



# Bayesian calibration of order and diffusivity parameters in a fractional diffusion equation

Hasnaa H Alzahrani, Marco Lucchesi, Kassem Mustapha, Olivier Le Maitre,  
Omar M Knio

## ► To cite this version:

Hasnaa H Alzahrani, Marco Lucchesi, Kassem Mustapha, Olivier Le Maitre, Omar M Knio. Bayesian calibration of order and diffusivity parameters in a fractional diffusion equation. *Journal of Physics Communications*, 2021, 5 (8), pp.085014. 10.1088/2399-6528/ac1507 . hal-03384280

**HAL Id: hal-03384280**

**<https://hal.science/hal-03384280v1>**

Submitted on 18 Oct 2021

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

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

# Bayesian calibration of order and diffusivity parameters in a fractional diffusion equation

Hasnaa H. Alzahrani<sup>a</sup>, Marco Lucchesi<sup>a</sup>, Kassem Mustapha<sup>b</sup>, Olivier P. Le Maître<sup>c</sup>, Omar M. Knio<sup>a,\*</sup>

<sup>a</sup>*Computer, Electrical, and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia*

<sup>b</sup>*Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, 31261, Saudi Arabia*

<sup>c</sup>*Centre de Mathématiques Appliquées, CNRS, Inria, École Polytechnique, 91128 Palaiseau, France*

---

## Abstract

This work focuses on parameter calibration of a variable-diffusivity fractional diffusion model. A random, spatially-varying diffusivity field with log-normal distribution is considered. The variance and correlation length of the diffusivity field are considered uncertain parameters, and the order of the fractional sub-diffusion operator is also taken uncertain and uniformly distributed in the range  $(0, 1)$ . A Karhunen-Loève (KL) decomposition of the random diffusivity field is used, leading to a stochastic problem defined in terms of a finite number of canonical random variables. Polynomial chaos (PC) techniques are used to express the dependence of the stochastic solution on these random variables. A non-intrusive methodology is used, and a deterministic finite-difference solver of the fractional diffusion model is utilized for this purpose. The PC surrogates are first used to assess the sensitivity of quantities of interest (QoIs) to uncertain inputs and to examine their statistics. In particular, the analysis indicates that the fractional order has a dominant effect on the variance of the QoIs considered, followed by the leading KL modes. The PC surrogates are further exploited to calibrate the uncertain parameters using a Bayesian methodology. Different setups are considered, including distributed and localized forcing functions and data consisting of either noisy observations of the solution or its first moments. In the broad range of parameters addressed, the analysis shows that the uncertain parameters having a significant impact on the variance of the solution can be reliably inferred, even from limited observations.

**Keywords:** Fractional diffusion; Bayesian Calibration; Karhunen-Loève expansion; Polynomial Chaos

---

## 1. Introduction

Anomalous diffusion occurs in a variety of natural phenomena, as well as applications involving particulate motion [1, 2, 3], viscoelastic materials [4], or subsurface transport [5]. Fractional diffusion models have shown to be particularly effective in representing and capturing such physical phenomena (see e.g. [6, 7, 8, 9]). They have also been successfully applied in various disciplines such as control, image processing, finance, and biological systems [10, 6, 11, 12, 13, 14, 15].

Except in simple or idealized settings, fractional diffusion models are not amenable to analytical solutions. Consequently, various efforts have focused on their numerical solution [16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36]. Specifically, in the case of the constant-diffusivity space fractional diffusion equation (FDE), various numerical solution schemes were developed, based on Galerkin [37, 38, 19, 39], spectral [22], discrete particle [40, 41, 42, 43, 44, 45], smooth particle [46, 47], and finite-difference methodologies [24, 25, 48]. In complex settings, models may be used that involve FDEs with non-uniform diffusivity fields. Various approaches have also been developed to solve these equations, including spectral representations based on Jacobi polynomials [26, 49, 50], as well as finite-difference techniques [51, 52, 17].

---

\*Corresponding author. Email: omar.knio@kaust.edu.sa

Despite recent advances, the application of fractional diffusion models remains challenging. One challenge concerns the numerical solution of the governing equation, which typically involves full matrices that arise from the discretization of nonlocal operators. Unless suitably addressed with fast, reduced storage techniques [47, 53, 54, 55], large memory and CPU requirements would be needed. Another challenge concerns the calibration of the parameters of fractional diffusion operators, including the fractional exponent and diffusivity field. These parameters are generally unknown *a priori*, and must be consequently inferred from measurements or observations. This work focuses on this challenge.

Bayesian inverse methods [56] are powerful statistical tools that enable calibration of uncertain parameters and random fields. These methods offer various advantages, including robust and systematic means for dealing with noisy measurements, quantifying information gained through specific measurements, and comparing solutions. In contrast to problems involving classical gradient diffusion [57, 58, 59, 60, 61, 62], the application of Bayesian methods to fractional diffusion appears to be quite limited. Most of these applications focused on time-fractional evolution equations, see [63, 64, 65, 66, 67]. To our knowledge, Bayesian methods have not been applied to space FDEs. Specifically, calibration of space FDEs has so far relied on identification approaches [68, 69].

This paper explores the application of a Bayesian inference approach to calibrate the fractional order parameter and infer a non-uniform diffusivity field in a space FDE. Similar to the case classical diffusion problems, this is motivated by the desire to capitalize on the advantages afforded by the Bayesian formalism, which generally yields a full posterior distribution of the parameters and their correlation, information concerning the value of particular observations, as well as a robust approach to experimental design. This work focuses on a forced, steady, one-dimensional (1D) space FDE with uncertain parameters. The restriction to the 1D case enables us to explore different scenarios involving different combinations of uncertain inputs, forcing functions, and types of observations, and consequently assess the performance of the Bayesian inversion under a broad range of conditions. Specifically, we will generally consider the situation of FDEs with an uncertain order and a random diffusivity field. The latter is a shifted log-normal field, with the underlying correlated Gaussian field having either known or uncertain parameters. Briefly, our approach relies on the application of (i) the truncated Karhunen–Loève (KL) decomposition [70, 71] of the Gaussian field to construct a compact representation of the random diffusivity field; (ii) the finite-difference scheme [17] to obtain deterministic solutions for individual realizations of the uncertain inputs; (iii) polynomial chaos expansions (PCEs) [70, 71] to establish suitable functional representations of the solution or of the quantities of interest (QoIs). The PCEs are first exploited to analyze the sensitivity of the QoIs to the uncertain inputs [72, 73, 74], and then to accelerate the solution of the Bayesian calibration problem [75, 76, 77]. To assess the calibration quality, we rely on synthetically generated data and explore the impact of different ways of probing and observing the system. Within this framework, we explore a range of inverse problems based on situations in which only a limited amount of system observations is available, and other situations in which the inference is based on available solution statistics.

This paper is organized as follows. In section 2, we introduce the main notations and outline the governing equation as well as the deterministic solver used for its numerical solution. A stochastic formulation of the variable-diffusivity FDE is then introduced and discussed in section 3. Basic concepts of PCE and non-intrusive approaches used for their construction are outlined in section 4. Computational test cases are summarized in section 5. Results of the analysis are discussed in section 6, and concluding remarks are given in section 7.

## 2. Governing equation and deterministic solver

We focus on the steady, 1D, two-sided, fractional order diffusion equation with variable diffusivity:

$$-\partial_x (\kappa(x) \partial_x^\alpha u(x)) = f(x), \quad x \in [a, b], \quad (1)$$

subject to homogeneous Dirichlet boundary conditions  $u(x) = 0$  for  $x \leq a$  and  $x \geq b$ . In (1),  $\kappa$  denotes the diffusivity,  $f$  is the forcing term, whereas  $\partial_x$  denotes the first-order derivative, and  $\partial_x^\alpha$  is the two-sided

fractional differential operator of order  $\alpha \in (0, 1)$ ,

$$\partial_x^\alpha v(x) \equiv \frac{\partial}{\partial x} \int_a^b |\omega_{1-\alpha}(x-z)| v(z) dz,$$

with

$$\omega_{1-\alpha}(x) := \frac{x^{-\alpha}}{\Gamma(1-\alpha)}.$$

The variable diffusivity field,  $\kappa$ , is assumed to satisfy  $c_0 \leq \kappa(x) \leq c_1$  for some positive constants  $c_0$  and  $c_1$  and all  $x \in [a, b]$ .

We seek the numerical solution of problem (1) by adopting the finite difference method proposed in [17], which combines first-order forward and backward approximations of the Riemann-Liouville derivatives. Therein, the existence and uniqueness of the finite difference solution were established, and a rigorous truncation error analysis was provided. We consider a uniform partition of the domain using a sequence of  $(P+1)$  grid points with step size  $h = \frac{b-a}{P}$  and nodes  $x_n$ . We let  $x_n = a + nh$  for  $n = 0, \dots, P$ ,  $x_{n+1/2} = (x_n + x_{n+1})/2$ , and for a given function  $v$ ,  $v^n := v(x_n)$ . Let  $\mathbf{U} = (U^1, \dots, U^{P-1})^T$  denote the finite difference solution vector, and  $\mathbf{F} = (\tilde{f}_h^1, \dots, \tilde{f}_h^{P-1})^T$  the modified forcing term defined according to:

$$\tilde{f}_h^n := \frac{h^2}{\omega_{2-\alpha}(h)} f^n.$$

Following the details in [17],  $\mathbf{U}$  can be obtained by solving the system  $\mathbf{BU} = \mathbf{F}$ , where the full dense matrix

$$\mathbf{B} = \begin{pmatrix} l_{1,1} & l_{1,2} & d_{1,3} & d_{1,4} & d_{1,5} & \dots & d_{1,P-1} \\ l_{2,1} & l_{2,2} & l_{2,3} & d_{2,4} & d_{2,5} & \dots & d_{2,P-1} \\ c_{3,1} & l_{3,2} & l_{3,3} & l_{3,4} & d_{3,5} & \dots & d_{3,P-1} \\ c_{4,1} & c_{4,2} & l_{4,3} & l_{4,4} & l_{4,5} & \dots & d_{4,P-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ c_{P-1,1} & c_{P-1,2} & c_{P-1,3} & c_{P-1,4} & \dots & l_{P-1,P-2} & l_{P-1,P-1} \end{pmatrix}$$

with

$$\begin{aligned} l_{i,i} &= (\kappa^{i-1/2} + \kappa^{i+1/2})[3 - 2^{1-\alpha}], \\ l_{i+1,i} &= \kappa^{i+1/2}[2^{1-\alpha} - 3] - \kappa^{i+3/2}[3^{1-\alpha} - 2^{2-\alpha} + 1], \\ l_{i,i+1} &= \kappa^{i+1/2}[2^{1-\alpha} - 3] - \kappa^{i-1/2}[3^{1-\alpha} - 2^{2-\alpha} + 1], \\ c_{n,j} &= \begin{cases} \kappa^{n-1/2} + \kappa^{n+1/2}[2 - 2^{1-\alpha}] & j = n, \\ \kappa^{n-1/2}[w_{n,j} - w_{n-1,j}] - \kappa^{n+1/2}[w_{n+1,j} - w_{n,j}], & j < n, \end{cases} \\ d_{n,j} &= \begin{cases} -\kappa^{n-1/2}w_{j,n-1} + (\kappa^{n-1/2} + \kappa^{n+1/2})w_{j,n} - \kappa^{n+1/2}w_{j,n+1} & j > n, \\ \kappa^{n+1/2} + \kappa^{n-1/2}[2 - 2^{1-\alpha}], & j = n, \end{cases} \\ \kappa^{i+1/2} &:= \kappa(x_{i+1/2}), \\ w_{n,j} &:= (n+1-j)^{1-\alpha} - (n-j)^{1-\alpha} \quad n \geq j \geq 1. \end{aligned}$$

### 3. Stochastic Formulation

A stochastic framework is introduced by considering uncertainty in the order of the fractional exponent  $\alpha$ , and in the diffusivity  $\kappa$ . The former is specifically modeled as a random variable, whereas the latter is modeled as a log-normal random field of the form:

$$\kappa(x, \theta) = \mu_\kappa(x) \exp(M(x, \theta)), \quad (2)$$

where  $\mu_\kappa(x)$  is the median of the log-normal field and  $M(x, \theta)$  is a real-valued, centered, Gaussian process characterized by the square exponential covariance function:

$$C(x, x') := \sigma_M^2 \exp\left(-\frac{|x - x'|^2}{2L_M^2}\right). \quad (3)$$

In the covariance expression,  $L_M$  is the correlation length and  $\sigma_M$  is the standard deviation of the Gaussian field. In the following  $L_M$  and  $\sigma_M$  are either given, or alternatively treated as uniformly-distributed random variables. Without loss of generality, we set throughout the paper  $\mu_\kappa(x) = 1$ . The covariance function,  $C$  is continuous, symmetric and positive-definite. By Mercer's theorem [78], it can be decomposed spectrally as:

$$C(x, x') = \sum_{k=1}^{\infty} \lambda_k \phi_k(x) \phi_k(x'),$$

where  $\lambda_k$  are eigenvalues, and  $\phi_k$  are the associated  $L_2$ -orthonormal eigenfunctions of the Fredholm integral operator with kernel,  $C$ . Thus, the truncated KL expansion of  $\kappa$  is given by [70]:

$$\kappa(x, \theta) \approx \mu_\kappa(x) \exp\left(\sum_{k=1}^{N_{KL}} \sqrt{\lambda_k} \phi_k(x) \xi_k(\theta)\right), \quad (4)$$

where  $N_{KL}$  is the number of the terms retained in the KL expansion, and the  $\xi_k(\theta)$  are uncorrelated Gaussian random variables with zero mean and unit variance.

#### 4. Polynomial Chaos Surrogates

Surrogate models of the solution, or of QoIs, are used to conduct sensitivity analyses of the solution and to accelerate the Bayesian calibration of  $\kappa$  and/or  $\alpha$  [79]. We rely on a PC formalism for this purpose [71]. This section briefly outlines the PC methodologies.

Consider an uncertain model inputs parametrized by an  $N$ -dimensional vector of canonical random variables,  $\boldsymbol{\xi}$ , with independent components  $\xi_1, \dots, \xi_N$ . The random vector  $\boldsymbol{\xi}$  is sometimes called the stochastic germ. The PC approximation  $\widehat{U}(\boldsymbol{\xi})$  of a generic second-order functional  $U(\boldsymbol{\xi})$  of the model output is expressed as:

$$U(\boldsymbol{\xi}) \approx \widehat{U}(\boldsymbol{\xi}) = \sum_{\beta=0}^{N_t} U_\beta \Psi_\beta(\boldsymbol{\xi}), \quad (5)$$

where  $N_t$  is the number of terms retained in the PC expansion. Note that Roman subscripts are used to denote discrete spatial locations, whereas Latin subscripts denote PC modes index. Here,  $\{\Psi_\beta, \beta = 0, 1, \dots\}$  forms a complete orthogonal polynomial basis of the space of second-order random functionals in  $\boldsymbol{\xi}$ , denoted  $\mathcal{A} := \{U : \int_{\Omega} U^2(\boldsymbol{\xi}) p(\boldsymbol{\xi}) d\boldsymbol{\xi} < \infty\}$ , where  $p(\boldsymbol{\xi})$  is the joint density function of the  $\xi_i$ 's. Because the  $\xi_i$ 's are independent, the density  $p(\boldsymbol{\xi}) = \prod_{i=1}^N p_{\xi_i}(\xi_i)$ , where  $p_{\xi_i}$  is the density of  $i$ -th random variable,  $\xi_i$ .

We will rely on non-intrusive approaches [71] to determine the PC coefficients,  $U_\beta$ . These methods generally aim to minimize the  $L_2$  error  $\|U - \widehat{U}\|$ , based on a discrete set of deterministic model realizations. In this work, both regression and spectral projection approaches are used to perform the minimization.

The regression approach [80] calculates the coefficients  $U_{0 \leq \beta \leq N_t}$  that minimize the least-squares error between the solution  $U(\boldsymbol{\xi})$  and its approximation  $\widehat{U}(\boldsymbol{\xi})$ . Consider a sample set of  $N_s$  realizations,  $\mathbf{S} = (\boldsymbol{\xi}^{(1)}, \dots, \boldsymbol{\xi}^{(N_s)})^T$ , and denote  $\mathbf{V} = (V(\boldsymbol{\xi}^{(1)}), \dots, V(\boldsymbol{\xi}^{(N_s)}))^T$  the corresponding sample set of observations of the QoI  $V(\boldsymbol{\xi})$ . The  $L_2$  minimization problem can be written in matrix form as follows:

$$\mathbf{Z}^T \mathbf{Z} \mathbf{U} = \mathbf{Z}^T \mathbf{V}, \quad (6)$$

where the matrix  $\mathbf{Z} = [Z_{i,\beta}]$  is defined such that  $Z_{i,\beta} = \Psi_\beta(\boldsymbol{\xi}^{(i)})$  and  $\mathbf{U} = (U_0, \dots, U_{N_t})^T$  is the vector of PC coefficients. In this work, we will also consider the regularized  $L_2$  minimization problem, that is solved using LASSO [81]. In this case, the minimization problem can be expressed as:

$$\mathbf{U} = \arg \min_{\mathbf{W}} \|\mathbf{Z}\mathbf{W} - \mathbf{V}\|^2 + \gamma \|\mathbf{W}\|_{l_1}, \quad (7)$$

where  $\|\mathbf{W}\|_{l_1} := \sum_{\beta=0}^{N_t} |W_\beta|$  and  $\gamma \geq 0$  is a parameter that is also optimized by a cross validation procedure.

The non-intrusive spectral projection (NISP) approach [82] aims to compute the PC coefficients of the model output through an orthogonal projection onto the PC basis. Hence, the evaluation of the coefficients proceeds as follows:

$$U_\beta = \frac{\langle U, \Psi_\beta \rangle}{\langle \Psi_\beta, \Psi_\beta \rangle}, \quad \beta = 0, \dots, N_t, \quad (8)$$

where

$$\langle U, \Psi_\beta \rangle = \int \Psi_\beta(\boldsymbol{\xi}) U(\boldsymbol{\xi}) p(\boldsymbol{\xi}) d\boldsymbol{\xi}.$$

Estimating the expansion coefficients amounts to the computation of  $N$ -dimensional integrals in a product space. Classically, the integrals are estimated by employing a numerical quadrature or a sampling approach. Since the  $\Psi_\beta$  are polynomials, we will apply fully-tensorized Gauss quadrature rules when the dimensionality  $N$  is small. Otherwise, we rely on an isotropic sparse-grid method (Smolyak formula [83]) based on the nested quadrature rules [71].

## 5. Setup

This section outlines the scope of the forward uncertainty quantification, sensitivity analysis and inference studies. As outlined in Table 1, four cases are considered in section 6. Case 1 concerns only one uncertain parameter, namely the fractional order  $\alpha$ . In Case 2, we use fixed  $L_M$ ,  $\sigma_M$ , and  $\alpha$ , and first study the solution sensitivity on the diffusivity field's KL coordinates. We then tackle the problem of calibrating  $\kappa$ . Case 3 also considers the impact of the leading diffusivity field's KL coordinates but also includes variability in  $\alpha$ . Case 4 explores the sensitivity of the first two solution's moments, namely the mean,  $\mu(x)$ , and standard deviation,  $\sigma(x)$ , to the correlation length,  $L_M$ , and standard deviation,  $\sigma_M$ , defining the log-normal diffusivity field, and the fractional order,  $\alpha$ . When considered uncertain, the parameters are assumed to be independent and follow uniform priors:

$$L_M \sim \mathcal{U}([0.5, 1]), \quad \sigma_M \sim \mathcal{U}([0, 1]), \quad \alpha \sim \mathcal{U}([0.1, 0.9]). \quad (9)$$

Each case also considers two types of forcing functions, distributed and localized. In the first type, the forcing function is  $f(x) = \sin(\pi x)$ . In the second type, the source has a sharp peak at  $x_0$  in the interior of the domain and follows  $f(x) = \frac{1}{\sqrt{\pi}\nu} \exp(-\frac{(x-x_0)^2}{\nu^2})$ . Different values of  $x_0$  are considered in the analysis with a fixed source width,  $\nu = 1/40$ .

As outlined in Table 1, different PC methodologies are used depending on the case considered. In Case 1, we use the NISP approach to approximate the dependence of the solution  $u$  on  $\alpha$ . In Case 2, we rely on a LASSO regression approach to build the PC surrogate of  $u$  as a function of the KL coordinates. As for Case 3, the surrogate of  $u$  is built via a fully-tensorized NISP machinery. In Case 4, we use a sparse-grid, pseudo-spectral projection (PSP) technique to construct functional approximations of the solution moments, in terms of  $\alpha$ ,  $L_M$  and  $\sigma_M$ .

In all cases, we consider a discretization of the physical domain,  $[0, 1]$ , with  $P = 300$  nodes, and we apply the finite-difference scheme introduced in Section 2. This discretization follows a careful grid refinement study (not shown), which revealed that this resolution is sufficient to obtain accurate predictions for the entire range of physical parameters considered. Also, unless otherwise stated, the KL is truncated by retaining the first five dominant modes (i.e.  $N_{KL} = 5$ ), which enables us to capture 99.9% of the underlying Gaussian field variability at the lowest value of the correlation length considered,  $L_M = 0.5$ .

Table 1: Summary of the cases considered in this work.

Case	QoI	Known quantities	Uncertain inputs	Surrogate via
Case 1	$u(x)$	$\kappa$	$\alpha$	NISP
Case 2	$u(x)$	$\alpha, L_M, \sigma_M$	$\kappa$	regression LASSO
Case 3	$u(x)$	$L_M, \sigma_M$	$\alpha, \kappa$	NISP
Case 4	$\mu(x)$ and $\sigma(x)$	–	$L_M, \sigma_M, \alpha$	sparse grid PSP

## 6. Numerical Experiments

This section presents for each case the surrogate construction, the results of the sensitivity analysis, and, finally, the inference results. For Cases 2–4, we perform a global sensitivity analysis [84] by calculating the Sobol indices [85]. These indices provide a variance-based importance measure of the effects of individual parameters on the variability of the solution or the QoI. After that, we focus on the Bayesian calibration problem based on synthetic data. The quality of inferred parameters is assessed based on a limited amount of noisy observations for each case scenario. Specifically, particular values of the parameters, called hereafter the true values, are picked within their respective prior ranges, and noisy observations are generated from the corresponding surrogates using a multiplicative Gaussian noise. The Maximum Likelihood Estimation (MLE) and Maximum A Posteriori (MAP) values of the parameters are retrieved, and the posterior densities are explored using an adaptive Markov Chain Monte Carlo (MCMC) algorithm [86]. Sufficiently long MCMC chains (not shown) ensure that stationary densities are adequately sampled. Kernel density estimation (KDE) [87, 88] was subsequently applied to the MCMC samples to estimate the posterior densities.

### 6.1. CASE 1

We build a surrogate for the solution  $u$  considering one uncertain parameter,  $\alpha$ . For this purpose, we used fixed values  $L_M = 0.75$ ,  $\sigma_M = 0.5$ ,  $\xi_1 = 0.9631$ ,  $\xi_2 = 0.5202$ ,  $\xi_3 = \xi_4 = \xi_5 = 0$  and  $\mu_\kappa = 1$ . Figure 1 shows the corresponding diffusivity field.

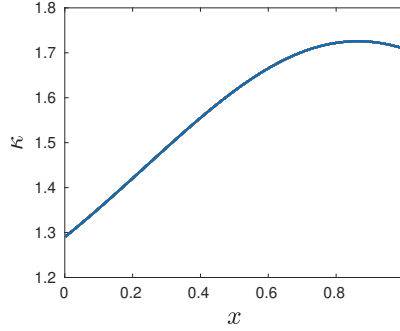


Figure 1: Diffusivity field,  $\kappa$ , for Case 1.

#### 6.1.1. Surrogate error

For the PC construction, we utilize a Gauss-Legendre quadrature rule with  $N_t = 8$  nodes, enabling us to determine a seventh-order expansion. In Figure 2 we compare the solution  $u$  and its surrogate  $\hat{u}$  for 16 randomly sampled values of  $\alpha$  in the prior range. The figure demonstrates the surrogate accuracy for both distributed and localized forcings. For a more quantitative accuracy assessment, the following relative error estimate,

$$\epsilon(\alpha) = \left( \frac{\int_0^1 (u(x; \alpha) - \hat{u}(x; \alpha))^2 dx}{\int_0^1 u(x; \alpha)^2 dx} \right)^{\frac{1}{2}}, \quad (10)$$

was computed and found to be less than 2% (results not shown) for all forcing functions considered and all values of  $\alpha$  in the prior range.

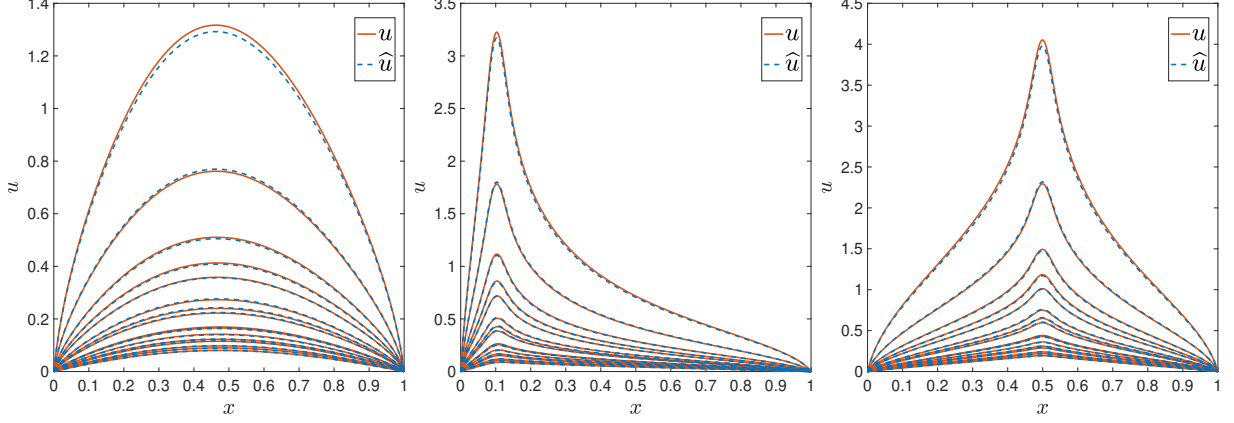


Figure 2: Case 1: solutions,  $u$ , and surrogates,  $\hat{u}$ , for 16 randomly selected values of  $\alpha$  corresponding to distributed forcing (left), localized forcing at  $x_0 = 0.1$  (center) and at  $x_0 = 0.5$  (right).

### 6.1.2. Inference of $\alpha$

In Case 1 we want to estimate  $\alpha$  based on observations of the solution at a collection of observation points  $x_i$ ,  $i = 1, \dots, m_d$ . A randomly selected true value  $\alpha_t = 0.4521$  is used. Noisy data,  $\tilde{u}_i$ , are generated at the observation points according to

$$\tilde{u}_i = \hat{u}(x_i; \alpha_t)(1 + \epsilon_i), \quad i = 1, \dots, m_d,$$

where  $\hat{u}(x_i; \alpha_t)$  is the surrogate prediction and  $\epsilon_i$ 's are i.i.d. Gaussian random variables with zero mean and standard deviation  $\tilde{\rho}_t = 0.01$ . The log-likelihood of the observations is consequently expressed as

$$\ln(\mathcal{L}(\mathbf{d}_u | \alpha, \rho)) \propto -\frac{1}{2} \sum_{i=1}^{m_d} \ln \rho_i^2 - \frac{1}{2} \sum_{i=1}^{m_d} \left( \frac{\tilde{u}_i - \hat{u}(x_i; \alpha)}{\rho_i} \right)^2, \quad (11)$$

where  $\mathbf{d}_u = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_{m_d})^T$  is the observation vector and  $\rho$  is a hyper-parameter that we also infer from the data. A Jeffrey's prior for  $\rho$  is assumed, with density  $p(\rho) = 1/\rho^2$ . Noisy observations are generated at  $m_d = 5$  observation points, located at  $x = 0.1, 0.3, 0.5, 0.7$  and  $0.9$ . For the distributed forcing case, the noisy observations are generated from a single experiment, observing the solution at the specified stations. In the localized forcing case, the observations come from  $m_d$  independent experiments involving a localized forcing centered at the corresponding observation station and using the same "system" (medium). The "true" solution profiles ( $\hat{u}$ ) and the noisy observations are depicted in Figure 3.

The marginal posterior densities of  $\alpha$  for the distributed and localized forcing scenarios are plotted in Figure 4. In both instances, the posterior of  $\alpha$  is a narrow distribution with a well-defined peak. The MAP estimates are  $\alpha = 0.4177$  for the distributed forcing experiment, and  $\alpha = 0.4788$  for the localized forcing experiments; the corresponding MLE results are  $\alpha = 0.4499$  and  $\alpha = 0.4787$ . The solution profiles corresponding to surrogates evaluated at the MAP and MLE values of  $\alpha$  are contrasted to the true profiles in Figure 3. A very close agreement between the true and inferred profiles is observed in the distributed case, whereas we notice some discrepancies for the localized forcing experiments. These higher discrepancies are consistent with the inferred MAP and MLE of  $\alpha$ , which are closer to the truth in the distributed forcing instance. Additional insight into the behavior of the two instances can be gained from Figures 5 and 6 which depict the solution,  $u$ , together with the 95% bounds of the pushed-forward prior and the 50% bounds of the posterior of  $\alpha$ , considering distributed and localized forcing, respectively. The plots indicate that the response of the stochastic solution and the inferred posterior can depend significantly on the imposed



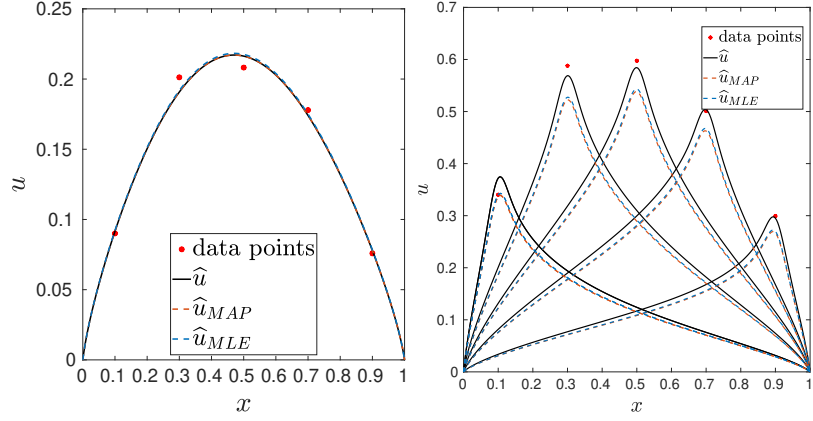


Figure 3: Case 1: profiles of MAP, MLE and true solutions, for distributed forcing (left) and localized forcing (right). The noisy observations used for the inference are also shown.

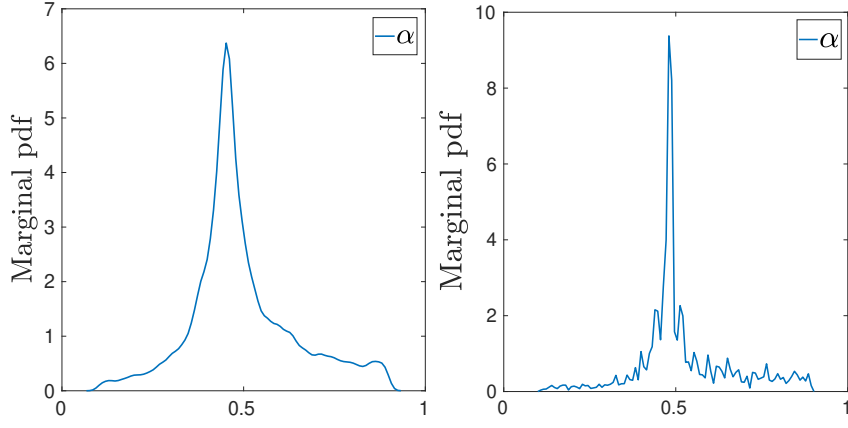


Figure 4: Case 1: marginal pdf of  $\alpha$  for distributed  $f(x)$  (left) and localized  $f(x)$  (right).

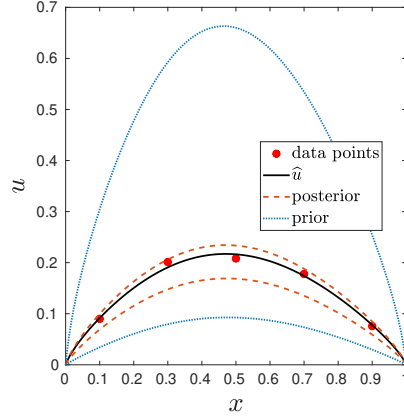


Figure 5: Case 1: profiles of the solution  $\hat{u}$  for distributed  $f(x)$ , the 95% bounds of the solution prior (blue dash), and the 50% bounds of the solution posterior (red dash). Also shown are the noisy measurements used to perform the inference.

forcing and provide insight into the computed posterior distributions. Specifically, one notes that though the posterior bounds in the distributed forcing case are tight, they clearly reflect the bias between the MAP

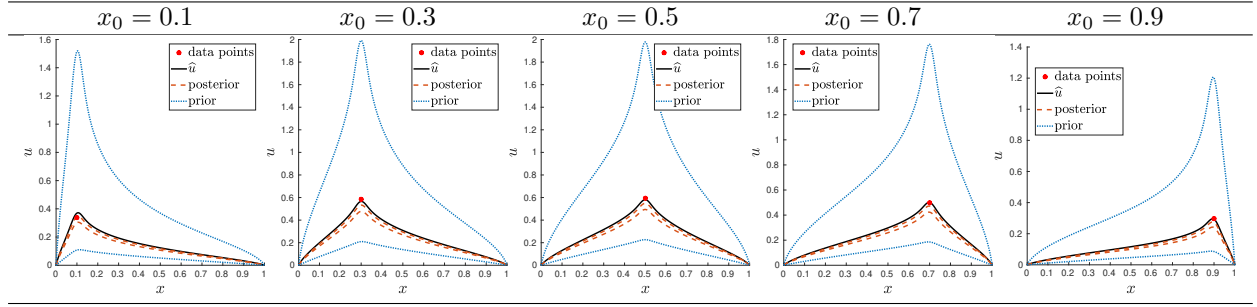


Figure 6: Case 1: profiles of the solution  $\hat{u}$  for localized forcing setups, as indicated. Also plotted are curves showing the 95% bounds of the solution prior (blue dash) and the 50% bounds of the solution posterior (red dash). The noisy measurements used to perform the inference are also depicted. Note that a single inference step is performed using all five data points simultaneously.

and true parameter (see Figure 4). On the other hand, the 50% bounds in the distributed forcing case are slightly broader but clearly capture the true solution. Of course, a more detailed analysis based on repeated realizations would be needed to provide a full characterization of the expected information gain associated with different strategies and its dependence on the number of observations. Such analysis is outside the scope of the present exploratory study and is left for future work.

## 6.2. CASE 2

Setting  $L_M = 0.75$ ,  $\sigma_M = 0.5$ , and  $\alpha = 0.5$ , we build the PC surrogate of  $u$  in terms of  $\xi$ . The PC coefficients are computed with the LASSO method, on a latin hypercube sample (LHS) set of size  $N_s = 50000$ . Specifically, we obtain a sample set of points  $\xi^i = (\xi_1^{(i)}, \dots, \xi_5^{(i)})^T$ , with  $i = 1, \dots, N_s$ ; then, for each sample point  $\xi^i$  we evaluate the corresponding solution  $u_i$  using the deterministic model, to constitute the training set of solutions used in the LASSO method. Moreover, surrogates using NISP were also established and produce surrogates (not shown) of similar quality.

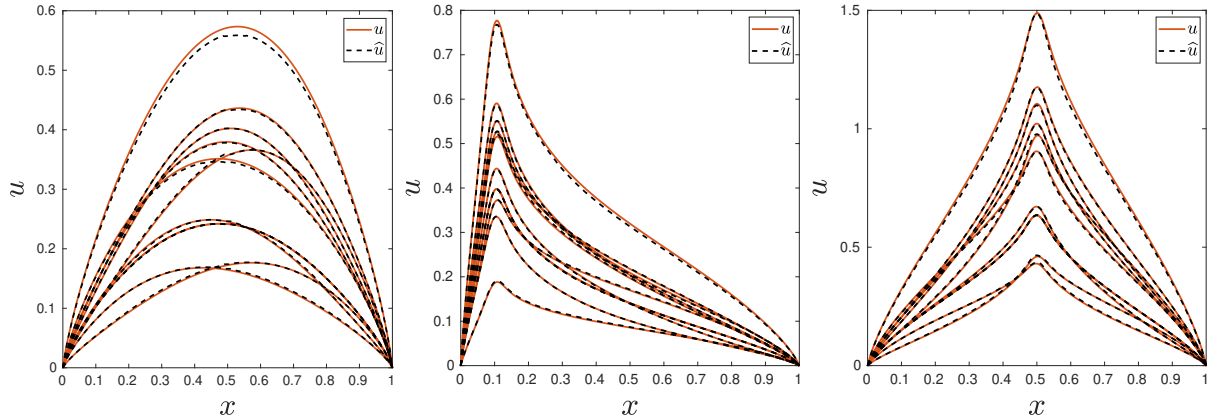


Figure 7: Case 2: solutions,  $u$ , and surrogates,  $\hat{u}$ , considering distributed  $f$  (left), and localized  $f$  with  $x_0 = 0.1$  (center) and  $x_0 = 0.5$  (right). Simulations are performed using 10 randomly sampled values of  $\xi$ . Surrogates are built using LASSO.

Figure 7 shows a comparison between the deterministic solution  $u$  and its surrogate for different forcing at randomly selected values of  $\xi$ . A good agreement is observed between the model and its surrogate. To provide a quantitative estimate of surrogate errors, we plot in Figure 8, profiles of local  $L_2$  errors between the surrogate and deterministic realizations generated using an independent coarse sample. Shown are results obtained for distributed forcing and forcing functions localized at  $x_0 = 0.1$  and  $0.5$ . As expected,

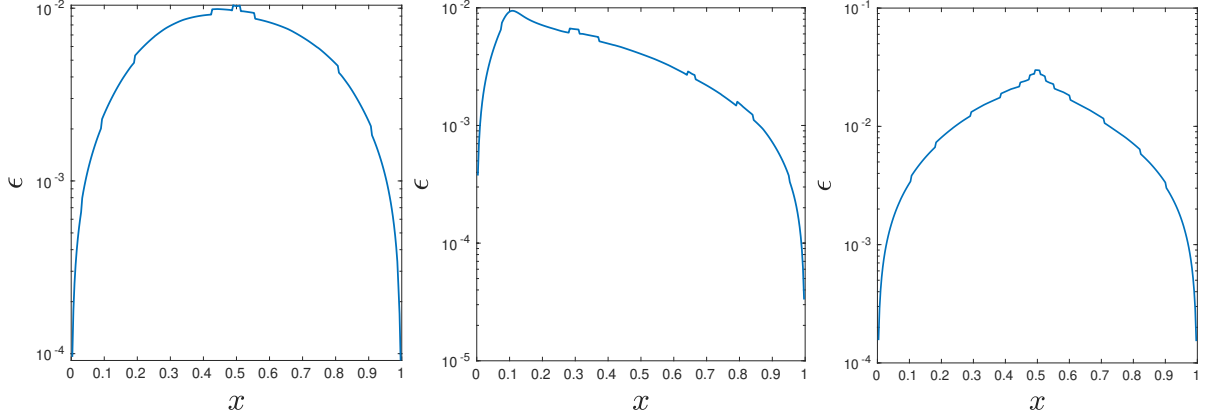


Figure 8: Case 2: local  $L_2$  error between surrogate and model realizations considering distributed (left) and localized forcing with  $x_0 = 0.1$  (center) and  $x_0 = 0.5$  (right). Errors are generated for an independent sample of size  $N_s = 100$ .

the localized profiles peak at the same locations as those of the forcing function. However, in all cases, the errors are small and fall below 2% of the peak values.

#### 6.2.1. Global Sensitivity Analysis

We first explore the variability of  $u$  due to the KL coordinates  $\xi_i$  of the diffusivity field, for the case of distributed forcing. Figure 9 displays the first-order and total-order partial variances of  $u$  due to  $\xi_i$ . The plots show clearly that the first KL coordinate,  $\xi_1$ , is responsible for most of the variance. The second coordinate,  $\xi_2$ , exhibits a small contribution to the variability of  $u$ , between the boundaries and midpoint of the domain. One also concludes from Figure 9 that other coordinates have an insignificant effect on the variability of the solution.

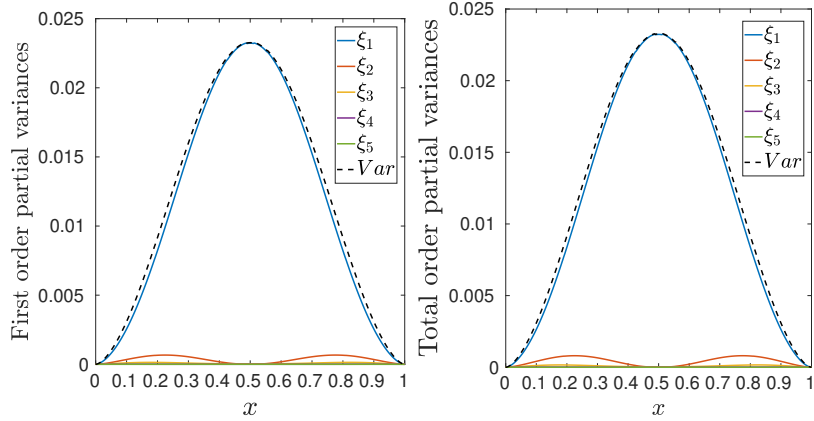


Figure 9: Case 2: a priori-variability of  $\hat{u}$  due to  $\xi$ , for distributed  $f$ . First-order partial variances (left) and total-order partial variances (right) are shown, as well as the variance.

For the sensitivity of  $u$  to the KL coordinates, Figure 10 reports first-order and total-order partial variances of  $u$  due to the KL coordinates  $\xi_i$  when  $f$  is localized at  $x_0 = 0.1$ , and at  $x_0 = 0.5$ . The figure indicates that most variability in  $u$  is due to the first mode  $\xi_1$  for the two forcing locations, while  $\xi_2$  has an appreciable effect only when  $x_0 = 0.1$ . Other coordinates are seen to be unimportant regardless of the forcing location.

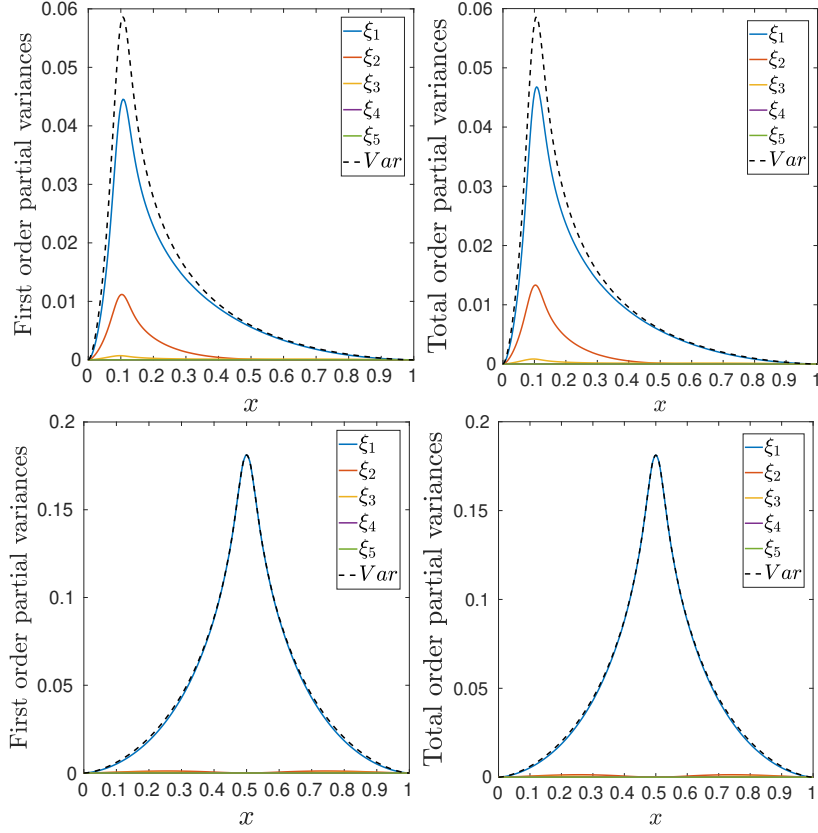


Figure 10: Case 2: a priori-variability of  $\hat{u}$  due to  $\xi$ , for localized  $f$  with  $x_0 = 0.1$  (top row) and  $x_0 = 0.5$  (bottom row). First-order partial variances (left) and total-order partial variances (right) are shown, as well as the variance.

### 6.2.2. Inference of the field's coordinates

We now seek to calibrate the coordinates  $\xi = (\xi_1, \dots, \xi_5)^T$  of the KL expansion of  $\log \kappa$ . The calibration uses noisy observations of the solution at a finite set of observation points,  $x_i$ . As for Case 1, we consider two different types of observations. In the first one, a single experiment is performed with a distributed forcing, and observations are collected at  $m_d = 5$  observation stations,  $x_i = 0.1, 0.3, 0.5, 0.7$  and  $0.9$ . In the second type, we consider  $m_d$  experiments, with localized forcing centered at  $x_0 = 0.1, 0.3, 0.5, 0.7$  and  $0.9$ , with a single observation point at  $x_0$ .

We performed the calibration for two sets of synthetic data. These correspond to different instances of the true  $\xi$ ; Instance I has

$$\xi_I = (0.5377, 1.8339, -2.2588, 0.8622, 0.3188)^T,$$

whereas Instance II has

$$\xi_{II} = (0.2521, 0.3124, -0.4588, 0.1622, 0.0188)^T.$$

The noisy observations are generated using

$$\tilde{u}_i = \hat{u}(x_i; \xi_{I,II})(1 + \epsilon_i), \quad i = 1, \dots, m_d,$$

where the  $\epsilon_i$ 's are i.i.d. Gaussian with zero mean and standard deviation  $\tilde{\rho}_t = 0.01$  and  $\xi_{I,II}$  denotes one of the two instances. The inference relies on the following expression of the log-likelihood:

$$\ln(\mathcal{L}(\mathbf{d}_u | \xi, \rho)) \propto -\frac{1}{2} \sum_{i=1}^{m_d} \ln \rho_i^2 - \frac{1}{2} \sum_{i=1}^{m_d} \left( \frac{\tilde{u}_i - \hat{u}(x_i, \xi)}{\rho_i} \right)^2, \quad (12)$$

where  $\mathbf{d}_u = (\tilde{u}_1, \dots, \tilde{u}_{m_d})^T$  is the observation vector, and  $\rho$  is the noise hyper-parameter that is assumed to follow a Jeffrey's prior.

We first comment on the inference results for Instance I. Table 2 presents the MAP and MLE values of the inferred  $\xi_i$  and noise level  $\rho$ , for the two forcing types. The table shows that the MLE values of the first three coordinates are in better agreement with the true values compared to the MAP. The MAP value of  $\rho$  is significantly higher than the actual value used to corrupt the synthetic data. Note that, in light of the sensitivity analysis presented in the previous section, one cannot expect to recover the coordinates of the higher modes from limited noisy observations because these coordinates have an insignificant impact on the solution.

Table 2: Case 2, Instance I: MLE and MAP estimates of the field coordinates and observation noise. The true values are also reported.

Coordinate	$\xi_1$	$\xi_2$	$\xi_3$	$\xi_4$	$\xi_5$	$\rho$
Distributed forcing						
MLE	0.4414	1.6003	-0.6153	-0.2453	-0.0063	0.009
MAP	0.4149	0.2657	0.1801	-0.1129	0.2921	0.0324
Localized forcings						
MLE	0.4858	1.8044	-1.4765	-0.1959	-0.7040	0.0088
MAP	0.4226	0.4337	-0.0086	0.1788	0.1184	0.1479
Truth	0.5377	1.8339	-2.2588	0.8622	0.3188	0.01

The MLE and MAP estimators are further studied in Figure 11 in terms of solution profiles and diffusivity fields. Consistently with the results of Table 2, the plots reveal a closer agreement between the true solution and the corresponding MLE estimators than is obtained for the MAP. This is seen for both the distributed and localized forcing setups. The discrepancies between the MLE estimator  $\kappa$  and the truth are much more significant than the discrepancies between the MLE estimator of  $u$  and the true solution. Further, we remark that the localized forcing experiment gives a better MLE estimate of the diffusivity field, while the MAP estimator remains far from its true value in all forcing scenarios.

Figure 12 shows the posterior marginals of the field coordinates. The first coordinate,  $\xi_1$ , exhibits a marked peak, while the other coordinates essentially retain their prior shapes. One can interpret the present results in light of the dominant sensitivity of  $\xi_1$ , and the insignificant contribution of the remaining coordinates. Further, one notes that the second and third coordinates of the true field have large magnitudes, with quite unlikely a priori value; it is well-known [89] that, for an informative prior such as the Gaussian one, the Bayesian calibration needs a substantial amount of data to localize the posterior in a region of low prior probability. These two remarks explain why, in the present case, the Bayesian inference leads to broad predictive distributions, as further discussed below.

The results outlined above for Instance I are in contrast with those obtained for Instance II. Specifically, the MLE and MAP values of the coordinate  $\xi_i$  reported in Table 3 are now in much closer agreement with their true values, with differences not exceeding a fraction of the a priori standard deviation. Consistently, the corresponding MLE and MAP solution profiles, reported in Figure 13, are also closer to the true profile, for both the distributed and localized forcings. Also, the MAP and MLE estimates of the diffusivity fields are also closer to the true  $\kappa$ , compared to the case of Instance I shown in Figure 11, although differences are still significant.

Finally, the posterior marginals of the coordinates presented in Figure 14 reveal a more appreciable information gain for  $\xi_2$ , compared to Instance I in Figure 12, with a clear peak value close to, but distinct from, the origin. The second coordinate's improvement in information gain for  $\xi_2$  is more pronounced for the localized forcing than for distributed  $f$ ; this suggests that the asymmetrical structure of the localized forcing can better sense the anti-symmetry of the mode associated with this coordinate.

To provide additional insight into the inferred solutions, we plot in Figures 15 and 16 the solution,  $u$ , the 95% bounds of the pushed-forward prior, and the 50% bounds of the pushed-forward posterior, respectively

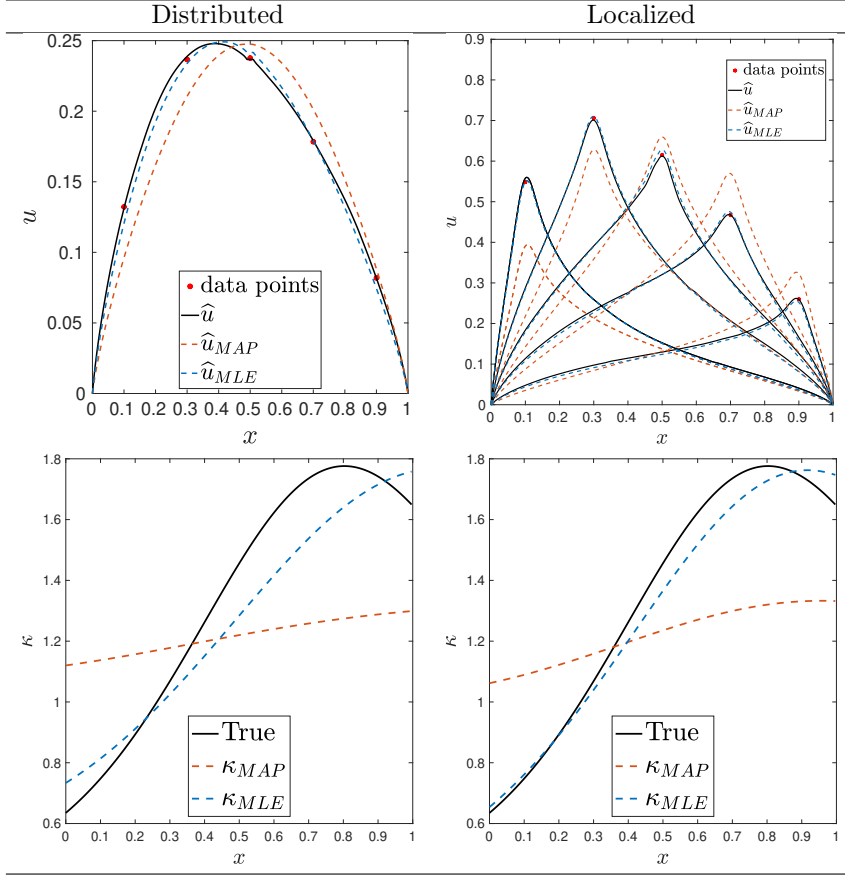


Figure 11: Case 2, Instance I. Top row: true, MAP and MLE solution profiles. Bottom row: true, MAP and MLE diffusivity fields. Inference is performed using noisy observations, for both distributed and localized forcing, as indicated.

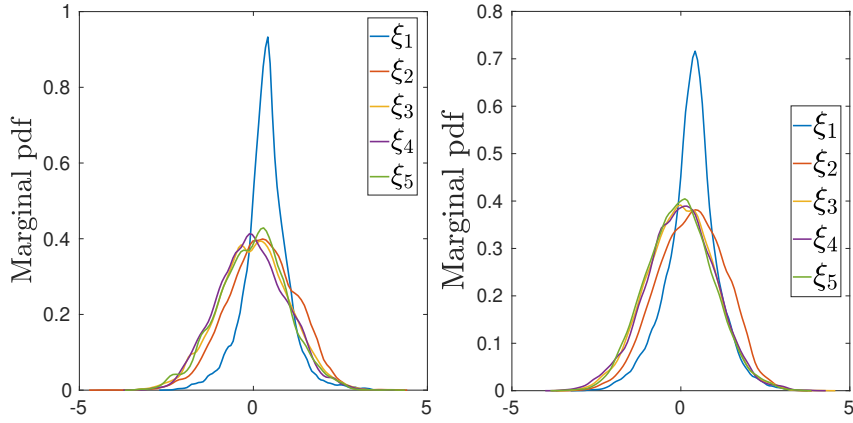


Figure 12: Case 2, Instance I: posterior marginal pdfs of diffusion field coordinates  $\xi_1, \dots, \xi_5$  for distributed (left) and localized (right) forcing.

for distributed and localized forcing. Contrasted in each of these figures are results obtained for Instances I and II. Unlike Instance II, one can observe that the true solution in Instance I can lie appreciably outside the bounds depicted. This is consistent with the previous discussion concerning the fact that a parameter with low prior probability was drawn in Instance I. A larger amount of data would be needed to better

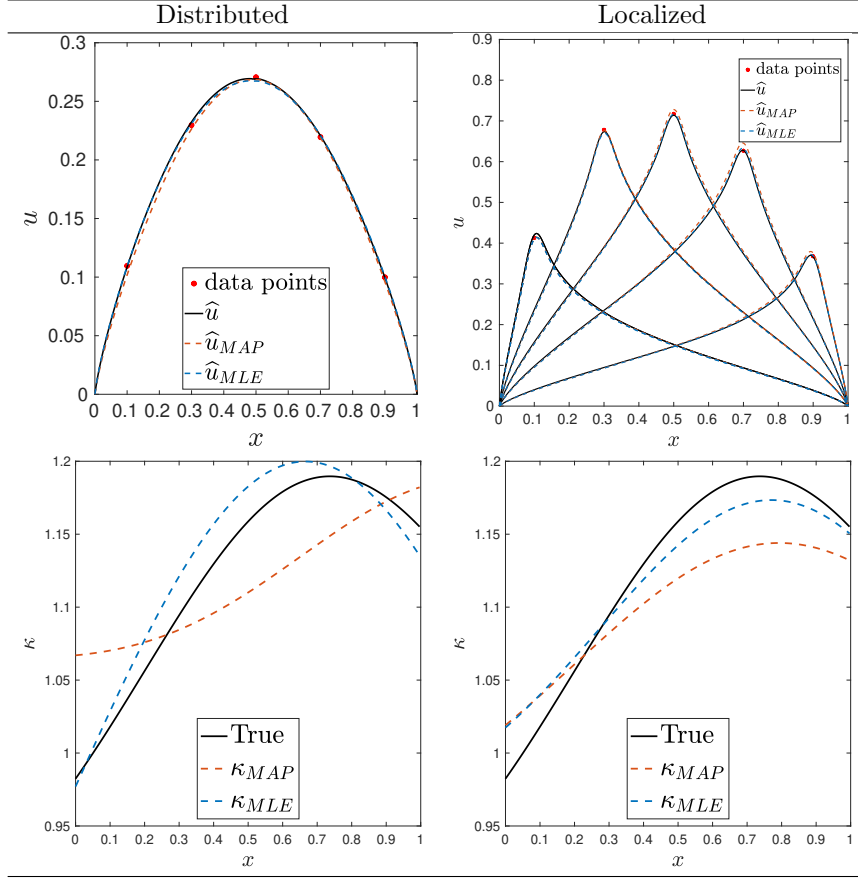


Figure 13: Case 2, Instance II. Top row: true, MAP and MLE solution profiles. Bottom row: true, MAP and MLE diffusivity fields. Inference is performed using oisy observations, for distributed and localized forcing, as indicated.

estimate the field in these situations.

### 6.3. CASE 3

In Case 3, we aim to infer  $\alpha$  and coordinates of the diffusivity field  $\kappa$ . Based on the results of Section 6.2, we restrict the parametrization of  $\kappa$  to the first two KL modes, setting  $\xi_i = 0$  for  $i > 2$ . The prior correlation length and standard deviation of the log-normal distribution of  $\kappa$  are held fixed,  $L_M = 0.75$  and  $\sigma_M = 0.5$ ,

Table 3: Case 2, Instance II: MLE and MAP estimates of the field coordinates and observation noise. The true values are also reported.

Coordinate	$\xi_1$	$\xi_2$	$\xi_3$	$\xi_4$	$\xi_5$	$\rho$
Distributed forcing						
MLE	0.2760	0.2712	-0.6679	-0.0906	-0.6957	0.0044
MAP	0.2321	0.1967	0.2396	0.0910	-0.1131	0.0253
Localized forcings						
MLE	0.2442	0.2499	-0.2431	0.2911	-0.0457	0.0066
MAP	0.2110	0.2016	-0.1698	0.0894	-0.0630	0.0882
Truth	0.2521	0.3124	-0.4588	0.1622	0.0188	0.01

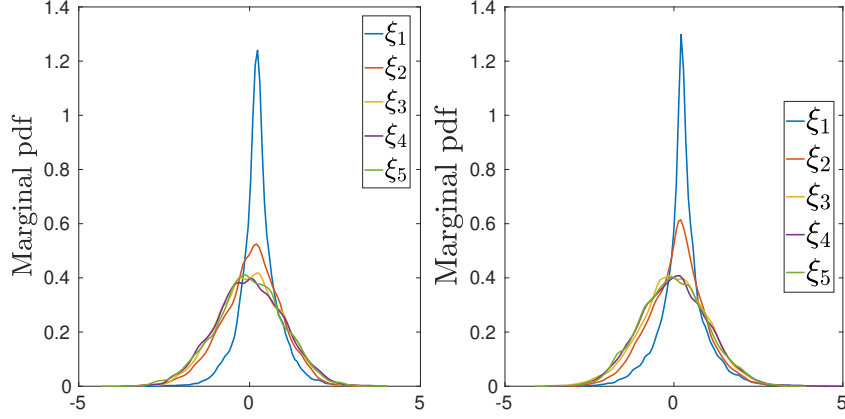


Figure 14: Case 2, Instance II: posterior marginal pdfs of diffusion field coordinates  $\xi_1, \dots, \xi_5$  for distributed (left) and localized (right) forcing.

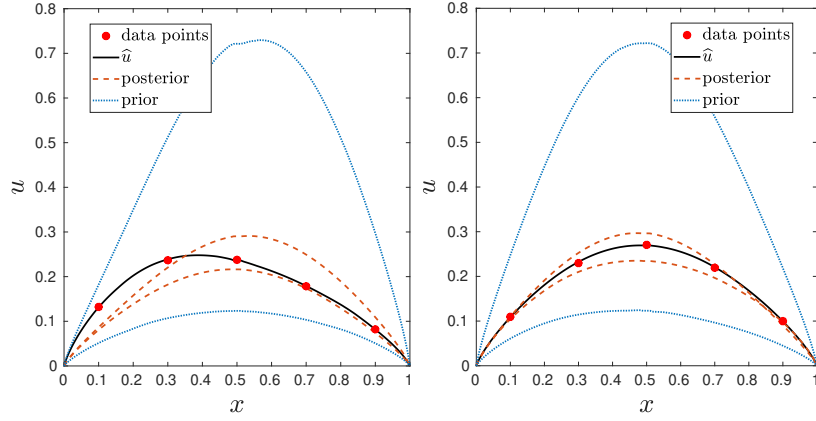


Figure 15: Case 2. Profiles of the solution  $u$ , the 95% bounds of the solution prior (blue dash) and the 50% bounds of the solution posterior (red dash) of  $\xi$ . Shown are solutions obtained using localized forcing; left: Instance I, right: Instance II. In both cases, the noisy measurements used for the inference are depicted.

respectively. We first construct an approximation of  $u$  in terms of  $\alpha$ ,  $\xi_1$ , and  $\xi_2$ , using a NISP approach to estimate the PC coefficients. Here, the multidimensional quadrature points resulting from the tensorization of 8 Gauss-Legendre ( $\alpha$  uniformly distributed) and Gauss-Hermite ( $\xi_1, \xi_2$  i.i.d. Gaussian) nodes. Similarly, the PC basis consists of fully tensorized, Legendre and Hermite polynomials with maximum (partial) order 7. In Figure 17, the PC surrogate,  $\hat{u}$ , of  $u$  is compared with the solution for a selected value of the random input vector, considering both distributed and localized forcing. A close agreement between  $u$  and  $\hat{u}$  is seen for all instances depicted in Figure 17.

### 6.3.1. Global Sensitivity Analysis

The Sobol decomposition of the variance of  $u$  is presented in Figure 18 for distributed forcing, and in Figure 19 for localized forcing. First-order and total-order variances are plotted to illustrate the contributions of the three uncertain inputs,  $\alpha$ ,  $\xi_1$ , and  $\xi_2$ . In all cases, the results indicate that  $\alpha$  has the dominant role on the variability of the solution, followed by  $\xi_1$ , whereas the impact of  $\xi_2$  is comparatively negligible. First and total-order partial variances due to  $\xi_1$  peak in the middle of the domain for the distributed forcing, and at  $x_0$  for the localized one. One can also deduce from the curves that mixed interactions between  $\alpha$  and  $\xi_1$  are appreciable at locations around the maximum of the forcing.



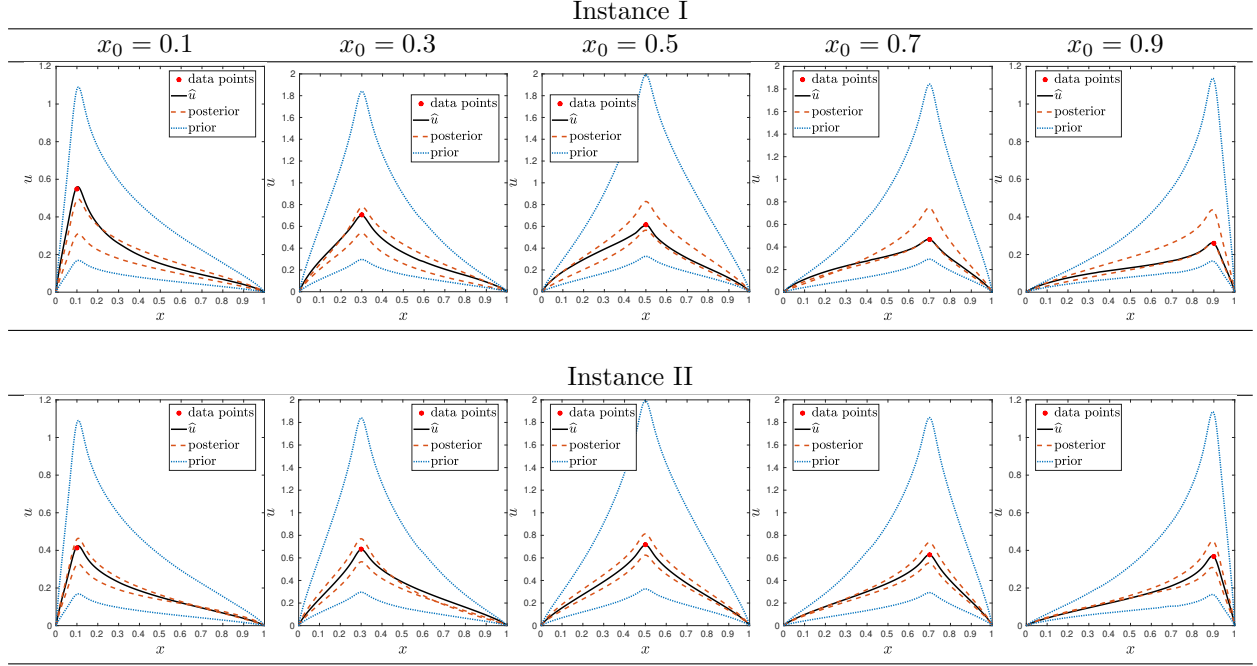


Figure 16: Case 2. Profiles of the solution,  $u$ , the 95% bounds of the pushed-forward prior (blue dash) and the 50% pushed-forward bounds of the posterior (red dash) of  $\xi$ . The plots correspond to localized forcing experiments as indicated. Top row: Instance I; bottom row: Instance II. The noisy measurements used in the inference are also depicted.

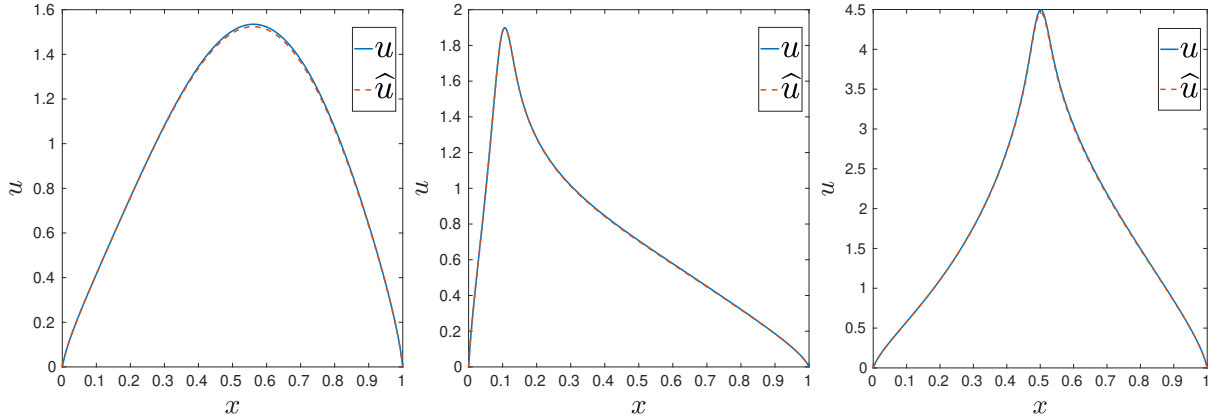


Figure 17: Case 3: solutions,  $u$ , and surrogates,  $\hat{u}$ , for distributed forcing (left), and localized  $f$  with  $x_0 = 0.1$  (center) and  $x_0 = 0.5$  (right). Solutions correspond to  $\alpha = 0.3377$ ,  $\xi_1 = -2.3668$ , and  $\xi_2 = -1.1544$ .

### 6.3.2. Inference of $\alpha$ and $\kappa$

To estimate  $\alpha$ ,  $\xi_1$  and  $\xi_2$ , we consider noisy observations of the solution,

$$\tilde{u}_i = \hat{u}(x_i; \alpha, \xi_t)(1 + \epsilon_i), \quad i = 1, \dots, m_d,$$

where  $m_d$  is the number of observations, the  $x_i$ 's are the observation locations, the  $\epsilon_i$ 's are i.i.d. centered Gaussians with standard deviation  $\tilde{\rho}_t = 0.01$ , and  $\xi_t$  is the vector of true coordinates. For the inference below, the true values of the parameters are randomly drawn from their respective prior distributions; we have  $\alpha = 0.5521$ ,  $\xi_t = (0.4124, -0.358)^T$ . Again, we consider both distributed and localized forcing experiments, with  $m_d = 5$  observations/forcing points located at  $x_i = 0.1, 0.3, 0.5, 0.7$ , and  $0.9$ . The same

noise level  $\tilde{\rho}_t = 0.01$  is used for the two forcing types. For both setups, the log-likelihood is as follows

$$\ln(\mathcal{L}(\mathbf{d}_u|\alpha, \rho)) \propto -\frac{1}{2} \sum_{i=1}^{m_d} \ln \rho_i^2 - \frac{1}{2} \sum_{i=1}^{m_d} \left( \frac{\tilde{u}_i - \hat{u}_i}{\rho_i} \right)^2, \quad (13)$$

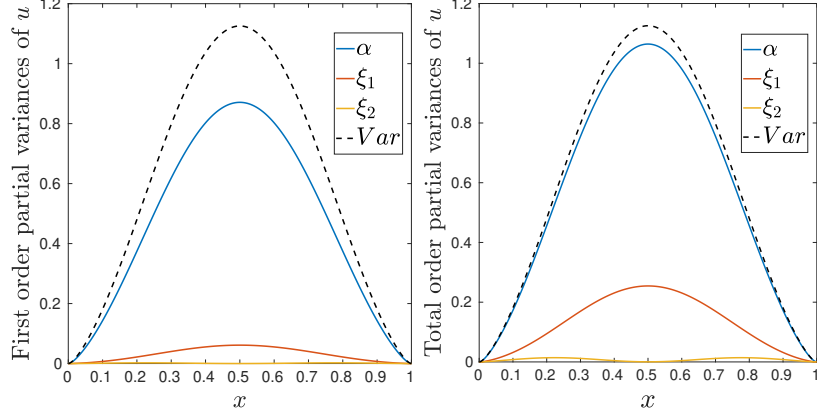


Figure 18: Case 3: first-order (left) and total-order (right) partial variances of  $\hat{u}$  due to  $\alpha$ ,  $\xi_1$  and  $\xi_2$ , for the distributed forcing case.

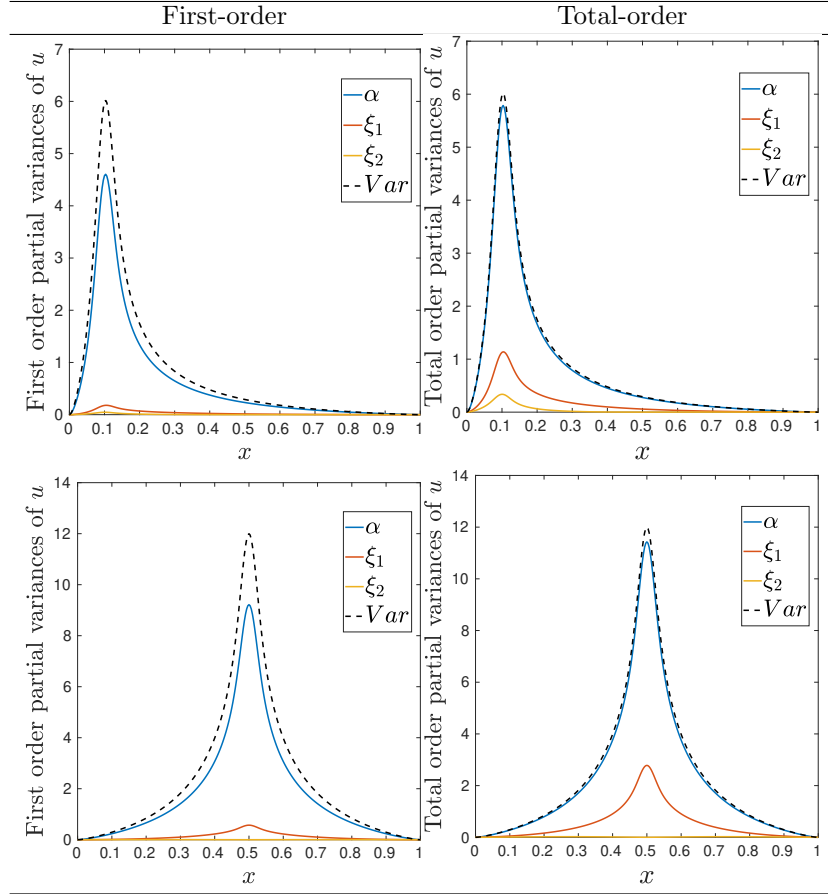


Figure 19: Case 3: first-order and total-order partial variances of  $\hat{u}$  due to  $\alpha$ ,  $\xi_1$  and  $\xi_2$ . Plotted are results obtained using localized forcing with  $x_0 = 0.1$  (top row) and  $x_0 = 0.5$  (bottom row).

where  $\hat{u}_i \equiv \hat{u}(x_i; \alpha, \xi_t)$ ,  $\mathbf{d}_u = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_{m_d})^T$  is the observation vector, and  $\rho$  is a hyper-parameter equipped with Jeffrey's prior.

Table 4 summarizes the MAP and MLE estimates obtained from the Bayesian calibration for the distributed and localized forcing setups. The table reveals that the MLE values of  $\alpha$ ,  $\xi_1$ , and  $\xi_2$  are in better agreement with their respective true values than the MAP estimates. In particular, the MAP estimators of  $\alpha$  over-estimate the true values for both setups. Further, the noise level is correctly estimated by the MLE value but over-estimated by the MAP value. Similar to Case 2, more observations would undoubtedly help to improve the MAP estimators.

Figure 20 compares the true solution profiles and the diffusivity field with their MAP and MLE counterparts in the distributed and localized forcing experiments. We see that the MLE profiles are in excellent agreement with the corresponding true profiles. In contrast, substantial discrepancies are observed between the MAP profiles, especially for the distributed forcing experiment. The differences between the MLE and MAP profiles reflect the calibration results in Table 4. Concerning the diffusivity fields, Figure 20 depicts the closer agreement of  $\kappa$  using MLE, especially in the distributed forcing scenario. The present results are also consistent with previous ones, which indicated that, at least for the present framework, sampling data at several locations in a single distributed forcing experiment provides a better means to recover the parameters than collecting a single response from the same number of localized forcing experiments.

The posterior marginals of the inferred parameters are plotted in Figure 21. The curves indicate that for both distributed and localized experiments, the posterior marginals for  $\alpha$  have a sharp peak. The marginals of  $\xi_1$  and  $\xi_2$  exhibit discernible peaks, although only the marginal of  $\xi_1$  appears to differ significantly from its (Gaussian) prior. Again, this behavior was expected from the results of the global sensitivity analysis.

#### 6.4. CASE 4

In the last case, we focus on the inference of the statistical characteristics of the stochastic diffusivity field jointly with the fractional order  $\alpha$ . Specifically, besides  $\alpha$ , we try to infer the correlation length,  $L_M$ , and standard deviation,  $\sigma_M$ , of the Gaussian field underlying  $\kappa$ . The inference uses the mean and standard deviation of the solution to calibrate the parameters. At given values of  $\alpha$ ,  $L_M$ , and  $\sigma_M$ , the mean,  $\mu$ , and standard deviation,  $\sigma$ , of the solution are estimated, sampling the first five KL coordinates of the corresponding  $\kappa$ . In this work, we rely on a large LHS strategy to obtain well-converged estimates of  $\mu$  and  $\sigma$ . Figure 22 compares the estimates of  $\mu$  and  $\sigma$  obtained on LHS sets with sizes  $N_s = 15,000$  and  $25,000$ . The differences are not significant and, consequently, we used  $N_s = 25,000$  for the surrogate construction.

To build the surrogates for the dependencies of  $\mu$  and  $\sigma$  on  $\alpha$ ,  $L_M$ , and  $\sigma_M$ , we considered a sparse-grid quadrature with Féjer rules [71]. A detailed convergence study of the surrogates was conducted to assess the suitability of the PC approximation; a level 5 quadrature was found sufficient in this case. Figure 22 illustrates this study, contrasting the surrogates prediction for level 4 and level 5 in the sparse grid construction, considering a distributed forcing with the same values of  $L_M$ ,  $\sigma_M$  and  $\alpha$  as before. A similar exercise for the localized forcing functions, with peaks located at  $x_0 = 0.1, 0.3, 0.5, 0.7$ , and  $0.9$ , showed an excellent agreement between the direct estimates of  $\mu$  and  $\sigma$  and their sparse grid surrogate predictions, see Figure 23 for an illustration.

Table 4: Case 3: MLE and MAP values of  $\alpha$ ,  $\xi_1$ ,  $\xi_2$  and  $\rho$ , considering distributed and localized forcing experiments.

	$\alpha$	$\xi_1$	$\xi_2$	$\rho$
Distributed				
MLE	0.5642	0.3901	-0.5018	0.0020
MAP	0.6444	0.1946	-0.4802	0.0074
Localized				
MLE	0.5141	0.4688	-0.4631	0.0019
MAP	0.6741	0.1785	0.0622	0.0632
Truth	0.5521	0.4124	-0.358	0.0010

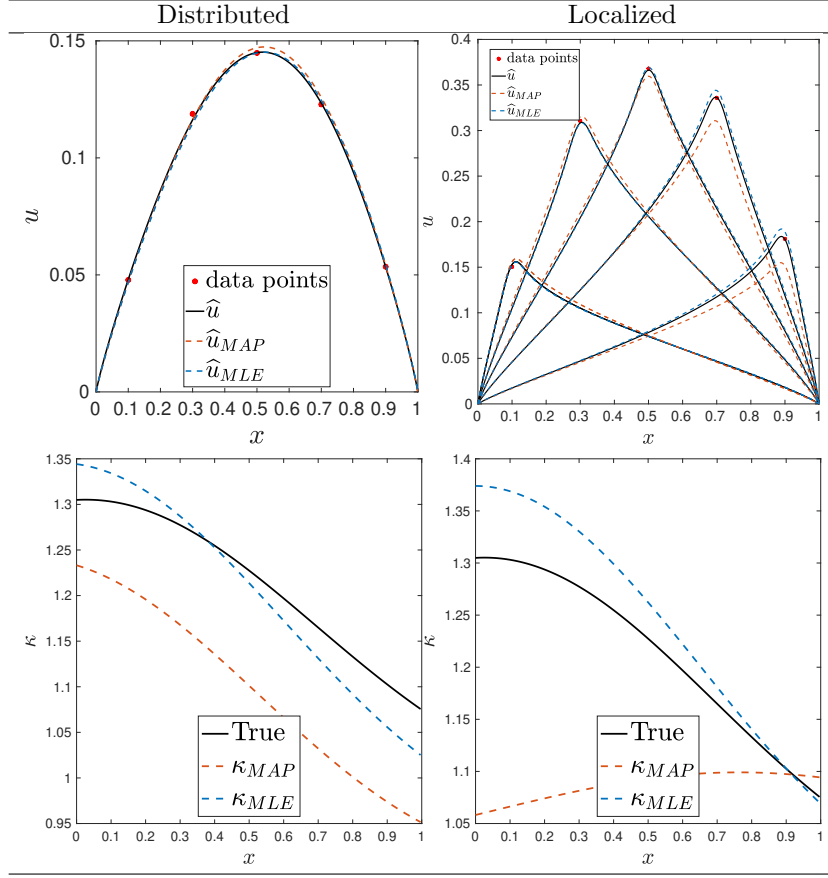


Figure 20: Case 3: MAP, MLE and true solution profiles (top row) and diffusivity fields (bottom row) for distributed and localized forcing experiments.

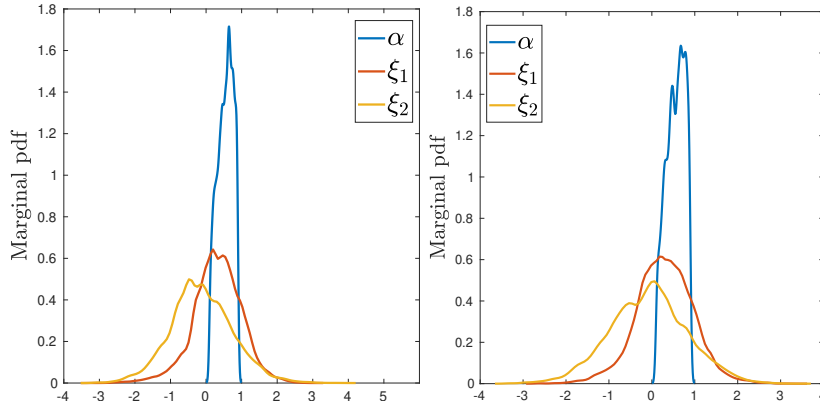


Figure 21: Case 3: marginal posterior of  $\alpha$ ,  $\xi_1$  and  $\xi_2$  for distributed (left) and localized (right) forcing experiments.

The sparse grid projection methods are susceptible to evaluation noise. Therefore, a high number of LHS samples are necessary to ensure a sufficiently converged estimation of  $\mu$  and  $\sigma$  at each of the sparse grid nodes. To alleviate this computational burden, one can instead rely on an alternative approach presenting a more robust behavior against evaluation noise. As a final check of our surrogates for  $\mu$  and  $\nu$ , we reduced the LHS set size to  $N_s = 1,000$  for the estimation of the  $\mu$  and  $\sigma$ , and rely on a regression approach on LHS

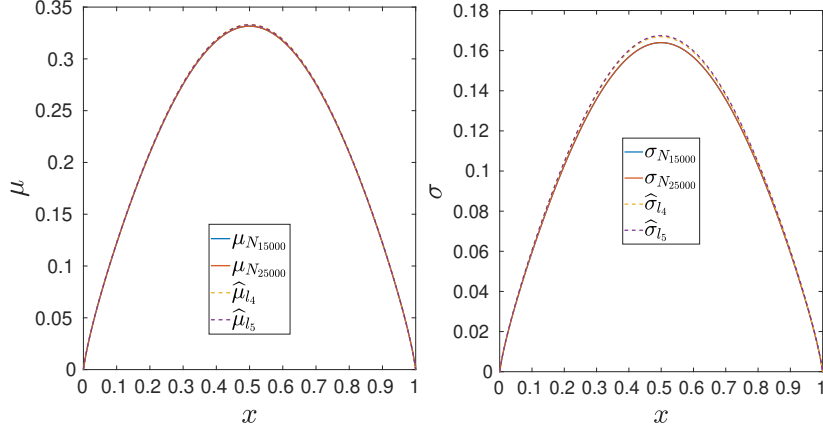


Figure 22: Case 4: estimates of  $\mu$  and  $\sigma$  using a LHS set with size  $N_s = 15,000$  and  $25,000$ , as well as level 4 and 5 NISP surrogates. Distributed forcing is considered, with a diffusivity field having  $L_M = 0.8889$  and  $\sigma_M = 0.5$ . The fractional order  $\alpha = 0.5$ .

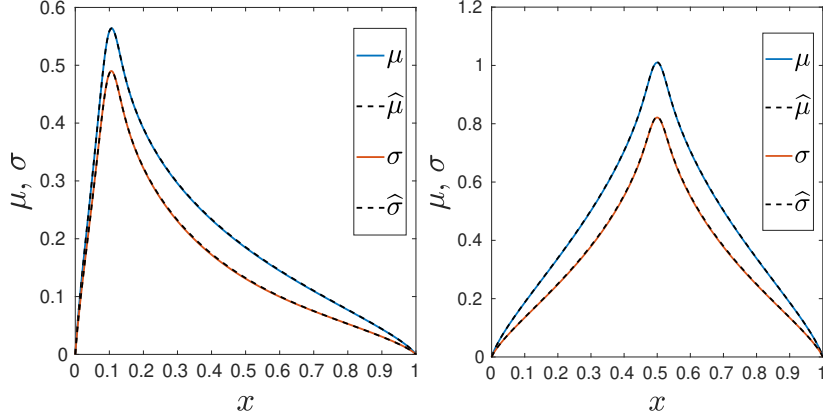


Figure 23: Case 4: profiles of  $\mu$  and  $\sigma$  compared to their surrogate predictions  $\hat{\mu}$  and  $\hat{\sigma}$ , for localized forcing at  $x_0 = 0.1$  (left) and  $x_0 = 0.5$  (right) and using  $L_M = 0.573$ ,  $\sigma_M = 0.778$  and  $\alpha = 0.5$ .

set of 300 triplets  $(L_M, \sigma_M, \alpha)$ . Figure 24 compares a particular realization of  $\mu$  and  $\sigma$  and their regression surrogate approximations  $\hat{\mu}$  and  $\hat{\sigma}$  for distributed and localized forcing at  $x_0 = 0.1$  and  $x_0 = 0.5$ . A detailed analysis (not shown) proved that the regression surrogates achieve an accuracy comparable to the sparse grid ones, but for just a fraction of the computational cost. These experiments also demonstrated that the sparse grid surrogates are not affected by the estimation noise, and they will be used in the remainder of the section.

#### 6.4.1. Global Sensitivity Analysis

Results of the variance decomposition are provided in Figure 25, considering distributed and localized forcings. Concerning the solution mean  $\mu$  (top row), we only show the total-order partial variances due to  $\alpha$ ,  $L_M$  and  $\sigma_M$ , because the contribution of  $\alpha$  is highly dominant,  $L_M$  and  $\sigma_M$  have weak impacts, and interaction effects are negligible for the three cases considered. In contrast, the variance decomposition of the solution's standard deviations shows non-negligible contributions and interactions between  $\alpha$  and  $\sigma_M$ . At the same time, the correlation length has an insignificant impact on  $\sigma$ , with a total-order variance due to  $L_M$  substantially smaller than the others.

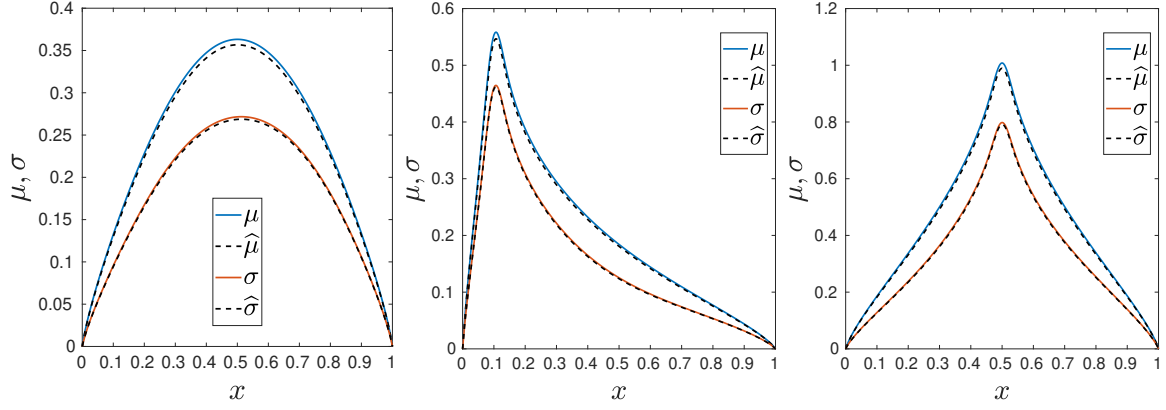


Figure 24: Case 4: direct estimates of  $\mu$  and  $\sigma$  and their regression surrogates  $\hat{\mu}$  and  $\hat{\sigma}$ ; distributed (left) and localized forcing at  $x_0 = 0.1$  (center) and  $x_0 = 0.5$  (right). In all cases,  $L_M = 0.573$ ,  $\sigma_M = 0.778$ , and  $\alpha = 0.5$ .

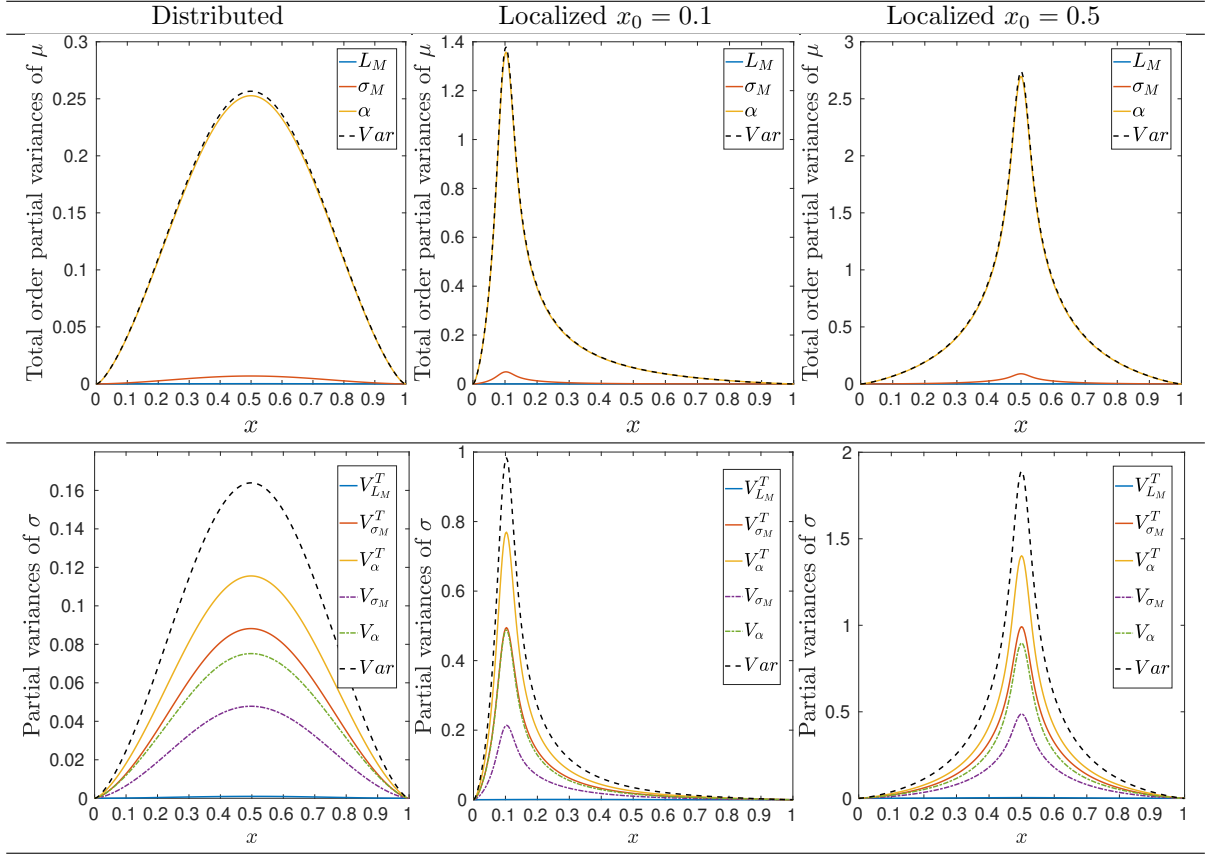


Figure 25: Case 4: total-order partial variances of  $\mu$  (top) and first-order and total-order partial variances of  $\sigma$  (bottom), for different forcing setups as indicated.

#### 6.4.2. Inference of $L_M$ , $\sigma_M$ and $\alpha$

We consider the problem of inferring  $L_M$ ,  $\sigma_M$ , and  $\alpha$  from observations of the solution statistics, namely the solution's mean  $\mu$  and standard deviation  $\sigma$ . Noisy observations of the mean,  $\tilde{\mu}_i$ , and standard deviation,

$\tilde{\sigma}_i$ , are generated at selected points  $x_i$ , from their respective PC surrogates, according to:

$$\tilde{\mu}_i = \hat{\mu}(x_i; \boldsymbol{\eta}_t)(1 + \epsilon_{1,i}), \quad \tilde{\sigma}_i = \hat{\sigma}(x_i; \boldsymbol{\eta}_t)(1 + \epsilon_{2,i}), \quad i = 1, \dots, m_d,$$

where  $\boldsymbol{\eta}_t$  is the vector of true values for  $\boldsymbol{\eta} = (L_M, \sigma_M, \alpha)^T$ , and  $\epsilon_{1,i}$  and  $\epsilon_{2,i}$  are assumed to be i.i.d. Gaussian random variables with zero mean and standard deviations  $\tilde{\rho}_1$  and  $\tilde{\rho}_2$ , respectively. We randomly selected  $m_d = 16$  observation points,  $x_i$ , and performed the inference using the log-likelihood is defined by

$$\ln(\mathcal{L}(\mathbf{d}_\mu, \mathbf{d}_\sigma | \boldsymbol{\eta}, \boldsymbol{\rho})) \propto -\frac{1}{2} \sum_{i=1}^{m_d} \left[ \ln \rho_{1,i}^2 + \ln \rho_{2,i}^2 + \left( \frac{\tilde{\mu}_i - \hat{\mu}(x_i; \boldsymbol{\eta})}{\rho_{1,i}} \right)^2 + \left( \frac{\tilde{\sigma}_i - \hat{\sigma}(x_i; \boldsymbol{\eta})}{\rho_{2,i}} \right)^2 \right], \quad (14)$$

where  $\mathbf{d}_\mu = (\tilde{\mu}_1, \dots, \tilde{\mu}_{m_d})^T$  and  $\mathbf{d}_\sigma = (\tilde{\sigma}_1, \dots, \tilde{\sigma}_{m_d})^T$  respectively denote the vector of observations of the solution mean and standard deviation, and  $\boldsymbol{\rho} = (\rho_1, \rho_2)^T$  is the vector of hyper-parameters, whose components are assumed to follow independent Jeffrey's prior distributions.

Table 5: Case 4: MLE and MAP values of  $L_M, \sigma_M, \alpha$  and noise hyper-parameters  $\rho_1$  and  $\rho_2$  for different forcings as indicated.

Forcing	$L_M$	$\sigma_M$	$\alpha$	$\rho_1$	$\rho_2$
MLE					
Distributed	0.8021	0.4235	0.6244	0.0331	0.0058
Localized ( $x_0 = 0.1$ )	0.6595	0.3981	0.5922	0.0275	0.0070
Localized ( $x_0 = 0.5$ )	0.5278	0.4128	0.6038	0.0552	0.0146
MAP					
Distributed	0.9512	0.4135	0.6170	0.0622	0.0302
Localized ( $x_0 = 0.1$ )	0.6832	0.3981	0.5999	0.0613	0.0463
Localized ( $x_0 = 0.5$ )	0.9655	0.4058	0.6111	0.0954	0.0213
True values	0.7500	0.4000	0.6000	0.0700	0.0350

The calibration is performed considering distributed and localized forcing. Table 5 provides the computed MLE and MAP values of  $L_M, \sigma_M, \alpha$ , and noise parameters; the last row of Table 5 also shows the true values used for the generation of the noisy observations. The MAP and MLE values of  $\sigma_M$  and  $\alpha$  are in excellent agreement and close to the true values. In contrast, the estimates of the correlation length differ greatly depending on the experiment. This finding is consistent with the negligible impact of  $L_M$  on the mean and standard deviation of the solution, reported in the sensitivity analysis. The inferred hyper-parameters,  $\rho_1$  and  $\rho_2$ , differ significantly from their true values, the MAP estimates being closer than the MLE values, which consistently underestimate the noise level.

Figure 26 shows the true profiles of  $\mu$  and  $\sigma$  and their corresponding MAP and MLE estimates using the values reported in Table 5. Also shown are the  $m_d = 16$  noisy observations used for the calibration. The inferred MAP and MLE profiles are in excellent agreement with the corresponding truth profiles for both the mean and standard deviation and all three forcing functions considered.

Finally, Figure 27 presents the marginal posteriors of  $L_M, \sigma_M$  and  $\alpha$ . As expected, the marginal posteriors of  $\sigma_M$  and  $\alpha$  are concentrated and exhibit unique, well-defined maxima. On the other hand, we note that the marginal pdf of  $L_M$  remains flat and exhibits several equivalent local maxima. The flatness of the marginal posterior of  $L_M$  is consistent with the findings of the sensitivity analysis, discussed above, which indicates that it has an insignificant impact on  $\mu$  and  $\sigma$ . Consequently, one expects the present observations to be essentially uninformative about this parameter, such that it essentially retains its flat uniform prior.

## 7. Conclusions

In this paper, we have investigated a steady, 1D, fractional diffusion equation involving uncertain fractional order and uncertain, spatially-varying, diffusivity field. Four different cases were analyzed, corre-

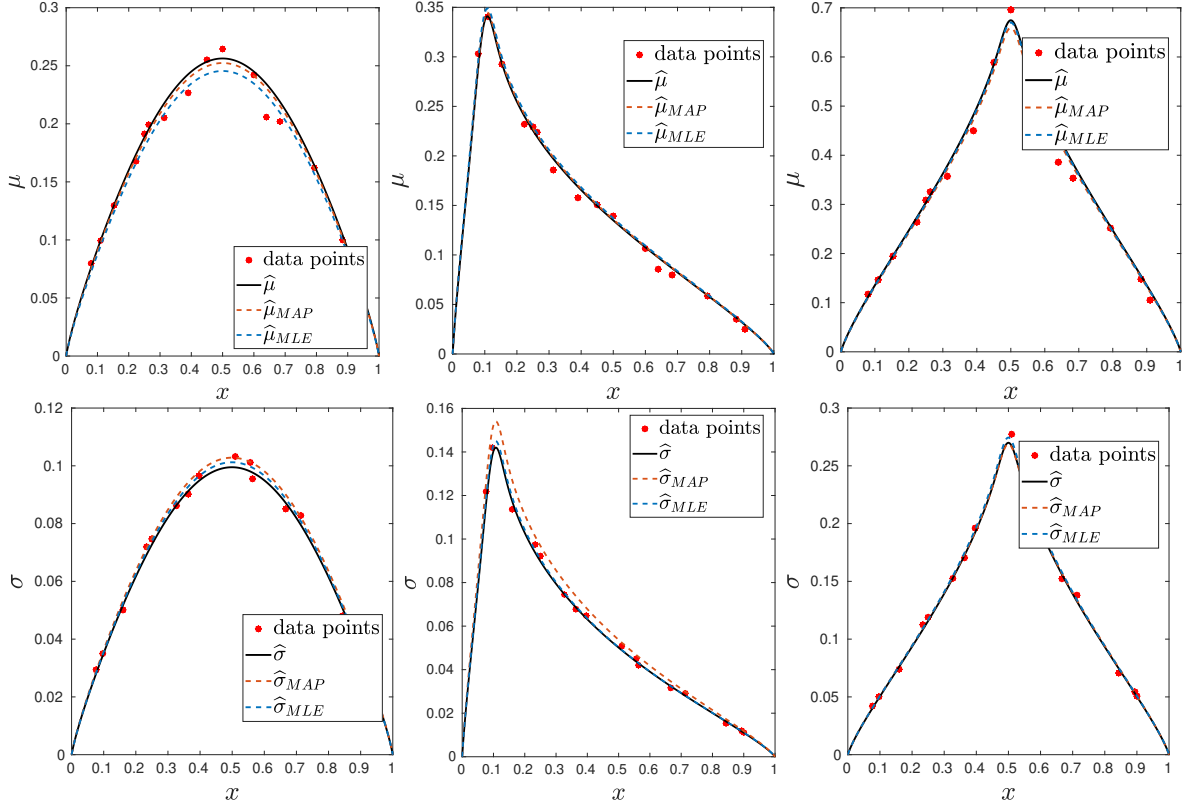


Figure 26: Case 4: MAP and MLE profiles of  $\mu$  (top) and  $\sigma$  (bottom) for distributed forcing (left) and localized forcing with  $x_0 = 0.1$  (center) and  $x_0 = 0.5$  (right). The true profiles are also plotted.

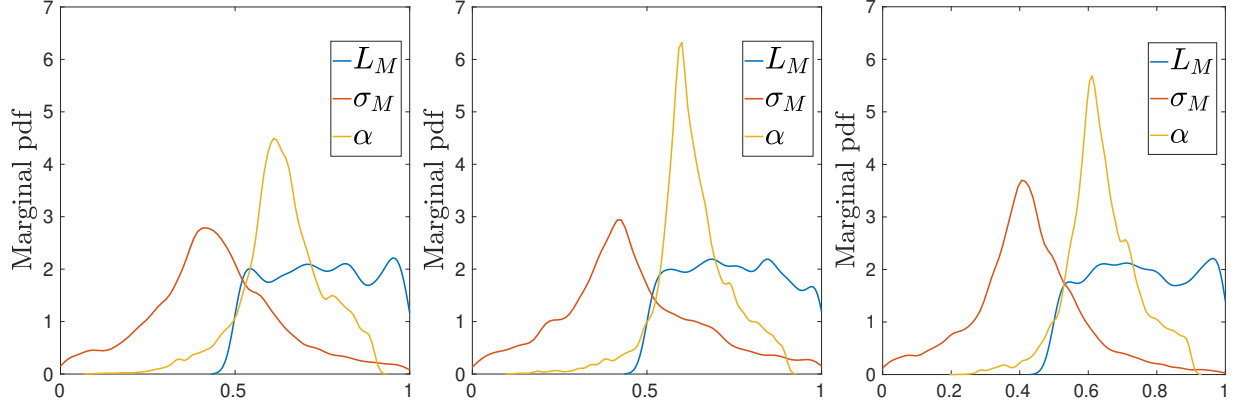


Figure 27: Case 4: marginal posterior pdfs of  $L_M$ ,  $\sigma_M$  and  $\alpha$  for distributed (left) and localized forcing at  $x_0 = 0.1$  (center) and  $x_0 = 0.5$  (right).

sponding to different combinations of random inputs, QoIs, and localized or distributed forcing terms. For all cases, we have constructed surrogate models for the QoIs, using non-intrusive PC approaches. With these PC surrogates, we have conducted sensitivity analyses of the QoIs. In particular, these analyses revealed the dominant role of the fractional order,  $\alpha$ , followed by the first KL coordinates involved in the diffusion field's parametrization (large scale fluctuations). The sensitivity analyses also showed that the correlation length of the Gaussian process underlying the diffusion field has a weak impact on the solution mean and



variance, at least for the range of values considered.

We also explored the calibration of the uncertain inputs, using direct observations of the solution or some of its statistics. A Bayesian formalism was specifically adopted with synthetic observation derived from the surrogates. In particular, the experiments showed that the dominant parameters, such as the fractional order,  $\alpha$ , can be robustly inferred using limited information. The investigations also showed that using the same number of noisy measurements in different experimental setups may lead to appreciably different inverse solutions. This suggests the application of optimal experimental design methodologies [90, 91] to quantify the utility of different setups. This will be addressed in future work, which will also consider the extension of the existing approaches to higher dimensions, as well as the incorporation of fast solvers [53, 54, 47] to enhance the efficiency of the surrogate construction.

## Acknowledgments

Research reported in this publication was supported by research funding from King Abdullah University of Science and Technology (KAUST). The computer codes for the simulation of the space-fractional diffusion equation are available by request from the authors.

## References

### References

- [1] Cushman J and Moroni M 2001 *Phys. Fluids* **13** 75
- [2] Moroni M and Cushman J 2001 *Phys. Fluids* **13** 81
- [3] Snyder M, Knio O, Katz J and Le Maître O 2007 *Phys. Fluids* **19** 065108
- [4] Ferrás L, Ford N, Morgado M, Rebelo M, McKinley G and Nóbrega J 2018 *Comput. Fluids* **174** 14–33
- [5] Amir S and Sun S 2018 *J. Petrol. Sci. Eng.* **163** 616–639
- [6] Du Q 2019 *Nonlocal Modeling, Analysis, and Computation* (Philadelphia: SIAM)
- [7] Benson D A, Wheatcraft S W and Meerschaert M M 2000 *Water Resour. Res.* **36** 1413–1423
- [8] Metzler R and Klafter J 2000 *Phys. Rep.* **339** 1–77
- [9] Podlubny I 1999 *Fractional Differential Equations* (San Diego, CA: Academic Press)
- [10] El-Khazali R 2013 *Comp. Math. Appl.* **66** 639–646
- [11] Klafter J and Sokolov I M 2005 *Phys. World* **18** 29
- [12] Gatto P and Hesthaven J S 2015 *J. Sci. Comput.* **65** 249–270
- [13] Kou S G 2002 *Manage. Sci.* **48** 1086–1101
- [14] Alderremy A, Saad K, Gómez-Aguilar J, Aly S, Kumar D and Singh J 2021 *Math. Meth. Appl. Sci.* URL <https://doi.org/10.1002/mma.7188>
- [15] Bahloul M and Laleg Kirati T 2021 *Physiol. Meas.* **42** 045008
- [16] Li C and Cai M 2020 *Theory and Numerical Approximations of Fractional Integrals and Derivatives* (Philadelphia: SIAM)
- [17] Mustapha K, Furati K, Knio O and Le Maître O 2020 *Comm. Appl. Math. Comput.* **2** 671–688
- [18] Çelik C and Duman M 2012 *J. Comput. Phys.* **231** 1743–1750
- [19] Deng W H and Hesthaven J S 2013 *Esaim Math. Model. Numer. Anal.* **47** 1845–1864
- [20] Ding H, Li C and Chen Y 2015 *J. Comput. Phys.* **293** 218–237
- [21] Kharazmi E, Zayernouri M and Karniadakis G 2017 *Comput Methods Appl Mech Eng* **324** 512–536
- [22] Li X and Xu C 2010 *Commun. Comput. Phys.* **8** 1016–1051
- [23] Lynch V, Carreras B, del Castillo-Negrete D, Ferreira-Mejias K and Hicks H 2003 *J. Comput. Phys.* **192** 406–421
- [24] Tadjeran C, Meerschaert M M and Scheffler H P 2006 *J. Comput. Phys.* **213** 205–213
- [25] Tian W, Zhou H and Deng W 2012 *Math. Comput.* **84** 1703–1727
- [26] Zheng X, Ervin V J and Wang H 2020 *Comput. Methods Appl. Math.* **20** 573–589
- [27] Lynch V E, Carreras B A, del Castillo-Negrete D, Ferreira-Mejias K and Hicks H 2003 *J. Comput. Phys.* **192** 406–421
- [28] Zhang H, Liu F and Anh V 2010 *Appl. Math. Comput.* **217** 2534–2545
- [29] Li X and Xu C 2010 *Commun. Comput. Phys.* **8** 1016–1051
- [30] Mustapha K 2010 *IMA J. Numer. Anal.* **31** 719–739
- [31] Mustapha K and McLean W 2013 *SIAM J. Numer. Anal.* **51** 491–515
- [32] Mustapha K, Abdallah B and Furati K 2014 *SIAM J. Numer. Anal.* **52** 2512–2529
- [33] Zhou H, Tian W and Deng W 2013 *J. Sci. Comp.* **56** 45–66
- [34] Mustapha K 2015 *Numer. Math.* **130** 497–516
- [35] Saad K and Alqhtani M 2021 *AIMS Math.* **6** 3788–3804
- [36] Alomari A, Abdeljawad T, Baleanu D, Saad K and Al-Mdallal Q 2021 *Numer. Methods Partial Differ. Equ.* URL <https://doi.org/10.1002/num.22699>
- [37] Ervin V J and Roop J P 2006 *Numer. Methods Partial Differ. Equ.* **22** 558–576

- [38] Chechkin A V, Klafter J and Sokolov I M 2003 *EPL* **63** 326–332
- [39] Liu Y, Yan Y and Khan M 2017 *Appl. Numer. Math.* **115** 200–213
- [40] Gorenflo R, Mainardi F, Moretti D, Pagnini G and Paradisi P 2002 *Chem. Phys.* **284** 521–541
- [41] Gorenflo R, Vivoli A and Mainardi F 2004 *Nonlinear Dynam.* **38** 101–116
- [42] Gorenflo R, Mainardi F and Vivoli A 2007 *Chaos Soliton. Fract.* **34** 87–103
- [43] Zhang H, Liu F and Anh V 2007 *J. Comput. App. Math.* **206** 1098–1115
- [44] Shi L, Yu Z, Mao Z, Xiao A and Huang H 2013 *Physica A* **392** 5801–5807
- [45] Luchko Y 2016 *Math. Model. Nat. Phen.* **11** 1–17
- [46] Lucchesi M, Allouch S, Le Maître O, Mustapha K and Knio O 2020 *Comput. Part. Mech.* **7** 491–507
- [47] Boukaram W, Lucchesi M, Turkiyyah G, Maître O L, Knio O and Keyes D 2020 *Comput. Methods Appl. Mech. Engrg.* **369** 113191
- [48] Wang H and Basu T S 2012 *SIAM J. Sci. Comput.* **34** A2444–A2458
- [49] Wang H and Zhang X 2015 *J. Comput. Phys.* **281** 67–81
- [50] Zheng X, Ervin V and Wang H 2019 *Appl. Math. Comput.* **361** 98–111
- [51] Feng L, Zhuang P, Liu F, Turner I and Yang Q 2015 *Prog. Fract. Differ. Appl.* **1** 23–35
- [52] Zhao M, Wang H and Cheng A 2018 *J. Sci. Comput.* **74** 1009–1033
- [53] Zhao X, Hu X, Cai W and Karniadakis G 2017 *Comput. Methods Appl. Mech. Engrg.* **325** 56–76
- [54] Karkulik M and Melenk J 2019 *Adv. Comput. Math.* **45** 2893–2919
- [55] Xu K and Darve E 2018 Efficient numerical method for models driven by Lévy process via hierarchical matrices arXiv:1812.08324
- [56] Gelman A, Carlin J B, Stern H S and Rubin D B 2004 *Bayesian Data Analysis* (Boca Raton: Chapman and Hall/CRC)
- [57] Zabaras N and Ganapathysubramanian B 2008 *J. Comput. Phys.* **227** 4697 – 4735
- [58] Ma X and Zabaras N 2009 *Inverse Probl.* **25** 035013
- [59] Koutsourelakis P S 2012 *Int. J. Numer. Methods Eng.* **91** 249–268
- [60] Iglesias M, Lin K and Stuart A 2014 *Inverse Probl.* **30** 114001
- [61] Contreras A, Le Maître O, Aquino W and Knio O 2016 *Probabilistic Eng. Mech.* **46** 107–119
- [62] Giraldi L, Le Maître O, Hoteit I and Knio O 2018 *Comput. Stat. Data Anal.* **124** 252–276
- [63] Ba Y, Jiang L and Ou N 2018 *J. Comput. Phys.* **374** 300–330
- [64] Jiang L and Ou N 2018 *Multiscale Model. Simul.* **16** 327–355
- [65] Saussereau B 2014 *Bernoulli* **20** 878–918
- [66] Pereira A, Fernandes J, Atman A and Acebal J 2018 *Physica A* **509** 369–382
- [67] Ba Y, Jiang L and Ou N 2019 *Int. J. Uncertain. Quantif.* **9** 245–273
- [68] Aldoghaither A and Laleg-Kirati T M 2020 *J. Comput. Appl. Math.* **369** 112570
- [69] El Hamidi A and Tfayli A 2020 *Math. Meth. Appl Sci.* **44** 8397–8413
- [70] Ghanem R G and Spanos P D 1991 *Stochastic Finite Elements: A Spectral Approach* (Berlin, Heidelberg: Springer-Verlag)
- [71] Le Maître O and Knio O 2010 *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics* (Berlin: Springer)
- [72] Sudret B 2008 *Reliab. Eng. Syst. Saf.* **93** 964–979
- [73] Crestaux T, Le Maître O and Martinez J 2009 *Rel. Eng. Syst. Saf.* **94** 1161–1182
- [74] Alexanderian A, Winokur J, Sraji I, Srinivasan A, Iskandarani M, Thacker W and Knio O 2012 *Comput. Geosci.* **16** 757–778
- [75] Tarantola A 2005 *Inverse Problem Theory and Methods for Model Parameter Estimation* (Philadelphia: SIAM)
- [76] Marzouk Y, Najm H and Rahn L 2007 *J. Comput. Phys.* **224** 560 – 586
- [77] Marzouk Y and Najm H 2009 *J. Comput. Phys.* **228** 1862–1902
- [78] Grigoriu M 2012 *Stochastic Calculus: Applications in Science and Engineering* (Boston: Birkhäuser Basel)
- [79] Marzouk Y, Najm H and Rahn L 2007 *J. Comput. Phys.* **224** 560–586
- [80] Blatman G and Sudret B 2008 *C.R. Mécanique* **336** 518–523
- [81] Tibshirani R 1996 *J. R. Stat. Soc. Ser. B Methodol.* **58** 267–288
- [82] Ng L and Eldred M 2012 *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*
- [83] Smolyak S 1963 *Dokl. Akad. Nauk SSSR* **148** 1042–1045
- [84] Sudret B 2008 *Reliab. Eng. Syst. Saf.* **93** 964–979
- [85] Sobol’ I 1993 *Math. Modeling Comput. Exp.* **1** 407–414
- [86] Haario H, Saksman E and Tamminen J 2001 *Bernoulli* **7** 223–242
- [87] Silverman B 1986 *Density Estimation for Statistics and Data Analysis* (London: Chapman and Hall)
- [88] Scott D 2015 *Multivariate density estimation* (New York: John Wiley & Sons, Inc.)
- [89] Sivia D and Skilling J 2006 *Data Analysis: A Bayesian Tutorial* (Oxford: Oxford University Press)
- [90] Huan X and Marzouk Y 2013 *J. Comput. Phys.* **232** 288–317
- [91] Bisetti F, Kim D, Knio O, Long Q and Tempone R 2016 *Int. J. Num. Meth. Eng.* **108** 136–155