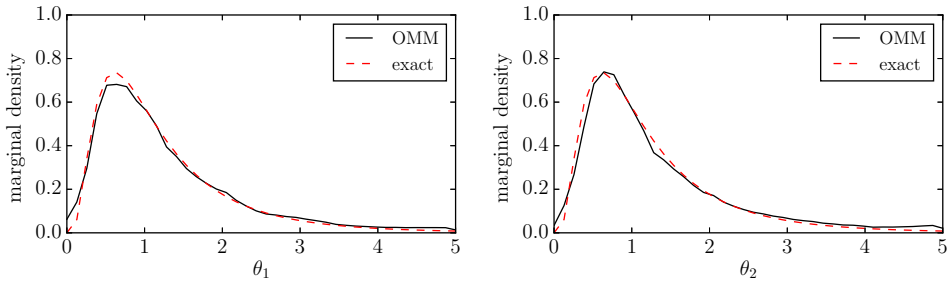

Figure 15: PDF estimation of a bivariate log-normal distribution: marginal densities.

reached at $n_{iter} = 1$ and $N_k = 31$. The PDF values are shown in Figure 16 and the marginal densities in Figure 17. Note that strong correlation between θ_1 and θ_2 is fully captured.

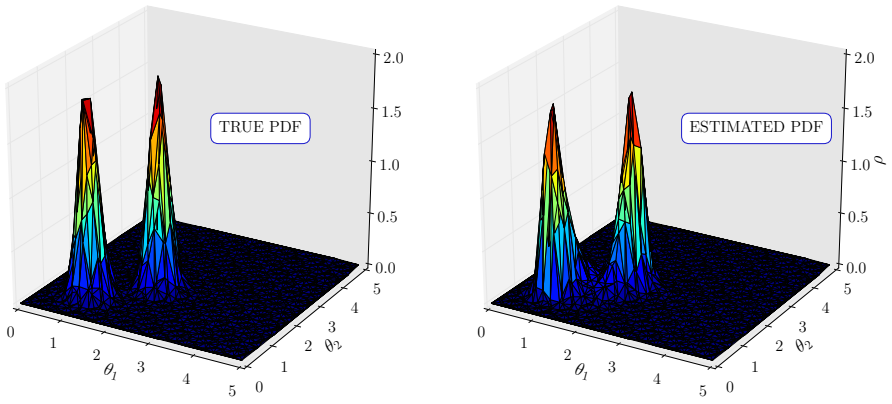

Figure 16: PDF estimation of a bivariate Gaussian mixture: direct visualization.

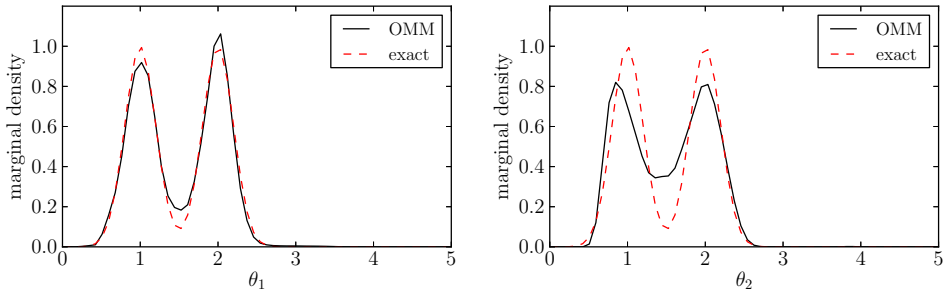

Figure 17: PDF estimation of a bivariate Gaussian mixture: marginal densities.

A.3 Comparison with existing techniques

In this section, the proposed approach is compared to existing techniques on the reference test case described in A.2. We show that all three approaches achieve the same precision on the parameters estimations but with a different number of function evaluations.

Least-squares moment-matching An alternative to the present approach is to directly minimize the moment difference using a least-squares method. In [14], the quantity of interest is represented as a finite polynomial expansion of d uncorrelated random variables $\xi = \{\xi^1, \ldots, \xi^d\}$. The

methodology was applied to inverse heat conduction problems and to microstructure reconstruction. This approach was applied to our test case. Here, the quantities of interest are the two parameters θ_1 and θ_2 and the observable is the variable u. Using the methodology presented in [14], the parameter θ_j is expanded on a sparse grid as follows:

$$\theta_j(\xi) = \sum_{k=1}^{n_k} \theta_j(\xi_k) L_k(\xi) = \sum_{k=1}^{n_k} q_{j,k} L_k(\xi),$$

where $q_{j,k} = \theta_j(\xi_k)$, the ξ_k are the sparse grid collocation points and L_k the members of the polynomial basis. Then, one can approximate the moments of the observable using the sparse grid quadrature rule:

$$\mu_m^*(\mathbf{x}_j) = \sum_{k=1}^{n_k} w_k g(\mathbf{x}_j, \xi_k)^m,$$

where the w_k are the sparse grid weights. The cost function J is defined as the squared difference between the approximated and experimental moments:

$$J = \frac{1}{2} \sum_{m=1}^{N_m} \sum_{j=1}^{N_\mathbf{x}} \alpha_m \left(\mu_m^*(\mathbf{x}_j) - \mu_m(\mathbf{x}_j)\right)^2,$$

where the α_m are user-defined weights. The problem now consists in minimizing J with respect to the coefficients $q_{j,k}$. In [14], this is done by a gradient descent method which involves solving the sensitivity equations associated with the model. For the sake of simplicity, to avoid the tedious derivation of the sensitivity equations of the MV model, we used the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) evolutionary algorithm [39] to minimize J. Since the minimization strategy differs from that of [14], the number of model evaluations needed to reach convergence may differ. This is to be taken into account when comparing the three methods in Table A.3.

Population approach (SAEM) Here we tackle the inverse problem from a radically different perspective, belonging to the so-called population approaches. It consists in seeking a Maximum Likelihood (ML) estimate of the unknown parameters. The MV test case can be seen as a mixed effects model where the observed data are the $y_{i,j}$, $i = 1, \ldots, N$, $j = 1, \ldots, N_\mathbf{x}$ and the parameters θ_1, θ_2 are the non-observed data. We assume that the observed data are outputs of the MV model with an additive noise $\epsilon_{i,j}$ assumed to be normally distributed: $\epsilon \sim \mathcal{N}\left(0, \tau^2\right)$.

$$y_{i,j} = g(\boldsymbol{\theta}_i, \mathbf{x_j}) + \epsilon_{i,j}.$$

Assuming each θ_j is normally distributed, $\theta_j \sim \mathcal{N}\left(\mu_j, \sigma_j^2\right)$, the likelihood L reads:

$$L(y,\theta;\tau,\mu_k,\sigma_k) = \left(2\pi\sigma_1^2\sigma_2^2\right)^{-N/2}\left(2\pi\tau^2\right)^{-NN_{\mathbf{x}}/2}\exp\left[-\frac{1}{2\tau^2}\sum_{i,j}\left(y_{i,j}-g(\boldsymbol{\theta}_i,\mathbf{x_j})\right)^2\right.$$
$$\left.-\frac{1}{2\sigma_1^2}\sum_i\left(\theta_{1,i}-\mu_1\right)^2-\frac{1}{2\sigma_2^2}\sum_i\left(\theta_{2,i}-\mu_2\right)^2\right].$$

Note that this approach differs from the other two on two major aspects. First, it is a *parametric* approach, meaning we are not seeking a pointwise estimate of the PDF but a parameterization of it (here a Gaussian parameterization). Second, the method provides, by construction, an estimation of the noise level of the measurements. In the other two approaches, the noise structure and amplitude is assumed to be known. The parameters τ, μ_k, σ_k are found by maximizing the log-likelihood $\log(L)$, which is challenging due to the nonlinear relationship between g and θ_1, θ_2. This is called the Maximum Likelihood Estimation (MLE) method. In the case of linear models, the maximum likelihood is usually found using the Expectation Maximization (EM) algorithm [40]. The paper by E. Kuhn and M. Lavielle [15] introduces a modified version of the EM algorithm to tackle cases where the models are nonlinear. The authors developed a Stochastic Approximation of the Expectation Maximization algorithm (SAEM) to solve the MLE problem. For the comparison study, we used Monolix® [41], the Matlab® implementation of the SAEM algorithm. This software was initially designed to perform the parameter estimation of pharmacokinetics-pharmacodynamics (PK-PD) models. Compared to PDEs, those models are usually computationally cheap so that the software does not look for a solution with minimum model evaluations. However, one may reduce the computational cost by constructing a pre-computed grid of solutions and then interpolate in that grid instead of evaluating the full model. The Monolix software was successfully used in [42] to estimate the parameters of a 1–D PDE model. Such a strategy was not adopted in this paper and the software was used as is.

Comparison We applied the Clustered Sensitivities / Observable Moment Matching algorithms and both the previously described methods to the reference test case described in A.2. The numerical settings for our method are: $N_c = 512$, $N_m = 3$ and $N_k = 50$. For the least-squares method, we used a two-dimensional sparse grid using the Smolyak rule [43] to discretize the parameter space with $N_c = 9$ and the first $N_m = 3$ moments were matched. As explained before, the SAEM algorithm was applied using the Monolix software with default settings.
Table A.3 shows the estimations of the parameters moments and the number of model evaluations needed for the three methods. For all three approaches,

the errors on the means are less than 1% and the errors on the standard deviations are less than 10%. Even though the SAEM appears to be more precise than the other two, the main difference lies in the number of model evaluations needed. Our approach requires much less model evaluations and those evaluations are made offline, once and for all. Again, our implementation of the least squares method presented in [14] may require more model evaluations due to the minimization strategy adopted.

Table 7: Comparison with existing techniques

	Exact		SAEM		least-squares		OMM	
moment order	θ_1	θ_2	θ_1	θ_2	θ_1	θ_2	θ_1	θ_2
1	**1.0972**	**1.1042**	1.0975	1.1051	1.0972	1.1019	1.0963	1.1015
2	**1.2147**	**1.2297**	-	-	1.2133	1.2215	1.2125	1.2224
3	**1.3566**	**1.3810**	-	-	1.3522	1.3616	1.3520	1.3663
std	**0.104**	**0.102**	0.104	0.102	0.098	0.086	0.103	0.095
model evaluations	-		2.98×10^6		1.67×10^5		512	