Inversion of Robin coefficient by a spectral stochastic finite element approach

Bangti Jin, Jun Zou [*],[1]

Department of Mathematics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

Abstract

This paper investigates a variational approach to the nonlinear stochastic inverse problem of probabilistically calibrating the Robin coefficient from boundary measurements for the steady-state heat conduction. The problem is formulated into an optimization problem, and mathematical properties relevant to its numerical computations are investigated. The spectral stochastic finite element method using polynomial chaos is utilized for the discretization of the optimization problem, and its convergence is analyzed. The nonlinear conjugate gradient method is derived for the optimization system. Numerical results for several two-dimensional problems are presented to illustrate the accuracy and efficiency of the stochastic finite element method.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Robin inverse problem; Spectral stochastic finite element method; Conjugate gradient method; Uncertainty quantification; Stochastic inverse problems

1. Introduction

Inverse problems arise naturally in assorted engineering design, control and identification problems, and inverse theory has witnessed great success during the past few decades. In this paper, we are interested in probabilistically calibrating the heat transfer coefficient, hereafter denoted as the Robin coefficient, of a Robin boundary condition, which models the convection between the conducting body and the ambient environment, from the boundary measurements of the temperature and heat flux.

The values of the Robin coefficient are of immense practical interest in thermal problems, such as the design of gas-turbine blades and nuclear reactors and the analysis of quenching processes [1]. But its accurate values are difficult to obtain experimentally since they depend strongly on at least twelve variables or eight non-dimensional groups [2]. Instead engineers seek to infer it from measured data, which leads naturally to the nonlinear inverse problem of estimating the Robin coefficient from boundary or interior measurements

[3,4]. Several non-destructive evaluation methods also give rise to this inverse problem. In corrosion detection, the Robin coefficient could represent the corrosion damage profile [5,6]; and in the study of MOSFET semiconductor devices, it contains information about the contact resistance and location of the metal-to-silicon contact window [7].

Several numerical methods [6,8–12,14] have been proposed for the Robin inverse problem, in particular in the context of corrosion detection, among which the least-squares method [9–12] has received intensive investigations and it has been implemented in the boundary integral equation method [9,10], the boundary element method [11] and the finite element method [12]. The Robin inverse problem studied in the literature so far is deterministic in the sense that the stochastic nature of the measurement errors is not rigorously considered and analyzed, and these deterministic inverse techniques yield only a single estimate of the large ensemble of solutions that are consistent with the given data without quantifying the uncertainties in the inverse solution.

Uncertainties are ubiquitous due to imperfect data acquisition, modeling errors and solution errors. The data is always contaminated by the inherent measurement errors, and only known with certain confidence, e.g. bounds on the error or the standard deviation of the noise. The forward model may be imprecise due to the presence of unmodeled physics, inherent variability of the physical system and limited ability of building realistic models. This occurs naturally for complex models with significant microstructures, e.g. the permeability of porous media [15], the heat conductivity of composite materials and polymeric materials, thermo-mechanical properties of polycrystalline materials, and domains with rough surfaces. Finally, the model may be so complex that the numerical approximation commits considerable amount of numerical noise [16], which in turn can affect the inverse solution significantly [17]. Uncertainties can be described in several ways, depending on the amount of information available, e.g. evidence theory, fuzzy set theory, worst-case scenario analysis and probabilistic setting (see e.g. [18] and extensive references therein). In this paper, we focus on probabilistic description of uncertainties, where the uncertain input data is modeled as random variables, or more generally, as stochastic processes with given correlation structures, and the resulting mathematical model is often a stochastic partial differential equation [19].

These uncertainties would certainly affect the inverse solution, and rigorous quantification of their effect on the inverse solution would allow enhanced decision-making strategies and deeper insight into the physical problem. However, the numerical analysis of inverse problems under uncertainties remains largely unexplored due to unprecedented computational challenges associated with stochastic numerics. Therefore, the development of computationally efficient and mathematically founded stochastic inverse methods that could account for uncertainties and that are able to provide the inverse solution with its uncertainties quantified is compelling.

Recent efforts on numerical methods for stochastic inverse problems include the Bayesian inference approach [17] and the spectral stochastic approach [20]. The Bayesian inference approach [17] could provide a complete probabilistic description of the inverse solution, and it also alleviates the difficult problem of selecting a suitable regularization parameter via hierarchical Bayesian models. However, it could be computationally expensive for nonlinear inverse problems. The spectral stochastic approach [20] elegantly integrates the powerful and efficient variational method with the state-of-art uncertainty quantification techniques, and it has been applied to reconstructing the heat flux. But no regularization has been considered in the mathematical formulation [20]. Considering the ill-posedness of the inverse problem, regularization should be incorporated in the variational formulation in order to rigorously justify the formulation, its stability and discretization as well as the convergence of the discrete system.

Motivated by the encouraging numerical results reported in [20], we shall propose a spectral stochastic finite element method using polynomial chaos for the stochastic Robin inverse problem, with some regularization incorporated in the variational formulation of the considered inverse problem. We shall also provide a systematic mathematical justification of the algorithm. This seems to be the first work of the kind for a stochastic inverse problem. The outline of the paper is as follows. Section 2 describes the mathematical formulation of the inverse problem and its reformulation as an optimization problem by defining a certain functional. Section 3 investigates mathematical properties of the functional relevant to its numerical computations. Section 4 describes the conjugate gradient method implemented in the spectral stochastic finite element method for the numerical solution of the optimization problem, and analyzes the convergence of the stochastic finite element approximation. Numerical results for several two-dimensional problems with the input data given in terms of polynomial chaos expansion (PCE) coefficients are presented and discussed in

Section 5 to illustrate the efficiency and accuracy of the numerical algorithm. Some existing techniques for constructing PCE coefficients from experimental data are also briefly discussed in Section 5.

We will use standard notations for Sobolev spaces in the sequel. For a non-negative integer s and $1 \leqslant p \leqslant +\infty$, let $W^{s,p}(\Omega)$ be the Sobolev space of functions with generalized derivatives up to order s in the space $L^p(\Omega)$, and its norm be denoted by $\|\cdot\|_{W^{s,p}(\Omega)}$. Whenever $p = 2$, we shall use the notation $H^s(\Omega)$ instead of $W^{s,2}(\Omega)$. Hereafter C denotes a generic constant, which may differ at different occurrences, and it is independent of the discretization parameters, i.e. the finite element mesh size h and the polynomial chaos expansion order p.

2. Mathematical formulation of the problem

2.1. Representation of stochastic processes

In the following, we confine our discussion to the case that the probability space has a finite dimensionality. This is the case for problems depending on a few parameters. Infinite-dimensional stochastic processes can be approximated by a finite number of random variables using the Karhunen–Loève expansion described next.

Let $(\mathcal{D}, \mathcal{E}, \mathcal{P})$ be a probability space, with $\mathcal{D}$ being the sample space, $\mathcal{E}$ the minimal σ-algebra of elementary events and $\mathcal{P}$ the probability measure. We will use $\mathcal{D}$ to denote the triplet $(\mathcal{D}, \mathcal{E}, \mathcal{P})$ for short. Let Y be an $\mathbb{R}^n$-valued random variable in $\mathcal{D}$, i.e. a measurable function from $\mathcal{D}$ to $\mathbb{R}^n$. Moreover, we assume that its distribution is absolutely continuous with respect to the Lebesgue measure. Then its expected value $E[Y]$ is given by

$$E[Y] = \int_{\mathcal{D}} Y(\omega) \mathrm{d}\mathcal{P}(\omega) = \int_{\mathbb{R}^n} y\rho \, \mathrm{d}y,$$

where ρ is the probability density function. The covariance matrix of Y, $\mathrm{Cov}[Y] \in \mathbb{R}^{n \times n}$, is defined by

$$\mathrm{Cov}[Y](i,j) = \mathrm{Cov}(Y_i, Y_j) = E[(Y_i - E[Y_i])(Y_j - E[Y_j])], \quad i, j = 1, \ldots, n.$$

For a stochastic process $w(x, \omega)$ with index $x \in \Omega$, its covariance function $C(x_1, x_2)$ is defined as

$$C(x_1, x_2) = \mathrm{Cov}[w(x_1, \cdot), w(x_2, \cdot)].$$

Then the covariance operator, i.e. the integral operator associated with the covariance function $C(x_1, x_2)$, is compact, real symmetric and semi-positive definite. Therefore it is Hilbert–Schmidt, and its eigenfunctions can be chosen to be mutually orthogonal and form a complete set. The Karhunen–Loève expansion (KLE) [19] employs the eigenpairs of the covariance operator to represent the stochastic process $w(x, \omega)$ by

$$w(x, \omega) = \bar{w}(x) + \sum_{i=1}^{\infty} \sqrt{\lambda_i}\xi_i(\omega)f_i(x), \tag{1}$$

where $\bar{w}(x)$ denotes the mean of the stochastic process and $\{\xi_i(\omega)\}_{i=1}^{\infty}$ form a set of uncorrelated random variables spanning the probability space $\mathcal{D}$. The random variables $\{\xi_i\}_{i=1}^{\infty}$ arising from the KLE are not necessarily independent, but they are often assumed to be independent in the literature. The scalars λ_i and the deterministic functions $f_i(x)$ are the eigenpairs of the covariance operator, i.e.

$$\int_{\Omega} C(x_1, x_2)f_i(x_1) \, \mathrm{d}x_1 = \lambda_i f_i(x_2), \tag{2}$$

and the eigenvalues are non-increasingly ordered

$$\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3 \geqslant \cdots.$$

The KLE is optimal in the mean-square sense among all possible decomposition of a stochastic process [19]. However, its application is limited by the fact that the covariance function needs to be known *a priori*. This can be complemented using the polynomial chaos expansion described in Section 4.2. The infinite sum (1) is not amenable with numerical computations, and a truncated KLE is often employed

$$w(x, \omega) \approx \bar{w}(x) + \sum_{i=1}^{N} \sqrt{\lambda_i}\xi_i(\omega)f_i(x).$$

The number of terms N needed to give an accurate approximation depends on the decay rate of the eigenvalues λ_i, which in turn depends on the regularity of the covariance function $C(x_1, x_2)$. If the eigenvalues λ_i decay rapidly, e.g. in the case of smooth kernels, a small N would suffice. In practice, an *ad hoc* choice of the covariance function, e.g. the exponential distribution, is often made, without or with little physical justification [19]. Constructing the covariance function out of experimental data has been considered by Babuška et al. [21].

2.2. Formulation of the Robin inverse problem

Let Ω be an open and bounded domain in $\mathbb{R}^d$ $(d = 2, 3)$ with a boundary $\Gamma = \partial\Omega$. The boundary Γ consists of two disjointed parts $\Gamma = \Gamma_i \cup \Gamma_c$. The boundaries Γ_i and Γ_c refer to the part of the boundary Γ that is inaccessible and accessible to experimental measurement, respectively. Then the steady-state heat conduction problem under uncertainties could be described by the stochastic elliptic partial differential equation [22,23]

$$-\nabla \cdot (\alpha \nabla u) = f, \quad (x, \omega) \in \Omega \times \mathcal{D}, \tag{3}$$

where the stochastic function $f(\cdot, \omega) \in (H^1(\Omega))'$ for almost every (a.e.) $\omega \in \mathcal{D}$ represents the heat source. Unless otherwise specified, the operator ∇ always takes the gradient with respect to the variable x. The stochastic function $\alpha(x, \omega)$ denotes the heat conductivity, and we assume that there exist positive constants α_0 and α_1 such that $\alpha_0 \leqslant \alpha(x, \omega) \leqslant \alpha_1$. The stochasticity of the functions α and f represents the variability or uncertainty of the heat conductivity and the thermal load, respectively, which can never be exactly measured or applied.

When convective heat transfer occurs through the boundary Γ_i, Newton's law states that the problem is subjected to the Robin boundary condition

$$\alpha \frac{\partial u}{\partial n} + \gamma u = \gamma u_a, \quad (x, \omega) \in \Gamma_i \times \mathcal{D},$$

where $u_a(x)$ is the ambient temperature, and $\gamma(x, \omega)$ is the Robin coefficient.

The inverse problem is to recover the Robin coefficient $\gamma(x, \omega)$ on the boundary Γ_i from the Cauchy data on the accessible part of the boundary Γ_c

$$u = g, \qquad \alpha \frac{\partial u}{\partial n} = q, \quad (x, \omega) \in \Gamma_c \times \mathcal{D},$$

where the stochastic functions $g(\cdot, \omega) \in H^{\frac{1}{2}}(\Gamma_c)$ and $q(\cdot, \omega) \in H^{-\frac{1}{2}}(\Gamma_c)$ for a.e. $\omega \in \mathcal{D}$. In this paper we confine ourselves to the practical case of finite-dimensional noise.

Assumption 2.1. The stochastic functions α, f and $q : \Omega \times \mathcal{D} \mapsto \mathbb{R}$ have the form

$$\alpha(x, \omega) = \alpha(x, Y_1(\omega), \ldots, Y_N(\omega)),$$
$$f(x, \omega) = f(x, Y_1(\omega), \ldots, Y_N(\omega)),$$
$$q(x, \omega) = q(x, Y_1(\omega), \ldots, Y_N(\omega)),$$

where N is a positive integer, $\{Y_n\}_{n=1}^N$ are real-valued independent random variables with zero mean and unit standard deviation. Furthermore, we denote with $I_n = Y_n(\mathcal{D})$ the image of Y_n, with I_n being a bounded interval in $\mathbb{R}$, and each Y_i has a density function $\rho_i : I_i \mapsto \mathbb{R}^+$ for $i = 1, 2, \ldots, N$.

The finite-dimensional noise assumption has been assumed by several authors [19,24,25,18]. It corresponds to the situation where the uncertainty of the problem is inherently associated with a finite number of random variables. It can also be satisfied for situations where the uncertainty derives from infinite-dimensional stochastic processes by dimensionality-reduction techniques such as the Karhunen–Loève expansion or the polynomial chaos expansion. The Doob–Dynkin theorem [25] asserts that the solution $u(x, \omega)$ depends only on $\{Y_i\}_{i=1}^N$, and thus $u(x, \omega) = u(x, Y_1(\omega), \ldots, Y_N(\omega))$. Denote the joint probability density function of the random variables $Y_1(\omega), \ldots, Y_N(\omega)$ by $\rho(y) \equiv \prod_{i=1}^N \rho_i(y_i)$ for $y \in I \equiv \prod_{i=1}^N I_i$. Then the stochastic functions $\alpha(x, \omega), f(x, \omega)$ and $q(x, \omega)$ can be succinctly written as $\alpha(x, \omega) \equiv \alpha(x, y), f(x, \omega) \equiv f(x, y)$ and $q(x, \omega) \equiv q(x, y)$ with $y \in I$. The admissible set $\mathcal{A}$ of the Robin coefficient $\gamma(x, y)$ is taken to be

$$\mathcal{A} = \{\gamma : \gamma \in L^\infty(\Gamma_i \times I) \cap W^1(\Gamma_i \times I), \ c_0 \leqslant \gamma(x,y) \leqslant c_1\},$$

where c_0 and c_1 are positive constants. The weighted Sobolev space $W^1(\Gamma_i \times I)$ is given by

$$W^1(\Gamma_i \times I) = \{\gamma(x,y) \in L^2(\Gamma_i \times I) : \gamma \in L^2(I; H^1(\Gamma_i)), \gamma \in H^1(I; L^2(\Gamma_i))\},$$

and the norm $\| \cdot \|_{W^1(\Gamma_i \times I)}$ is defined as

$$\|\gamma\|^2_{W^1(\Gamma_i \times I)} = \int_I \rho \left(\|\gamma\|^2_{H^1(\Gamma_i)} + \|\nabla_y \gamma\|^2_{L^2(\Gamma_i)} \right) \mathrm{d}y,$$

where ∇_y denotes taking gradient with respect to y.

The inverse problem is ill-posed in the sense that the existence and uniqueness of the solution may be violated, and more importantly, a small perturbation in the data could cause an enormous deviation in the solution. Thus we resort to its variational formulation: minimizing the functional $\mathcal{J}$ over the admissible set $\mathcal{A}$

$$\mathcal{J}(\gamma) = \frac{1}{2} \int_I \rho \int_{\Gamma_c} (u(\gamma) - g)^2 \, \mathrm{d}s \, \mathrm{d}y + \frac{\eta}{2} \|\gamma\|^2_{W^1(\Gamma_i \times I)}. \tag{4}$$

Here the first term represents the mean-squares error between the solution $u(\gamma)$ and the data g. The second term refers to the well-known Tikhonov regularization in the W^1 norm, and the regularization parameter $\eta > 0$ controls the tradeoff between the data fitting and *a priori* information. Note that for notational simplicity we abuse slightly g to denote the measured data, i.e. $g(x,y) \in L^2(\Gamma_c)$ for a.e. $y \in I$. In particular, the functional $\mathcal{J}$ will be denoted by $\mathcal{J}_0$ in the case of $\eta = 0$. The function $u(\gamma)(x,y)$ is the solution to the stochastic boundary value problem

$$\begin{aligned} -\nabla \cdot (\alpha \nabla u) &= f, \quad (x,y) \in \Omega \times I, \\ \alpha \frac{\partial u}{\partial n} &= q, \quad (x,y) \in \Gamma_c \times I, \\ \alpha \frac{\partial u}{\partial n} + \gamma u &= \gamma u_a, \quad (x,y) \in \Gamma_i \times I. \end{aligned} \tag{5}$$

Assumption 2.1 reduces the stochastic forward problem to a parametric yet deterministic problem, and allows the use of the finite element method to approximate its solution. In the next subsection we collect some basic properties of the forward problem.

2.3. Well-posedness of the forward problem

Since stochastic functions have intrinsically different structure with respect to x and with respect to y, the analysis of numerical approximations requires tensor spaces. In particular, we will need the Hilbert space $L^2(I; H^1(\Omega))$ for discussing the well-posedness of the forward problem. The Hilbert space $L^2(I; H^1(\Omega))$ is defined by

$$L^2(I; H^1(\Omega)) = \{w : \text{a measurable function from } \Omega \times I \text{ to } \mathbb{R} \text{ and } E\left[\|w(\cdot,y)\|^2_{H^1(\Omega)}\right] < +\infty\},$$

and its norm $\| \cdot \|_{L^2(I; H^1(\Omega))}$ is given by

$$\|w\|^2_{L^2(I; H^1(\Omega))} = E\left[\|w(\cdot,y)\|^2_{H^1(\Omega)}\right] \equiv \int_I \rho \|w(\cdot,y)\|^2_{H^1(\Omega)} \, \mathrm{d}y.$$

Here we have omitted the weight ρ in the norm notation, and hereafter we assume this convention for the L^2 and H^1 norms involving I. Note that the space has an obvious probabilistic interpretation as expectation, but we will use the integral notation $\int_I \rho \cdot \mathrm{d}y$ consistently in place of the expectation operator $E[\cdot]$ in the sequel for its simple variational justification.

With these notations, the forward problem (5) could be recast into the variational formulation

$$\mathcal{B}(u, \phi) = \mathcal{L}(\phi) \quad \forall \phi \in L^2(I; H^1(\Omega)), \tag{6}$$

where the bilinear form $\mathcal{B} : L^2(I; H^1(\Omega)) \times L^2(I; H^1(\Omega)) \mapsto \mathbb{R}$ and the linear functional $\mathcal{L} : L^2(I; H^1(\Omega)) \mapsto \mathbb{R}$ are respectively defined by

$$\mathcal{B}(u, \phi) \equiv \int_I \rho \int_\Omega \alpha \nabla u \cdot \nabla \phi \, dx \, dy + \int_I \rho \int_{\Gamma_i} \gamma u \phi \, ds \, dy \quad \forall u, \phi \in L^2(I; H^1(\Omega)),$$

$$\mathcal{L}(\phi) \equiv \int_I \rho \int_\Omega f \phi \, dx \, dy + \int_I \rho \int_{\Gamma_c} q \phi \, ds \, dy + \int_I \rho \int_{\Gamma_i} \gamma u_a \phi \, ds \, dy \quad \forall \phi \in L^2(I; H^1(\Omega)).$$

One can verify that the variational formulation (6) has a unique weak solution $u \in L^2(I; H^1(\Omega))$, and the solution is bounded independent of γ; see [13] for details.

3. Properties of the functional $\mathcal{J}$

This section investigates mathematical properties of the functional $\mathcal{J}$. We need the following lemma for establishing the existence result.

Lemma 3.1. *Assume that $\{\gamma^n\} \subset \mathcal{A}$ and $\gamma^* \in \mathcal{A}$. Then $\lim_{n \to \infty} \gamma^n = \gamma^*$ in $L^1(\Gamma_i \times I)$ implies that there exists a subsequence of $u(\gamma^n)$, also denoted as $u(\gamma^n)$, such that*

$$\lim_{n \mapsto \infty} u(\gamma^n) = u(\gamma^*) \quad \text{in } L^2(I; H^1(\Omega)).$$

Proof. Let $w^n = u(\gamma^n) - u(\gamma^*)$. Then the definition of $u^n \equiv u(\gamma^n)$ and $u(\gamma^*)$ gives

$$\int_I \rho \int_\Omega \alpha \nabla w^n \cdot \nabla \phi \, dx \, dy + \int_I \rho \int_{\Gamma_i} \gamma^n w^n \phi \, ds \, dy = \int_I \rho \int_{\Gamma_i} (\gamma^n - \gamma^*)(u_a - u(\gamma^*)) \phi \, ds \, dy.$$

Taking $\phi = w^n$, we obtain

$$\alpha_0 \int_I \rho \int_\Omega |\nabla w^n|^2 \, dx \, dy + c_0 \int_I \rho \int_{\Gamma_i} (w^n)^2 \, ds \, dy \leqslant \int_I \rho \int_{\Gamma_i} (\gamma^n - \gamma^*)(u_a - u(\gamma^*)) w^n \, ds \, dy.$$

By the Cauchy–Schwarz inequality and the Young's inequality, we derive

$$\int_I \rho \int_{\Gamma_i} (\gamma^n - \gamma^*)(u_a - u(\gamma^*))(u^n - u(\gamma^*)) \, ds \, dy$$
$$\leqslant \|w^n\|_{L^2(I;L^2(\Gamma_i))} \|(\gamma^n - \gamma^*)(u_a - u(\gamma^*))\|_{L^2(I;L^2(\Gamma_i))}$$
$$\leqslant \frac{c_0}{2} \|w^n\|_{L^2(I;L^2(\Gamma_i))}^2 + \frac{1}{2c_0} \|(\gamma^n - \gamma^*)(u_a - u(\gamma^*))\|_{L^2(I;L^2(\Gamma_i))}^2.$$

Therefore,

$$\alpha_0 \int_I \rho \int_\Omega |\nabla w^n|^2 \, dx \, dy + \frac{c_0}{2} \int_I \rho \int_{\Gamma_i} (w^n)^2 \, ds \, dy \leqslant \frac{1}{2c_0} \|(\gamma^n - \gamma^*)(u_a - u(\gamma^*))\|_{L^2(I;L^2(\Gamma_i))}^2.$$

Recall that the $L^1(\Gamma_i \times I)$ convergence of the sequence γ^n to γ^* implies that there exists a subsequence of γ^n converges almost everywhere [26]. By passing to a subsequence, the term on the right hand side tends to 0 as n tends to infinity using Lebesgue dominated convergence theorem [26]. This concludes the proof of the lemma. $\square$

Assisted with Lemma 3.1, one can justify the existence of a minimizer to the optimization problem (4) and the stability of the formulation; see [13] for the detailed proof.

We attempt to solve the optimization problem (4) by gradient type methods, and thus the gradient of the functional $\mathcal{J}$ with respect to the Robin coefficient γ is of much interest. We first establish the differentiability of the solution $u(\gamma)$ with respect to γ. To do so, we define the sensitivity problem for $u^1 \equiv u^1(\gamma, \lambda)$ for any $\gamma \in \mathcal{A}$ and $\lambda \in L^\infty(\Gamma_i \times I)$ by

$$-\nabla \cdot (\alpha \nabla u^1) = 0, \quad (x, y) \in \Omega \times I,$$

$$\alpha \frac{\partial u^1}{\partial n} = 0, \quad (x, y) \in \Gamma_c \times I,$$

$$\alpha \frac{\partial u^1}{\partial n} + \gamma u^1 = \lambda(u_a - u(\gamma)), \quad (x, y) \in \Gamma_i \times I.$$

(7)

The sensitivity problem will also be utilized in the nonlinear conjugate gradient method for determining an appropriate step length.

Lemma 3.2. *For any $\gamma \in \mathcal{A}$, the solution $u(\gamma)$ is differential in the sense that*

$$\frac{\|u(\gamma + \lambda) - u(\gamma) - u^1(\gamma, \lambda)\|_{L^2(I; H^1(\Omega))}}{\|\lambda\|_{L^\infty(\Gamma_i \times I)}} \to 0 \quad as \ \lambda \to 0 \quad in \ L^\infty(\Gamma_i \times I).$$

Proof. The function $w \equiv u(\gamma + \lambda) - u(\gamma) - u^1(\gamma, \lambda)$ satisfies

$$-\nabla \cdot (\alpha \nabla w) = 0, \quad (x, y) \in \Omega \times I,$$

$$\alpha \frac{\partial w}{\partial n} = 0, \quad (x, y) \in \Gamma_c \times I,$$

$$\alpha \frac{\partial w}{\partial n} + \gamma w = -\lambda(u(\gamma + \lambda) - u(\gamma)), \quad (x, y) \in \Gamma_i \times I.$$

Therefore we have

$$\int_I \rho \int_\Omega \alpha |\nabla w|^2 \, dx \, dy + \int_I \rho \int_{\Gamma_i} \gamma w^2 \, ds \, dy = -\int_I \rho \int_{\Gamma_i} \lambda(u(\gamma + \lambda) - u(\gamma)) w \, ds \, dy.$$

The Cauchy–Schwarz inequality and the Sobolev trace lemma imply

$$-\int_I \rho \int_{\Gamma_i} \lambda(u(\gamma + \lambda) - u(\gamma)) w \, ds \, dy \leqslant \|\lambda\|_{L^\infty(\Gamma_i \times I)} \|u(\gamma + \lambda) - u(\gamma)\|_{L^2(I; L^2(\Gamma_i))} \|w\|_{L^2(I; L^2(\Gamma_i))}$$

$$\leqslant C \|\lambda\|_{L^\infty(\Gamma_i \times I)} \|u(\gamma + \lambda) - u(\gamma)\|_{L^2(I; H^1(\Omega))} \|w\|_{L^2(I; H^1(\Omega))}.$$

Therefore, we derive that

$$\|w\|_{L^2(I; H^1(\Omega))} \leqslant C \|\lambda\|_{L^\infty(\Gamma_i \times I)} \|u(\gamma + \lambda) - u(\gamma)\|_{L^2(I; H^1(\Omega))}.$$

Similarly we can show that

$$\|u(\gamma + \lambda) - u(\gamma)\|_{L^2(I; H^1(\Omega))} \leqslant C \|\lambda\|_{L^\infty(\Gamma_i \times I)} (\|u_a\|_{L^2(I; L^2(\Gamma_i))} + \|u(\gamma + \lambda)\|_{L^2(I; H^1(\Omega))}).$$

However, for $\|\lambda\|_{L^\infty(\Gamma_i \times I)}$ sufficiently small $\|u(\gamma + \lambda)\|_{L^2(I; H^1(\Omega))}$ can be uniformly bounded independent of λ. Thus it follows directly

$$\frac{\|u(\gamma + \lambda) - u(\gamma) - u^1(\gamma, \lambda)\|_{L^2(I; H^1(\Omega))}}{\|\lambda\|_{L^\infty(\Gamma_i \times I)}} \to 0 \quad as \ \lambda \to 0 \quad in \ L^\infty(\Gamma_i \times I). \quad \square$$

From the proof of Lemma 3.2, we also have the following expansion

$$u(\gamma + \lambda) = u(\gamma) + u^1(\gamma, \lambda) + \mathcal{O}(\|\lambda\|_{L^\infty(\Gamma_i \times I)}^2).$$

For the efficient evaluation of the gradient, we introduce the adjoint equation associated with $u(\gamma)$: Find $v(\gamma) \in L^2(I; H^1(\Omega))$ such that

$$-\nabla \cdot (\alpha \nabla v) = 0, \quad (x, y) \in \Omega \times I,$$

$$\alpha \frac{\partial v}{\partial n} = u(\gamma) - g, \quad (x, y) \in \Gamma_c \times I,$$

$$\alpha \frac{\partial v}{\partial n} + \gamma v = 0, \quad (x, y) \in \Gamma_i \times I.$$

(8)

Theorem 3.1. *The functional $\mathcal{J}$ is Fréchet differentiable, and its Fréchet derivative $\mathcal{J}'(\gamma)$ at $\gamma \in \mathcal{A}$ is given by*

$$\mathcal{J}'(\gamma)[\lambda] = \int_I \rho \int_{\Gamma_i} \lambda(u_a - u(\gamma))v(\gamma) + \eta(\lambda\gamma + \nabla\lambda \cdot \nabla\gamma + \nabla_y\lambda \cdot \nabla_y\gamma) \, ds \, dy. \tag{9}$$

Proof. By Lemma 3.2 we observe

$$
\begin{aligned}
\mathcal{J}_0(\gamma + \lambda) - \mathcal{J}_0(\gamma) &= \frac{1}{2} \int_I \rho \int_{\Gamma_c} (u(\gamma + \lambda) - g)^2 \, ds \, dy - \frac{1}{2} \int_I \rho \int_{\Gamma_c} (u(\gamma) - g)^2 \, ds \, dy \\
&= \frac{1}{2} \int_I \rho \int_{\Gamma_c} \left\{ (u(\gamma) + u^1(\gamma, \lambda) + \mathcal{O}(\|\lambda\|^2_{L^\infty(\Gamma_i \times I)}) - g)^2 - (u(\gamma) - g)^2 \right\} ds \, dy \\
&= \int_I \rho \int_{\Gamma_c} (u(\gamma) - g)u^1(\gamma, \lambda) \, ds \, dy + \mathcal{O}(\|\lambda\|^2_{L^\infty(\Gamma_i \times I)}),
\end{aligned}
$$

and consequently

$$\mathcal{J}_0'(\gamma)[\lambda] = \int_I \rho \int_{\Gamma_c} (u(\gamma) - g)u^1(\gamma, \lambda) \, ds \, dy.$$

Now the Green's second identity applied to $u^1(\gamma, \lambda)$ and $v(\gamma)$ reads

$$
\begin{aligned}
0 &= \int_I \rho \int_\Omega \{v(\gamma)\nabla \cdot (\alpha\nabla u^1(\gamma, \lambda)) - u^1(\gamma, \lambda)\nabla \cdot (\alpha\nabla v(\gamma))\} \, dx \, dy \\
&= \int_I \rho \int_\Gamma \left\{ \alpha\frac{\partial u^1(\gamma, \lambda)}{\partial n}v(\gamma) - \alpha\frac{\partial v(\gamma)}{\partial n}u^1(\gamma, \lambda) \right\} ds \, dy.
\end{aligned}
$$

Substituting the boundary conditions for $v(\gamma)$ and $u^1(\gamma, \lambda)$ we obtain

$$\int_I \rho \int_{\Gamma_c} (u(\gamma) - g)u^1(\gamma, \lambda) \, ds \, dy = \int_I \rho \int_{\Gamma_i} \lambda(u_a - u(\gamma))v(\gamma) \, ds \, dy.$$

Therefore

$$\mathcal{J}_0'(\gamma)[\lambda] = \int_I \rho \int_{\Gamma_i} \lambda(u_a - u(\gamma))v(\gamma) \, ds \, dy,$$

and consequently

$$\mathcal{J}'(\gamma)[\lambda] = \int_I \rho \int_{\Gamma_i} \lambda(u_a - u(\gamma))v(\gamma) + \eta(\lambda\gamma + \nabla\lambda \cdot \nabla\gamma + \nabla_y\lambda \cdot \nabla_y\gamma) \, ds \, dy. \quad \square$$

4. Numerical algorithm

This section describes the conjugate gradient method with its spectral stochastic finite element implementation for the numerical solution of the optimization problem, and analyzes the convergence of the stochastic finite element approximation.

4.1. Conjugate gradient method

The ill-posed nature of the inverse problem carries over to the optimization problem, and thus it suffers also from numerical instability. Regularization techniques are among the most popular methods for stabilizing ill-posed problems. The Tikhonov regularization method stabilizes the functional by penalizing large solution norms. An alternative attracting much recent attentions is iterative regularization methods, e.g. the conjugate gradient method (CGM) and the generalized minimal residual method. The CGM in conjunction with an appropriate stopping rule can serve as a regularization method, and it has been successfully applied to various inverse problems.

Since an expression for the gradient $\mathcal{J}'(\gamma)$ of the functional $\mathcal{J}$ is explicitly available, and easily obtained by solving the adjoint problem, the CGM can be readily implemented. The Fletcher-Reeves conjugate gradient algorithm [27] applied to the optimization problem takes the form

(a) Choose γ^0, and set $k = 0$;
(b) Solve the direct problem (5) with $\gamma = \gamma^k$, and determine the residual

$$\tilde{r}_k = u(\gamma^k)|_{\Gamma_c \times I} - g;$$

(c) Solve the adjoint problem (8) for $v(\gamma^k)$ with the Neumann boundary condition given by

$$\alpha \frac{\partial v(\gamma^k)}{\partial n} = \tilde{r}_k, \quad (x, y) \in \Gamma_c \times I.$$

Determine the gradient $\mathcal{J}'(\gamma^k)$ by

$$\mathcal{J}'(\gamma^k) = v(\gamma^k)(u_a - u(\gamma^k)) + \eta(\gamma^k - \Delta\gamma^k - \Delta_y\gamma^k),$$

then calculate the descent direction

$$d_k = -\mathcal{J}'(\gamma^k) + \beta_{k-1}d_{k-1}$$

with the conjugate coefficient β_k given by

$$\beta_{k-1} = \frac{\|\mathcal{J}'(\gamma^k)\|^2_{L^2(I;L^2(\Gamma_i))}}{\|\mathcal{J}'(\gamma^{k-1})\|^2_{L^2(I;L^2(\Gamma_i))}};$$

and assuming the convention that $\beta_{-1} = 0$.
(d) Solve the sensitivity problem (7) for $u^1(\gamma, \lambda)$ with $\lambda = d_k$;
(e) Calculate the step length α_k in the conjugate gradient direction d_k by

$$\alpha_k = -\frac{\langle \tilde{r}_k, u^1(\gamma, \lambda)\rangle_{L^2(I;L^2(\Gamma_c))} + \eta\langle\gamma^k, d_k\rangle_{W^1(\Gamma_i \times I)}}{\|u^1(\gamma, \lambda)\|^2_{L^2(I;L^2(\Gamma_c))} + \eta\|d_k\|^2_{W^1(\Gamma_i \times I)}};$$

(f) Update the Robin coefficient γ^k by

$$\gamma^{k+1} = \gamma^k + \alpha_k d_k;$$

(g) Increase k by one and go to Step (b), repeat the above procedure until a stopping criterion is satisfied.

The standard CGM typically employs a line search procedure to determine the step length α_k, which requires evaluating the functional multiple times per iteration. This is undesirable in the present context since each functional evaluation requires one forward solver. Instead, here the step length is determined by a quadratic approximation of the functional using the sensitivity problem. For all numerical examples considered in the present study and previous works (see e.g. [27]) with inverse problems, this step size calculation has proved to be effective.

4.2. Spectral stochastic finite element method

We employ the spectral stochastic finite element method (SSFEM) [19] to discretize the optimization problem and the forward problems arising from the CGM. The SSFEM originated from the pioneering works of Ghanem and Spanos [19], and since then it has witnessed considerable success in various engineering problems, e.g. solid mechanics, heat transfer and fluid flow. Its horizon has also been vastly extended in various aspects, e.g. the Wiener–Askey scheme for non-Gaussian random variables [24] and sparse grid-based stochastic collocation methods [28] and model reduction [29] for mitigating the curse of dimensionality. For the numerical analysis of the SSFEM for solving elliptic problems, we refer to the recent works [25,18].

To formulate the SSFEM approximation, we first triangulate the spatial domain Ω with a shape regular and quasi-uniform triangulation $\mathcal{T}_h$ of simplicial elements. Then we define the piecewise linear finite element space V_h by

$$V_h = \{\psi_h \in C^0(\Omega) : \psi_h|_{\mathcal{T}_i} \in P_1(\mathcal{T}_i), \forall \mathcal{T}_i \in \mathcal{T}_h\},$$

where $P_1(\mathcal{T}_i)$ denotes the space of linear polynomials on the finite element $\mathcal{T}_i$. Moreover, we define the space $V_{\Gamma_i,h}$ to be the restriction of V_h on Γ_i.

The SSFEM employs the polynomial chaos, i.e. orthogonal polynomials in multidimensional random variables with respect to the probability density ρ, based on Wiener's pioneering works on homogeneous chaos [30]. Consider a set of independent, identically distributed random variables $\{\xi_i(\omega)\}_{i=1}^\infty$. The polynomial chaos expansion (PCE) [19] represents a stochastic process $w(x,\omega)$ by

$$w(x,\omega) = a_0(x)H_0 + \sum_{i_1=1}^\infty a_{i_1}(x)H_1(\xi_{i_1}(\omega)) + \sum_{i_1=1}^\infty \sum_{i_2=1}^{i_1} a_{i_1 i_2}(x)H_2(\xi_{i_1}(\omega), \xi_{i_2}(\omega)) + \cdots,$$

where $H_p(\cdot)$, $p = 0, 1, 2, \ldots$ are orthogonal polynomials (known as generalized polynomial chaos). For computational convenience, the PCE is succinctly rewritten as

$$w(x,\omega) = \sum_{i=0}^\infty w_i(x)\varphi_i(\omega),$$

where there is a one-to-one correspondence between the functionals $H_p(\cdot)$ and $\varphi_i(\omega)$. Here $w_0(x)$ is the mean of the stochastic process, and $w_i(x)(i \geqslant 1)$ capture its probabilistic characteristics [19,24]. Cameron and Martin [31] established that such a choice of basis leads to a spectral expansion that converges in L^2 sense to any random variable with finite variance for the Wiener–Hermite chaos. Invoking Assumption 2.1 and truncating the order of the orthogonal polynomials to p yield

$$w(x,y) = \sum_{i=0}^M w_i(x)\varphi_i(y),$$

where $M + 1 = \frac{(N+p)!}{N!p!}$. The polynomial chaos is normalized such that

$$E[\varphi_i \varphi_j] \equiv \int_I \rho\varphi_i(y)\varphi_j(y)dy = \delta_{ij}, \quad i,j = 0,1,\ldots,N,$$

where δ_{ij} is the Kronecker delta. We denote the space $\text{span}\{\varphi_i, i = 0,1,\ldots,M\}$ by W_p.

In the SSFEM, the Robin coefficient $\gamma(x,y)$ and the solution $u(x,y)$ are approximated by

$$\gamma(x,y) \approx \sum_{i=0}^M \gamma_i(x)\varphi_i(y), \quad \text{and} \quad u(x,y) \approx \sum_{i=0}^M u_i(x)\varphi_i(y), \tag{10}$$

respectively, where $\gamma_i(x)$ and $u_i(x)$ are referred to as the ith PCE coefficient for the stochastic processes $\gamma(x,y)$ and $u(x,y)$, respectively. The SSFEM amounts to approximating the solution $u(x,y)$ in the tensor product space

$$W_p \otimes V_h \equiv \{\phi = \phi(x,y) \in L^2(I; H^1(\Omega)) : \phi \in \text{span}\{\psi(x)\varphi(y) : \psi \in V_h, \varphi \in W_p\}\},$$

followed by Galerkin-type procedures or collocation-type methods. Similarly, we can define the tensor product space $W_p \otimes V_{h,\Gamma_i}$. Then the discrete admissible set $\mathcal{A}_{h,p}$ is given by

$$\mathcal{A}_{h,p} = \{\gamma_{h,p} \in W_p \otimes V_{h,\Gamma_i} : c_0 \leqslant \gamma_{h,p} \leqslant c_1\}.$$

Now the discrete optimization problem can be formulated as:

Minimizing the following functional $\mathcal{J}_{h,p}$ over the discrete admissible set $\mathcal{A}_{h,p}$

$$\mathcal{J}_{h,p}(\gamma_{h,p}) = \frac{1}{2}\int_I \rho \int_{\Gamma_c} (u_{h,p}(\gamma_{h,p}) - g)^2 \, \mathrm{d}s \, \mathrm{d}y + \frac{\eta}{2}\|\gamma_{h,p}\|_{W^1(\Gamma_i \times I)}^2, \tag{11}$$

where the function $u_{h,p} \in W_p \otimes V_h$ is the SSFEM solution to the stochastic variational formulation (6):

$$\int_I \rho \int_\Omega \alpha \nabla u_{h,p}(\gamma_{h,p}) \cdot \nabla \phi_{h,p} \, \mathrm{d}x \, \mathrm{d}y + \int_I \rho \int_{\Gamma_i} \gamma_{h,p} u_{h,p}(\gamma_{h,p}) \phi_{h,p} \, \mathrm{d}s \, \mathrm{d}y$$

$$= \int_I \rho \int_\Omega f\phi_{h,p} \, \mathrm{d}x \, \mathrm{d}y + \int_I \rho \int_{\Gamma_c} q\phi_{h,p} \, \mathrm{d}s \, \mathrm{d}y + \int_I \rho \int_{\Gamma_i} \gamma_{h,p} u_a \phi_{h,p} \, \mathrm{d}s \, \mathrm{d}y, \quad \forall \phi_{h,p} \in W_p \otimes V_h. \tag{12}$$

This gives a finite-dimensional system, which is solved via a direct solver in the present study. In practice, an expression of the stochastic process $\gamma(x, y)$ in the form (10) constitutes a full probabilistic description of the inverse solution. Informative information in terms of relevant statistics, e.g. the mean and the variance of the stochastic process, can be directly calculated from the PCE. For example, the mean $E[\gamma(x, y)]$ and the variance $\text{var}[\gamma(x, y)]$ of the stochastic process γ are respectively given by

$$E[\gamma(x, y)] = \sum_{i=0}^{M} \gamma_i(x) E[\varphi_i] = \sum_{i=0}^{M} \gamma_i(x) \delta_{i0} = \gamma_0(x),$$

and

$$\text{var}[\gamma(x, y)] = E[(\gamma(x, y))^2] - E[\gamma(x, y)]^2 = \sum_{i=0}^{M} \sum_{j=0}^{M} \gamma_i(x) \gamma_j(x) E[\varphi_i(y) \varphi_j(y)] - \gamma_0^2(x) = \sum_{i=1}^{M} \gamma_i^2(x).$$

Note again that the mean is simply the zeroth PCE coefficient. Formulas for characteristic statistical quantities related to higher order moments, e.g. skewness and kurtosis, can be derived analogously.

4.3. Convergence analysis

This part analyzes the convergence of the SSFEM approximation in the case that the probability density ρ has a positive lower bound and a finite upper bound. Then the weighted spaces $L^2(I; L^2(\Gamma_i))$ and $W^1(\Gamma_i \times I)$ are isomorphic to that defined by the uniform distribution, i.e. the standard Sobolev space $L^2(I; L^2(\Gamma_i))$ and $H^1(\Gamma_i \times I)$ [32]. We will need the projection operator $\mathcal{R}_{h,p} : L^2(I; H^1(\Omega)) \to W_p \otimes V_h$ and the interpolation operator $\mathcal{I}_{h,p} : C^\infty(\Omega \times I) \to W_p \otimes V_h$ respectively defined by

$$\int_I \rho \int_\Omega \nabla(\mathcal{R}_{h,p} w - w) \cdot \nabla \phi_{h,p} \, dx \, dy + \int_I \rho \int_{\Gamma_i} (\mathcal{R}_{h,p} w - w) \phi_{h,p} \, ds \, dy = 0 \quad \forall \phi_{h,p} \in W_p \otimes V_h,$$

$$\mathcal{I}_{h,p} w = \sum_i \sum_p \langle w(x_i, y), \varphi_p(y) \rangle \psi_i(x) \varphi_p(y) \quad \forall w \in C^\infty(\Omega \times I).$$

The operators $\mathcal{R}_{h,p}$ and $\mathcal{I}_{h,p}$ have the following approximation properties; see [13] for details.

Lemma 4.1. *The operators $\mathcal{R}_{h,p}$ and $\mathcal{I}_{h,p}$ defined above satisfy the following error estimates*

$$\lim_{h \to 0, p \to \infty} \|\mathcal{R}_{h,p} w - w\|_{L^2(I; H^1(\Omega))} = 0 \quad \forall w \in L^2(I; H^1(\Omega)), \tag{13}$$

$$\lim_{h \to 0, p \to \infty} \|\mathcal{I}_{h,p} w - w\|_{L^\infty(\Omega \times I)} = 0 \quad \forall w \in C^\infty(\Omega \times I), \tag{14}$$

$$\lim_{h \to 0, p \to \infty} \|\mathcal{I}_{h,p} w - w\|_{W^1(\Omega \times I)} = 0 \quad \forall w \in C^\infty(\Omega \times I). \tag{15}$$

Assisted with Lemma 4.1, one can verify the existence of a minimizer to the discrete optimization problem (11) and the following convergence property, see [13] for details.

Lemma 4.2. *Assume $\{\gamma_{h,p}\} \subset \mathcal{A}_{h,p}$, and $\gamma^* \in \mathcal{A}$. Then $\lim_{h \to 0, p \to \infty} \gamma_{h,p} = \gamma^*$ in $L^1(I \times \Gamma_i)$ implies that there exists a subsequence of $u_{h,p}(\gamma_{h,p})$, also denoted as $u_{h,p}(\gamma_{h,p})$, such that*

$$u_{h,p}(\gamma_{h,p}) \to u(\gamma^*) \text{ in } L^2(I; H^1(\Omega)),$$

as h tends to 0 and p tends to ∞.

Now we are ready to demonstrate the convergence of the stochastic finite element solution to a minimizer of the continuous optimization problem.

Theorem 4.1. *Assume that the exact solution $\gamma(x, y)$ to the Robin inverse problem (4) satisfies*

$$c_0 + \delta < \gamma(x, y) < c_1 - \delta, \quad \text{a.e. } (x, y) \in \Gamma_i \times I$$

for some small number $\delta > 0$. Let $\gamma_{h,p}^*$ be a sequence of minimizers to the discrete optimization problem (11). *Then each subsequence of $\{\gamma_{h,p}^*\}_{h,p}$ has a subsequence converging weakly in $W^1(\Gamma_i \times I)$ to a minimizer of the continuous optimization problem* (4).

Proof. By taking $\gamma = c_0$, we see that

$$\mathcal{J}_{h,p}(\gamma_{h,p}^*) \leqslant \mathcal{J}_{h,p}(c_0) \leqslant \mathcal{J}(c_0) + C$$

for some constant C. Therefore, the sequence $\gamma_{h,p}^*$ in $W^1(\Gamma_i \times I)$ is uniformly bounded. This implies that there exists a subsequence, also denoted as $\gamma_{h,p}^*$, and some γ^* such that

$$\gamma_{h,p}^* \to \gamma^* \quad \text{weakly in } W^1(\Gamma_i \times I) \quad \text{as } h \to 0 \text{ and } p \to \infty,$$

and, moreover,

$$\gamma_{h,p}^* \to \gamma^* \text{ in } \quad L^2(I; L^2(\Gamma_i)) \quad \text{as } h \to 0 \text{ and } p \to \infty.$$

Now Lemma 4.2 implies that (possibly after passing to a subsequence)

$$u_{h,p}(\gamma_{h,p}^*) \to u(\gamma^*) \quad \text{in } L^2(I; L^2(\Gamma)).$$

Now for any $\gamma \in \mathcal{A}$ and $\Gamma_i \times I \subset \tilde{\Omega}$, we can find an $H^1(\tilde{\Omega})$ function with its maximum norm preserved (e.g. by solving the Laplace equation). Thus for fixed $\varepsilon > 0$, the density of $C^\infty(\Gamma_i \times I)$ in $W^1(\Gamma_i \times I)$ implies that there exists $\gamma^\varepsilon \in C^\infty(\Gamma_i \times I) \cap \mathcal{A}$ such that

$$\gamma^\varepsilon \to \gamma \quad \text{in } W^1(\Gamma_i \times I).$$

Denote $\gamma_{h,p}^\varepsilon = \mathcal{I}_{h,p}\gamma^\varepsilon$, then Lemma 4.1 ensures that $\gamma_{h,p}^\varepsilon \in \mathcal{A}_{h,p}$ for sufficiently small h and sufficiently large p. Furthermore, the property of the interpolation operator $\mathcal{I}_{h,p}$ implies

$$\lim_{h \to 0, p \to \infty} \|\gamma^\varepsilon - \gamma_{h,p}^\varepsilon\|_{W^1(\Gamma_i \times I)} = 0$$

$$\lim_{h \to 0, p \to \infty} \|u_{h,p}(\gamma_{h,p}^\varepsilon) - u(\gamma^\varepsilon)\|_{L^2(I; L^2(\Gamma))} = 0.$$

Noting that $\gamma_{h,p}^*$ is the minimizer of $\mathcal{J}_{h,p}(\cdot)$ over $\mathcal{A}_{h,p}$, we derive

$$\mathcal{J}_{h,p}(\gamma_{h,p}^*) = \frac{1}{2} \int_I \rho \int_{\Gamma_c} (u_{h,p}(\gamma_{h,p}^*) - g)^2 \, ds \, dy + \frac{\eta}{2} \|\gamma_{h,p}^*\|_{W^1(\Gamma_i \times I)}^2$$

$$\leqslant \frac{1}{2} \int_I \rho \int_{\Gamma_c} (u_{h,p}(\gamma_{h,p}^\varepsilon) - g)^2 \, ds \, dy + \frac{\eta}{2} \|\gamma_{h,p}^\varepsilon\|_{W^1(\Gamma_i \times I)}^2 = \mathcal{J}_{h,p}(\gamma_{h,p}^\varepsilon) = \mathcal{J}_{h,p}(\mathcal{I}_{h,p}\gamma^\varepsilon).$$

Then the use of Lemma 4.2 and the convergence property of the interpolation operator $\mathcal{I}_{h,p}$ give

$$\mathcal{J}(\gamma^*) = \frac{1}{2} \int_I \rho \int_{\Gamma_c} (u(\gamma^*) - g)^2 \, ds \, dy + \frac{\eta}{2} \|\gamma^*\|_{W^1(\Gamma_i \times I)}^2$$

$$\leqslant \lim_{h \to 0, p \to \infty} \frac{1}{2} \int_I \rho \int_{\Gamma_c} (u_{h,p}(\gamma_{h,p}^*) - g)^2 \, ds \, dy + \lim_{h \to 0, p \to \infty} \inf \frac{\eta}{2} \|\gamma_{h,p}^*\|_{W^1(\Gamma_i \times I)}^2 \leqslant \lim_{h \to 0, p \to \infty} \inf \mathcal{J}_{h,p}(\gamma_{h,p}^*)$$

$$\leqslant \lim_{h \to 0, p \to \infty} \inf \mathcal{J}_{h,p}(\mathcal{I}_{h,p}\gamma^\varepsilon) \leqslant \frac{1}{2} \int_I \rho \int_{\Gamma_c} (u(\gamma^\varepsilon) - g)^2 \, ds \, dy + \frac{\eta}{2} \|\gamma^\varepsilon\|_{W^1(\Gamma_i \times I)}^2.$$

The $W^1(\Gamma_i \times I)$ convergence of the sequence γ^ε to γ, Lemma 3.1 and the Sobolev embedding theorem imply that there exists a subsequence of $u(\gamma^\varepsilon)$, also denoted as $u(\gamma^\varepsilon)$, such that

$$u(\gamma^\varepsilon) \to u(\gamma) \quad \text{in } L^2(I; H^1(\Omega)),$$

$$u(\gamma^\varepsilon) \to u(\gamma) \quad \text{in } L^2(I; L^2(\Gamma_c)).$$

Now letting ε tend to 0, we obtain

$$\mathcal{J}(\gamma^*) \leqslant \mathcal{J}(\gamma) \ \forall \gamma \in \mathcal{A},$$

which indicates that γ^* is a minimizer of the functional $\mathcal{J}$. $\quad \square$

5. Numerical results and discussions

In this section we illustrate the efficiency and accuracy of the numerical algorithm, namely the CGM in conjunction with the SSFEM, by several two-dimensional numerical examples. The convergence and stability of the algorithm for both exact and noisy data are investigated. The following componentwise notations will be used in this section: $x = (x_1, x_2)$, and $y = (y_1, \ldots, y_N)$.

The domain Ω under consideration is a unit square $(0,1) \times (0,1)$, and the boundaries Γ_i and Γ_c are taken to be $\Gamma_i = (0,1) \times \{1\}$ and $\Gamma_c = \{0\} \times [0,1] \cup [0,1] \times \{0\} \cup \{1\} \times [0,1]$, respectively. Unless otherwise specified, the forward problem is discretized using 1800 uniform triangular finite elements, the order p of the PCE is taken to be 4, and the regularization parameter η is taken to be 1×10^{-6}. The variable y in the conductivity $\alpha(x, y)$ and the Robin coefficient $\gamma(x, y)$ is taken to be a random variable uniformly distributed in $[-1, 1]$, in order to ensure the well-posedness of the forward problem and the SSFEM formulation, and the corresponding generalized polynomial chaos is Legendre polynomials [24].

Since the analytical solution to the stochastic forward problem is unknown, the exact data is calculated by solving the stochastic boundary value problem. The forward solver for generating the data is performed on a finer spatial mesh and with a higher-order PCE, in order to avoid the most obvious form of the notorious 'inverse crime'. In real inverse problems, the boundary data is experimentally measured and thus inevitably contaminated by measurement errors. In our test cases, the simulated noisy data, which retains only the zeroth and first order PCE terms, is generated using the following formula:

$$\tilde{g}_i(x) = \begin{cases} g_i(x) + \max_{x \in \Gamma_i} |g_i(x)| \varepsilon \zeta, & i = 0, \ldots, N, \\ 0, & i = N+1, \ldots, M, \end{cases}$$

where $g_i(x)$ is the ith PCE coefficient of $g(x, y)$. Here ζ is a Gaussian random variable with zero mean and unit standard deviation, and ε dictates the level of the data noise. In the present study, the random variable ζ is realized using the MATLAB function randn(). The data for Examples 1 and 2 studied below can be considered as an entire range of data $[d_0(x) - d_1(x), d_0(x) + d_1(x)]$ with $d_0(x)$ and $d_1(x)$ being the mean of the data and the bound on the variability of the data, respectively. It seems reasonable to expect that in many practical situations this kind of bounds can be obtained experimentally. The existing deterministic inverse techniques are usually assessed only on one set of data out of this range, and thus provide only a single estimate of the inverse solution without quantifying the associated uncertainty. In contrast, the SSFEM propagates the entire range of data to the inverse solution, and thus provides an entire ensemble of solutions consistent with the given data.

The PCE coefficients are not directly experimentally measurable, and thus their accurate construction from experimental data is of paramount practical importance. Some major progress has recently been made by Ghanem et al. [33,34], and the maximum likelihood method [33] and the Bayesian inference approach [34] are proposed for estimating PCE coefficients from limited data. Mathematically, the problem might be formulated as an approximation problem by multivariate polynomials with *data-driven* stochastic dimensionality. Its statistical correlation with the model uncertainties, e.g. material property uncertainties, has to be taken into account. For our initial demonstration of the algorithm, we employ the PCE coefficients as the given data in the present investigation.

The initial guess γ^0 was taken to be

$$\gamma_0^0(x) = 1, \qquad \gamma_i^0(x) = 0, \quad i = 1, 2, \ldots M$$

for all the examples. The conventional CGM is known to converge slowly, and it might stagnate after a few iterations [11]. To accelerate the convergence of the CGM, a preconditioner via the Sobolev inner product (see [35,12] for more details) is applied to the mean $\gamma_0(x)$ of the Robin coefficient $\gamma(x, y)$.

5.1. Convergence of the algorithm

As the first example, we consider the case that the inverse solution is smooth and the uncertainty is present only in the data.

Example 1. In Eq. (3), the conductivity $\alpha(x, y)$ is taken to be $\alpha(x, y) = e^{x_1}$, the exact solution to the deterministic forward problem is given by $u(x) = \sin x_1 e^{x_2} + x_1 + x_2$, and the Robin coefficient $\gamma(x, y)$ is given by

$$\gamma(x, y) = 1 + \sin(\pi x_1) + \frac{x_1}{2} y, \quad (x, y) \in \Gamma_i \times I.$$

Here the deterministic forward problem is obtained by ignoring the stochasticity, and the functions $f(x, y)$, $q(x, y)$ and $u_a(x)$ are taken to be appropriate deterministic functions such that its exact solution is given as above.

First we investigate the convergence of the algorithm with respect to refining the spatial mesh size h and increasing the order p of the PCE. To do that, at every iteration we evaluate the accuracy error e_γ defined by

$$e_\gamma = \|\gamma^* - \gamma^k\|_{L^2(I; L^2(\Gamma_i))}, \tag{16}$$

where γ^* is the exact Robin coefficient, and γ^k is the Robin coefficient on the boundary Γ_i reconstructed after k iterations. The residual E after the kth iteration is given by

$$E = \|u(\gamma^k) - g\|_{L^2(I; L^2(\Gamma_c))}. \tag{17}$$

In the case of exact data, the residual E tends to zero as the approximation γ^k converges to the exact solution.

Figs. 1(a) and (b) show, respectively, the accuracy error e_γ and the residual E as functions of the number of iterations, k, obtained for Example 1 when using $h \in \{1/10, 1/20, 1/30, 1/40\}$ and exact data. The spatial discretization error diminishes as the mesh size h refines, and consequently both the accuracy error e_γ and the residual E decrease. However, little could be gained when the mesh size h is smaller than $1/30$. Therefore, we use $h = 1/30$ in our subsequent analysis. Figs. 2(a) and (b) show, respectively, the accuracy error e_γ and the residual E as functions of the number of iterations, k, obtained for Example 1 when using $p \in \{2, 3, 4, 5, 6\}$ and exact data. The discretization error with the parametric domain I decreases with the increase of the PCE order p. However, the results do not improve much when the PCE order p is beyond 4, and therefore we use the PCE of order $p = 4$ in our computations. Note that the number of unknowns increases drastically with the increase of the PCE order p, which might deteriorate the convergence rate. Fig. 2 shows that the convergence is relatively independent of the PCE order p up to 40 iterations, and thereafter the convergence deteriorates slightly as the PCE order p increases from $p = 2$ to $p = 5$. However, the convergence of the numerical approximation is already achieved well before the 40th iteration, especially in the case of noisy data, see Fig. 5. Thus increasing the PCE order p seems practically harmless to the
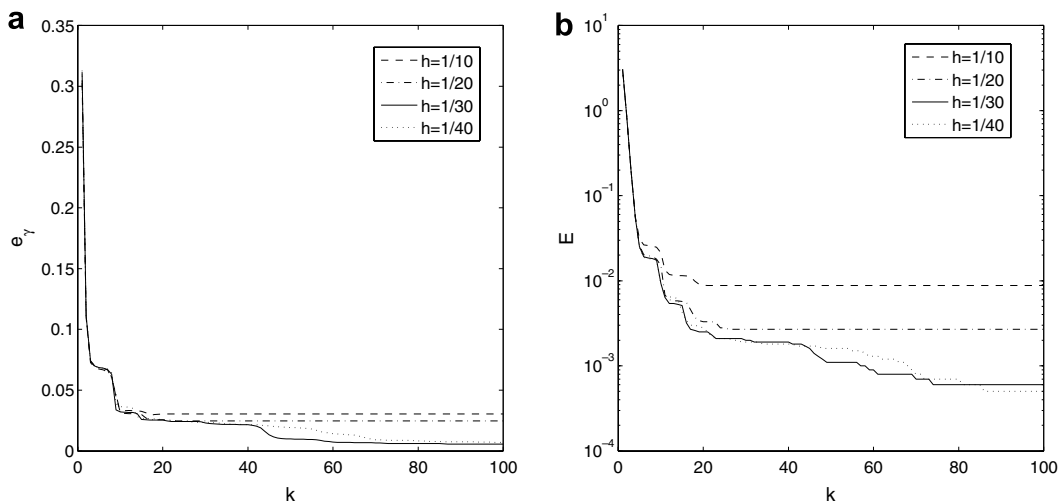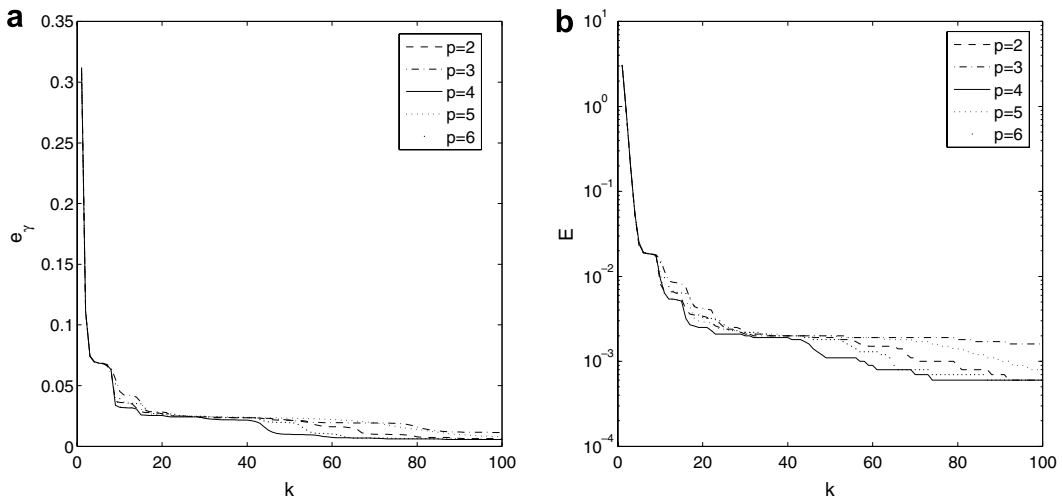

Fig. 1. (a) The accuracy error e_γ and (b) the residual E as functions of the number of iterations, k, for Example 1 with exact data and different mesh size h.

Fig. 2. (a) The accuracy error e_γ and (b) the residual E as functions of the number of iterations, k, for Example 1 with exact data and different PCE order p.

convergence of the algorithm. It is also observed that both the accuracy error e_γ and the residual E decrease as the number of iteration, k, increases, although the convergence slows down significantly as the iteration proceeds.

In the presence of data noise, the accuracy error of the numerical results by iterative regularization methods typically exhibits the so-called 'semi-convergence' phenomenon [36]. The approximation converges towards the exact solution up to a certain iteration number, and beyond this point, it deviates from the exact solution. To illustrate this point, we consider Example 1 with 1% noise in the data. Fig. 3(a) shows the convergence of the accuracy error e_γ obtained with different regularization parameters and without preconditioning. The semi-convergence phenomenon was observed in the case of small values for the regularization parameter η, e.g. $\eta = 0$ and $\eta = 1.0 \times 10^{-6}$. In particular, the formulation without incorporating the regularization term (i.e. $\eta = 0$) is unstable. An appropriate choice of the regularization parameter η, e.g. $\eta = 4 \times 10^{-5}$, can eliminate the semi-convergence phenomenon, and stabilizes the numerical algorithm. However, too large a value of the regularization parameter η, e.g. $\eta = 1 \times 10^{-4}$, can penalize the solution norm too much and lead to 'over-regularization'.
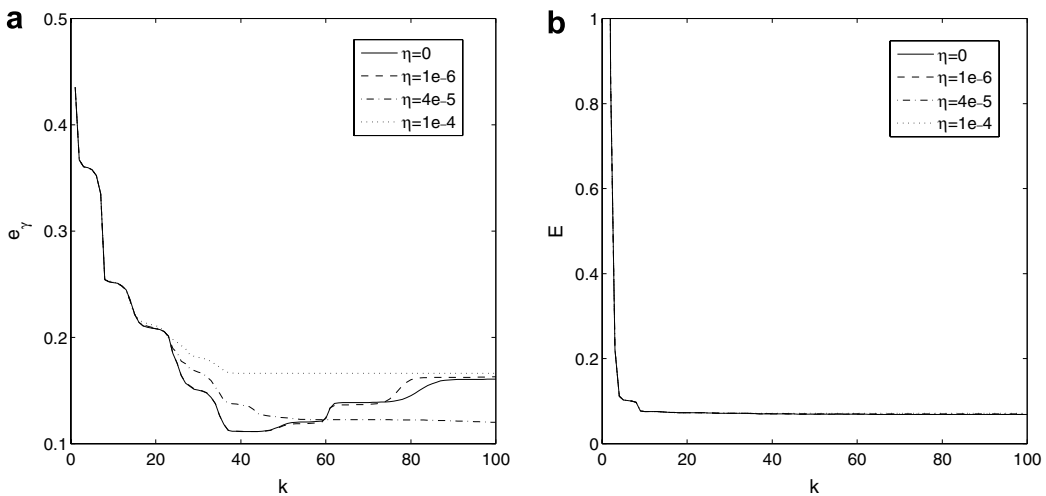

Fig. 3. (a) The accuracy error e_γ and (b) the residual E as functions of the number of iterations, k, for Example 1 with 1% noise in the data, obtained with different regularization parameter η and without preconditioning.

In the presence of the semi-convergence phenomenon, an appropriate stopping rule is necessary for obtaining an accurate and stable approximation to the inverse problem. Stopping criteria for iterative regularization methods are typically based on the discrepancy principle, the generalized cross-validation (GCV) or the *L*-curve criterion [36]. The basic underlying idea of the discrepancy principle is that one cannot expect that the approximation be more accurate than the data, and it ceases the iterative procedure for

$$k \in \mathbb{N} : E = \|u(\gamma^k) - g\|_{L^2(I;L^2(\Gamma_c))} \leqslant c\delta, \tag{18}$$

where δ is a measure of the amount of measurement errors of the Cauchy data, and the constant $c > 1$. Fig. 3(b) shows that the residual curve is rather flat over a wide range of iteration numbers, and thus the discrepancy principle requires a very accurate estimation of the amount of data noise in order that the CGM could yield a good approximation, which might not always be practical. And for heuristic methods disrespecting the noise level, e.g. the *L*-curve criterion and the GCV, it is generally difficult to obtain theoretical results, such as the convergence and convergence rate. Therefore these stopping criteria are not fully amenable with practical usage in the present context, and severe practical difficulties can arise in the applications of the mathematical formulation without incorporating a regularization term.

The use of the preconditioner via the Sobolev inner product could mitigate greatly the deleterious effect of data noise and numerical noise [12]. Fig. 4(a) shows the convergence of the approximate solutions with different regularization parameters in conjunction with preconditioning. The convergence of the approximate solution is rather steady, and no semi-convergence phenomenon is observed. Thus it undermines the crucial role of the number of iterations as the regularization parameter, such that a few extra iterations would not degrade the inverse solution much. Moreover, the value of the regularization parameter η is less crucial for the preconditioned CGM. The numerical results in terms of both the accuracy error e_γ and the residual E for all four regularization parameters considered are practically identical, see Figs. 4(a) and (b), respectively. A comparison of Fig. 4(a) with Fig. 3(a) indicates preconditioning can accelerate the convergence of the CGM significantly: The conventional CGM converges after approximately 40 iterations, while the preconditioned one converges within 20 iterations. The comparison in terms of the number of iterations is fair as the computational cost of applying the preconditioner is negligible compared with that of the forward solver.

Figs. 5(a) and (b) show the accuracy error e_γ and the residual E, respectively, for Example 1 with various levels of noise in the data. Both the accuracy error e_γ and the residual E decrease as the noise level ε decreases. The attainable minimum residual is restricted by the noise level and the discretization error, and it cannot be brought down to arbitrarily small levels.
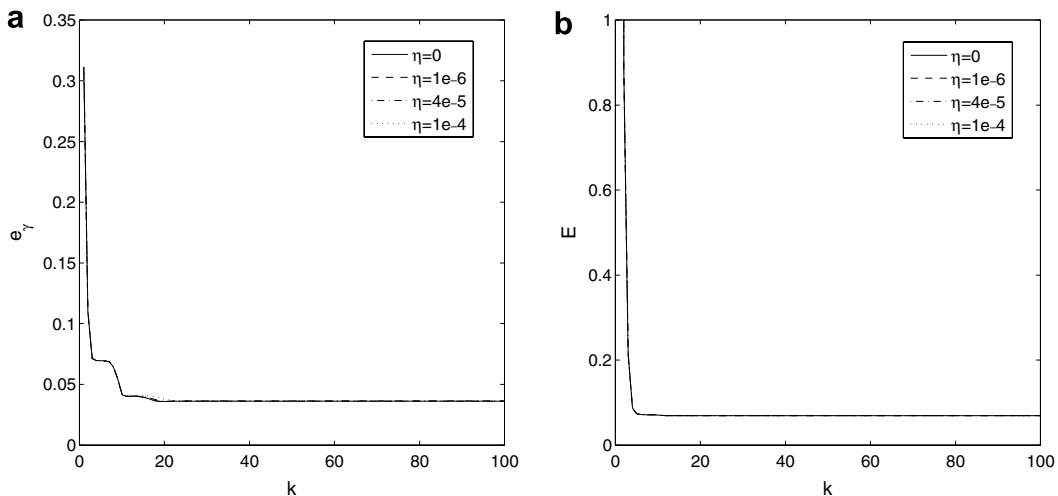

Fig. 4. (a) The accuracy error e_γ and (b) the residual E as functions of the number of iterations, k, for Example 1 with 1% noise in the data, obtained with different regularization parameter η and with preconditioning.
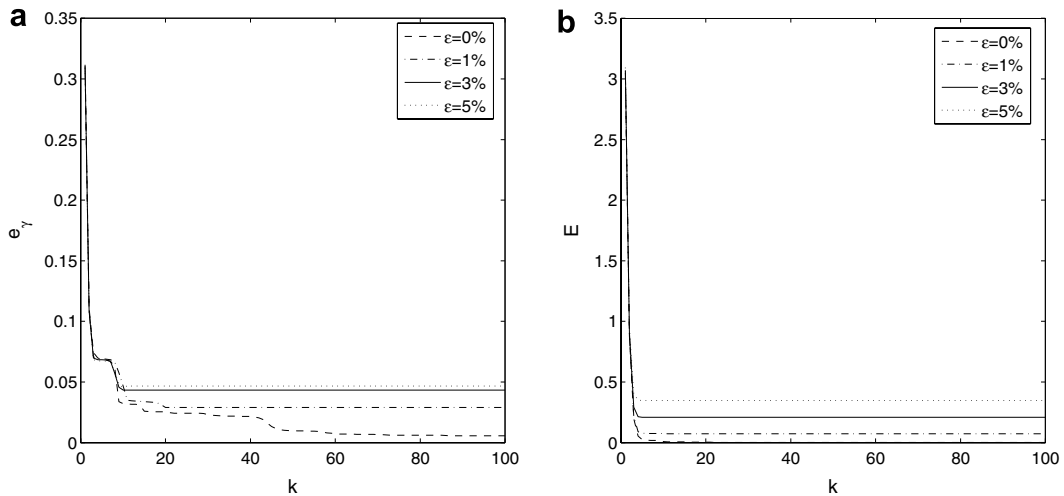
Fig. 5. (a) The accuracy error e_γ and (b) the residual E as functions of the number of iterations, k, for Example 1 with various levels of noise in the data.

5.2. Numerical results and discussions

The numerical results for Example 1 with various levels of noise in the data are shown in Fig. 6. The numerical solution is practically identical with the exact Robin coefficient for exact data. With up to 5% noise in the data, the mean of the Robin coefficient is still in excellent agreement with the exact one, and the first order PCE coefficient represents a reasonable approximation to the exact one. Higher order terms of the PCE, see Figs. 6(c) and (d), are less accurate, but their magnitude is maintained at a small level. It is noted that it usually takes more iterations for the higher order terms to converge, and thus the results presented in the figure are overall optimal, which may not be optimal for each coefficient.

From Figs. 6(a) and (b), it is observed that, while remaining stable, the numerical Robin coefficient converges and predicts the exact solution better as the amount of noise in the data decreases. Therefore, the numerical algorithm is convergent with respect to decreasing the amount of data noise. The accuracy of the numerical solutions for noisy data of level 5% are presented in Table 1. It is observed that the error in the numerical solution is maintained at a small level comparable with that in the data.

Accurate reconstruction of non-smooth solutions is numerically very challenging. As the second example, we consider the more difficult situation where the inverse solution is non-smooth. Again the uncertainty is present only in the data.

Example 2. In Eq. (3), the conductivity $\alpha(x, y)$ is taken to be $\alpha(x, y) = 1$, the exact solution to the deterministic forward problem is given by $u(x) = \sin x_1 e^{x_2}$, and the Robin coefficient $\gamma(x, y)$ is given by

$$\gamma(x, y) = \begin{cases} 1 + 2x_1 + \frac{x_1}{2}y, & (x, y) \in (0, \frac{1}{2}) \times \{1\} \times I, \\ 3 - 2x_1 + \frac{x_1}{2}y, & (x, y) \in [\frac{1}{2}, 1) \times \{1\} \times I. \end{cases}$$

The numerical results for Example 2 with various levels of noise in the data are shown in Fig. 7. The mean of the numerical Robin coefficient is in good agreement with the exact solution, taking into account the severe ill-posed nature of the problem and non-smoothness of the solution. However it fails to capture distinct features of non-smooth solutions, e.g. the corner. This is attributed to both the CGM and the smoothing nature of the Sobolev inner product preconditioner [12]. The total-variation or the bounded-variation regularization might be employed to overcome the undesirable smoothing effect. The inaccuracy in the approximation of the mean adversely affects the estimation of the first order PCE coefficient, see Fig. 7(b), which consequently loosens the probability bound.

We stress that the inverse solution is overall optimal, and not necessary optimal for each component. This point is evidenced by comparing the results for the cases of exact data and that of 1% noise in the data. The
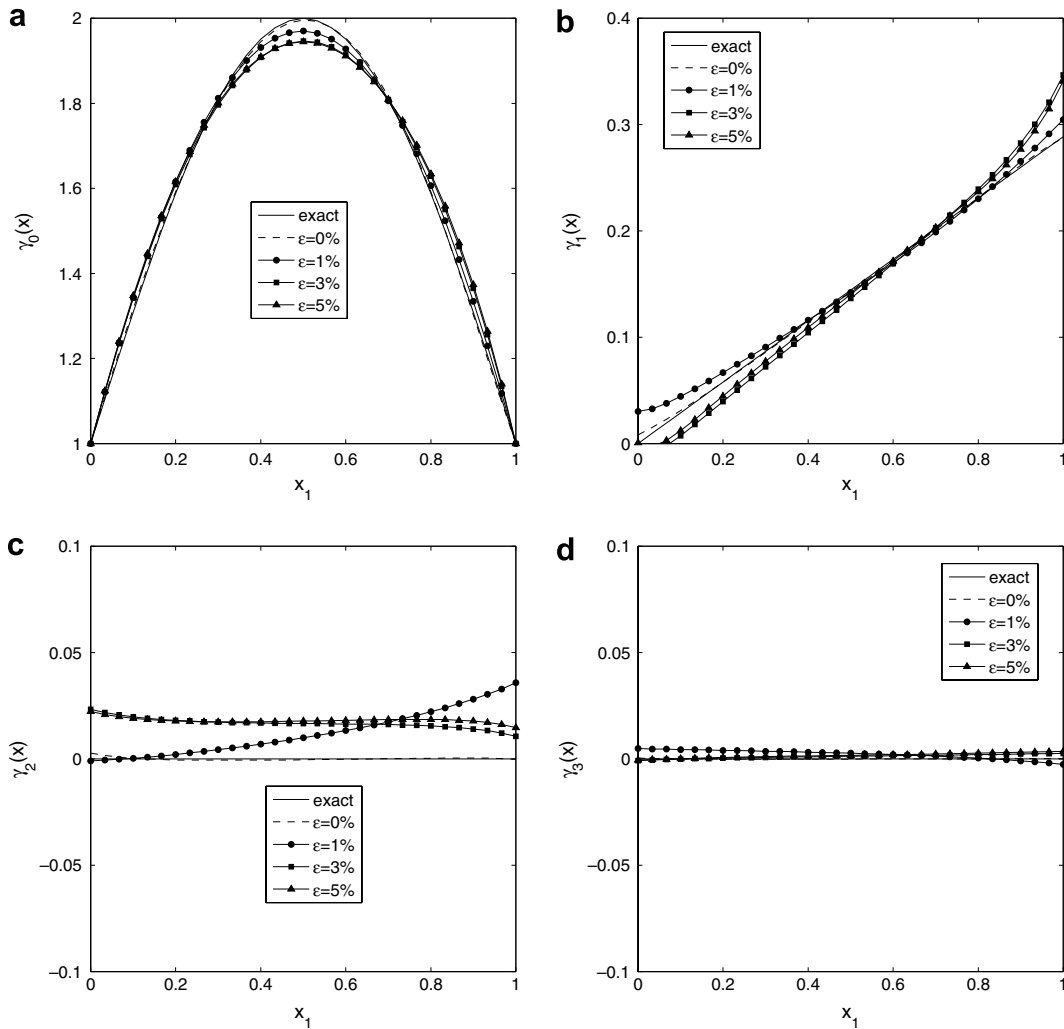
Fig. 6. The numerical results of (a) mean, (b) 1st, (c) 2nd, and (d) 3rd terms in the PCE for Example 1 with various levels of noise in the data.

Table 1
Numerical results for the examples with 5% noise in the data

Example	k	e_γ	E
1	10	4.53×10^{-2}	3.47×10^{-1}
2	14	7.32×10^{-2}	1.95×10^{-1}
3	18	3.80×10^{-2}	2.74×10^{-1}
4	12	5.07×10^{-2}	2.90×10^{-1}

first order PCE coefficient for the former seems less accurate than the latter, as its deviation near the end points is larger. However, this is compensated by the fact that the mean for the former is more accurate than the latter. The accuracy error for exact data is 5.35×10^{-2}, as compared with 7.01×10^{-2} for 1% noise in the data, and thus overall more accurate. The magnitude of higher order terms of the PCE are maintained at a small level comparable with the noise level, see Figs. 7(c) and (d). Therefore, the non-smoothness could affect the PCE coefficients, but its effect could be negligible on some terms.

Till now, we have considered only problems with deterministic thermal conductivity. We examine the effect of material property uncertainties on the inverse solution by considering Example 3.
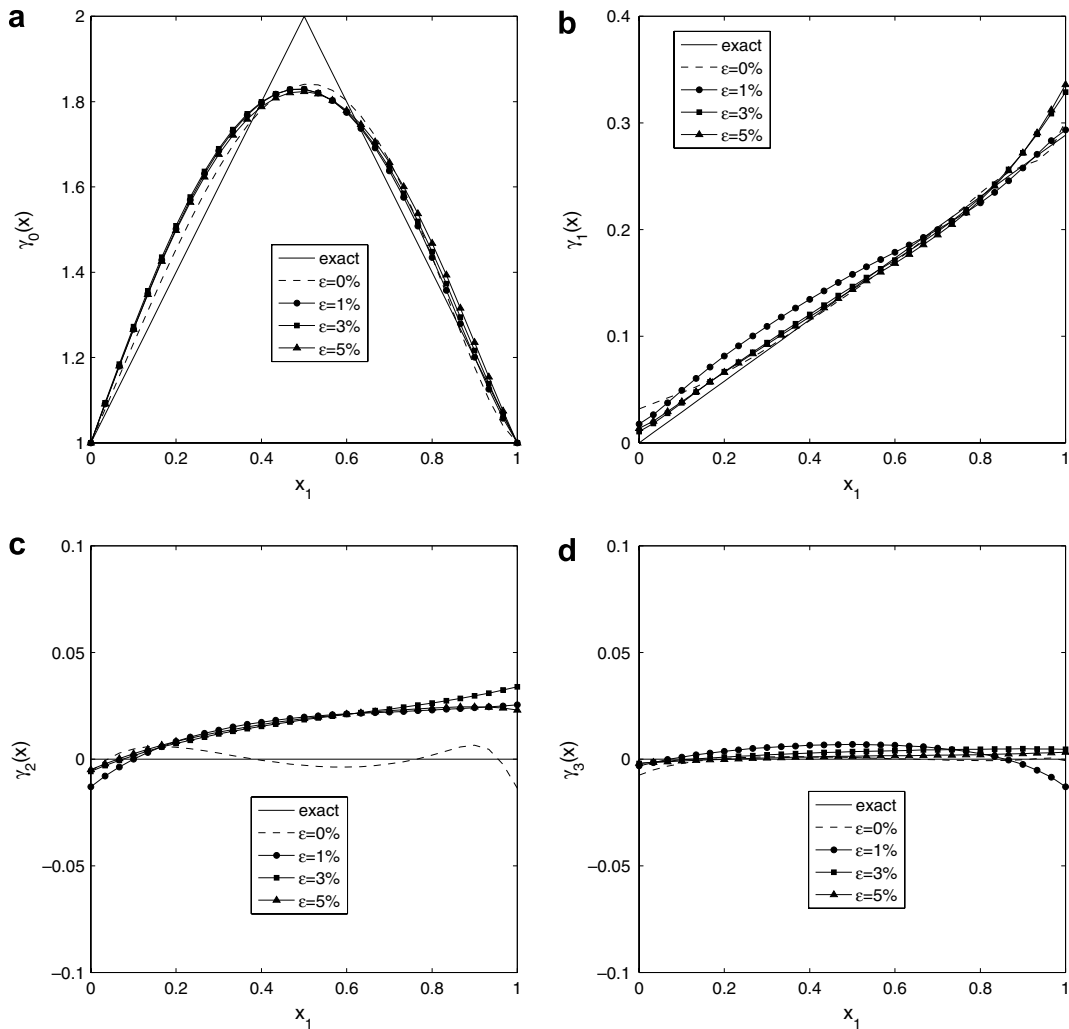
Fig. 7. The numerical results of (a) mean, (b) 1st, (c) 2nd, and (d) 3rd terms in the PCE for Example 2 with various levels of noise in the data.

Example 3. In Eq. (3), the conductivity $\alpha(x, y)$ is taken to be $\alpha(x, y) = e^{x_1} + \frac{1}{2} x_1 y_1$, the exact solution to the deterministic forward problem is given by $u(x) = x_1^2 + x_2^2$, and the Robin coefficient $\gamma(x, y)$ is given by

$$\gamma(x, y) = 1 + \sin(\pi x_1) + \frac{x_1}{2} y_2, \quad (x, y) \in \Gamma_i \times I.$$

The problem is discretized using 722 finite elements. For the ease of presentation, we differentiate the stochasticity of the Robin coefficient and that of the conductivity by the prefixes γ- and α-, respectively. We consider two cases for Example 3: Case 1 uses both γ- and α-first order PCEs of the Cauchy data, and Case 2 uses only the γ-first order PCE.

The numerical results for Case 1 of Example 3 with various levels of noise in the data are presented in Fig. 8, where the first and second subscripts indicate the order of the polynomial chaos with respect to γ- and α-random variables, respectively. Both the mean and the γ-first order PCE coefficient of the numerical Robin coefficient are in excellent agreement with the exact one with up to 5% noise in the data, and the magnitude of the γ-second order PCE coefficient is small. Therefore, the stochasticity of the conductivity has little effect on the mean and γ-PCE coefficients of the numerical Robin coefficient. Table 1 shows the accuracy error for Example 3 with 5% noise, which is also comparable to that of Example 1.
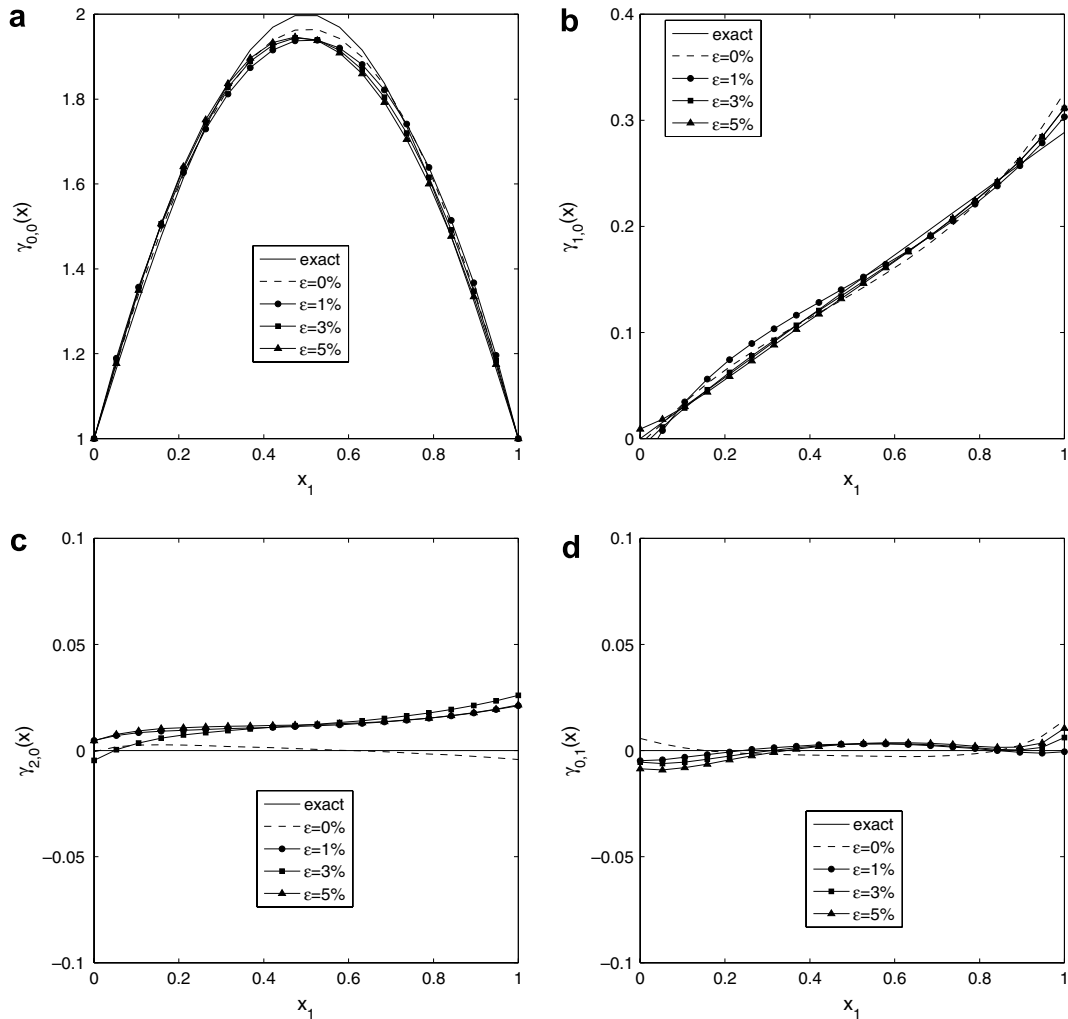
Fig. 8. The numerical results of (a) mean, (b) γ-1st, (c) γ-2nd, and (d) α-1st terms in the PCE for Case 1 of Example 3 with various levels of noise in the data.

The α-uncertainty indeed affects the numerical Robin coefficient in the sense that the random dimension of its PCE coefficient is increased by one. Fig. 8(d) shows the α-first order PCE coefficient of the numerical Robin coefficient. The magnitude of the α-first order PCE coefficient is only of the noise level, and thus the effect of α-uncertainty might be considered minor. To further assess its effect on the numerical Robin coefficient, we display its variance for Case 1 of Example 3 with exact data and 5% noise in the data in the logarithmic scale in Figs. 9(a) and (b), respectively. It is observed that both the γ-variance and the overall variance are in excellent agreement with that of the exact variance, and the α-variance, which here includes all terms not included in the γ-variance, contributes very little to the overall variance of the inverse solution. The magnitude of the α-variance increases slightly for noisy data, but it is still negligible compared with that of γ-variance up to 5% noise in the data. Therefore, the α-uncertainty has only immaterial effect on the inverse solution if it is appropriately reflected in the data, e.g. with the α-first order PCE characterized.

The numerical results for Case 2 of Example 3 with 5% noise in the data are presented in Fig. 10, where the results for Case 1 are also presented for the convenience of comparison. The mean and the γ-first order PCE coefficient for Case 2 deteriorate only slightly as compared with that of Case 1, and the γ-second PCE coefficient is practically of the same small magnitude. Thus inadequate α-uncertainty characterization of the Cauchy data has limited effect on the γ-PCE coefficients of the numerical Robin coefficient. Fig. 10(d) shows
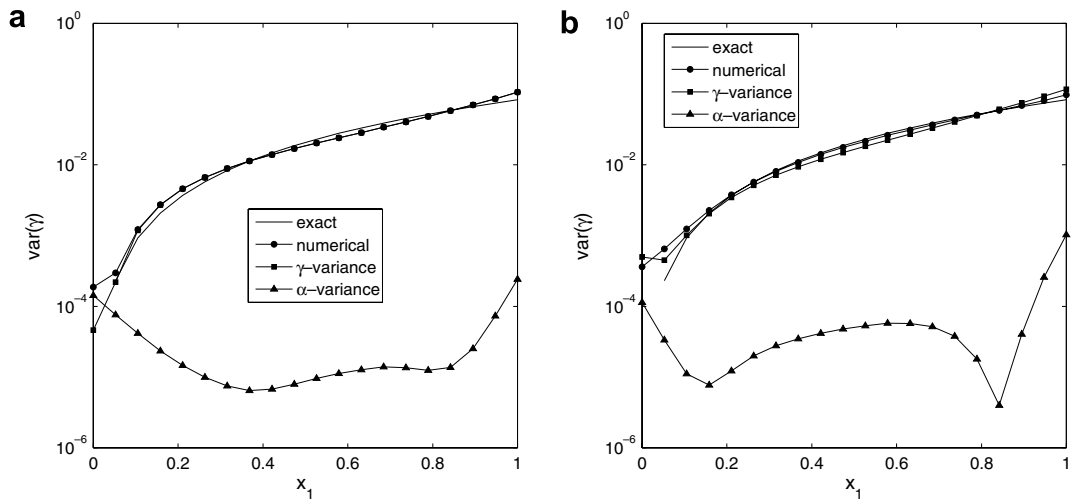
Fig. 9. The variance of the Robin coefficient for Example 3 with (a) exact data, and (b) 5% noise in the data.
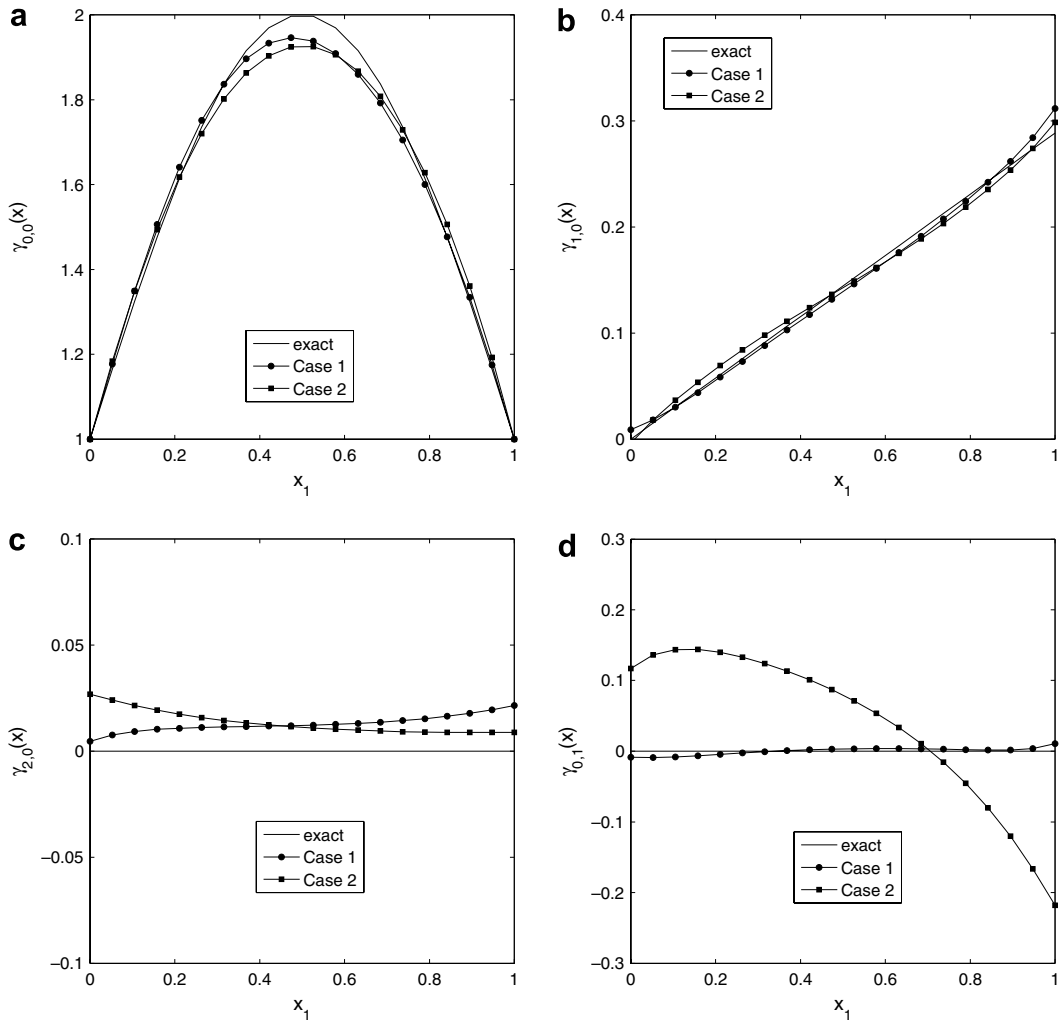

Fig. 10. The numerical results for Example 3 in Case 2: (a) mean, (b) γ-1st, (c) γ-2nd, and (d) α-1st PCE coefficients compared with Case 1.

the α-first order PCE coefficient of the inverse solution. The α-PCE coefficient for Case 2 is larger than that for Case 1 approximately by one order in magnitude, which consequently renders the probabilistic bounds of the numerical Robin coefficient less accurate. Therefore, the α-stochasticity could have significant effect on the inverse solution if the data is inadequately characterized in terms of the α-uncertainty. This has a significant practical implication that the model selection of the data plays a crucial role in obtaining realistic probabilistic calibration of the inverse solution.

So far we have considered only problems with the stochasticity given by random variables. We next consider a more realistic setting where the stochasticity is described by a second order random process.

Example 4. In Eq. (3), the conductivity $\alpha(x, \omega)$ is taken to be a second order process with its covariance function $C(r)$ given by

$$C(r) = \sigma^2 \frac{r}{b} K_1\left(\frac{r}{b}\right),$$

where K_1 is the modified Bessel function of the second kind of order one, b scales as the correlation length, r is the distance between two spatial points and σ^2 is the variance of the random field. For any given point in the random field, this type of covariance function takes into account the influences of its nearest neighbors in all
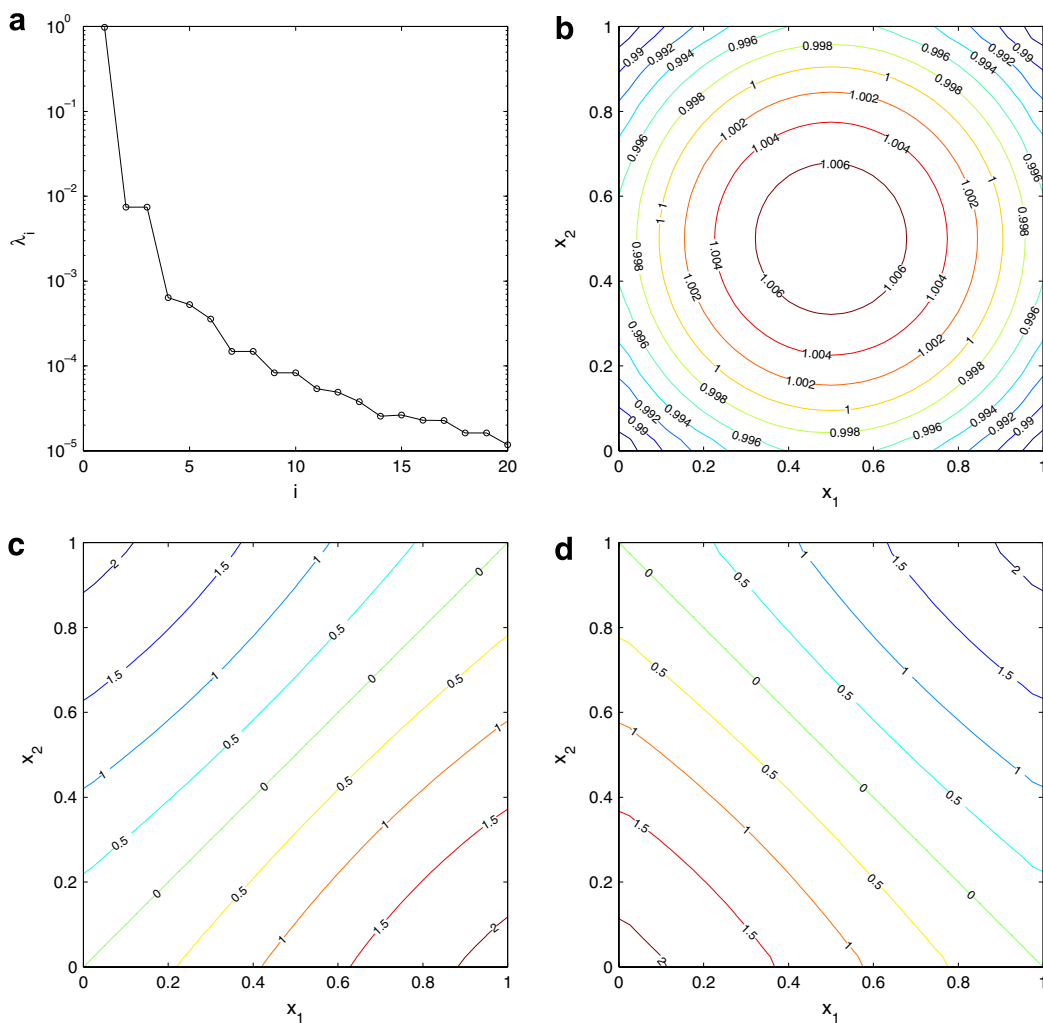

Fig. 11. (a) The distribution of the first 20 normalized eigenvalues λ_i of the covariance operator, and the contour of (b) the 1st, (c) the 2nd, and (d) the 3rd eigenfunctions.

directions and is considered as the 'elementary' covariance function in two-dimensional domains [22]. Here we employ this covariance function and apply the Galerkin procedure [19] to numerically solve the eigenvalue problem (2) for obtaining the eigenvalues λ_i and eigenfunctions $f_i(x)$. For the purpose of initial demonstration, a relatively strong auto-correlation with parameter $b = 5$ is assumed, which results in a fast decay of the eigenvalues. Fig. 11(a) shows the first twenty normalized eigenvalues of the covariance operator, and the first three eigenvalues are found to encompass over 99.9% of the total 'energy' $\sum_i \lambda_i$ of the eigenvalue spectrum. Therefore, the first three eigenfunctions, which are shown in Figs. 11(b)–(d), are sufficient for accurately characterizing the random field. We employ the first three eigenfunctions for describing $\alpha(x, \omega)$ in the sequel. The mean of the random field $\alpha(x, \omega)$ is taken to be e^{x_1}, and the variance σ^2 is taken to be 0.5. The exact solution to the deterministic forward problem is given by $u(x) = x_1^2 + x_2^2$, and the Robin coefficient $\gamma(x, y)$ is given by

$$\gamma(x, y) = 1 + \sin(\pi x_1) + \frac{x_1}{2} y_4, \quad (x, y) \in \Gamma_i \times I.$$

The problem is discretized using 800 finite elements, and the PCE order p is taken to be $p = 3$, which results in a PCE of 35 terms. The numerical results for Example 4 with various levels of noise in the data are shown in Fig. 12. The accuracy of the numerical results is comparable with that of Examples 1–3, see Table 1. The mean
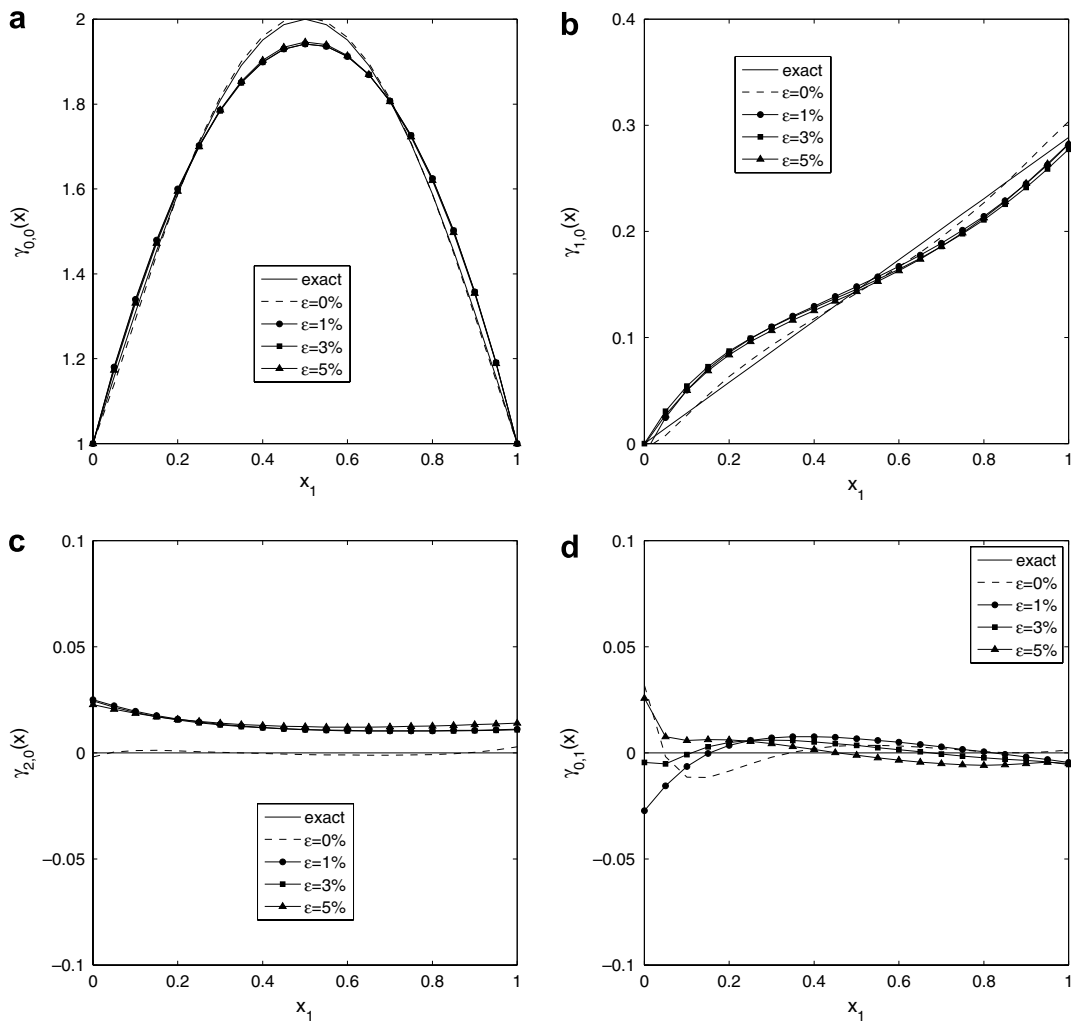

Fig. 12. The numerical results of (a) mean, (b) γ-1st, (c) γ-2nd, and (d) α-1st terms in the PCE for Example 4 with various levels of noise in the data.

of the numerical Robin coefficient is in good agreement with the exact solution for all four noise levels considered, and their accuracy is comparable with that of Example 3. The numerical estimates of γ-first order PCE coefficient also represent reasonable approximations to the exact one. The magnitude of other PCE terms are all maintained below the noise level, see Figs. 12(c) and (d) for the γ-second and α-first PCE coefficients, respectively. Note that the number of iterations necessary for achieving the convergence of the numerical results is observed to be comparable with that of Example 3, see Table 1. Therefore, the convergence of the algorithm is only mildly dependent of the stochastic dimension N for the Robin inverse problem. However, it still takes much longer computational time as the forward solver this time is computationally more expensive. Recent algorithm innovations of the SFEM, e.g. the sparse grid technique [28,37], model reduction [29,38] and the adaptive multi-element technique [39], may provide viable means for accelerating the algorithm.

Many other numerical experiments have been performed, and the numerical results obtained are similar to those presented above. The forgoing numerical verifications indicate that the proposed algorithm is efficient, stable and accurate for the reconstruction of smooth Robin coefficients, and also yields acceptable results for non-smooth solutions.

6. Conclusions

This paper investigates a variational approach to the stochastic Robin inverse problem. Mathematical properties relevant to its numerical computations of the optimization problem are investigated, and the convergence of the spectral stochastic finite element approximation is established. Numerical results by a preconditioned conjugate gradient method for several two-dimensional inverse problems are presented, and the results indicate that the algorithm is convergent with respect to refining the mesh size and decreasing the amount of data noise, reasonably accurate and computationally efficient.

This study presents only a first step into the fascinating field of numerical methods for stochastic inverse problems, and there are several avenues for further extensions, which are currently under investigation. Firstly, more works is required concerning the theoretical analysis of the numerical algorithm, e.g. the convergence of the SSFEM approximation for general probability distributions, and also compact embedding and density results for weighted Sobolev spaces. Secondly, the algorithm extends straightforwardly to other inverse problems, such as parameter identification and boundary identification problems. Thirdly, stochastic modeling can affect significantly the inverse solution. Therefore, the related model selection, verification and validation issues are imperative, and novel computing machineries from statistics may be necessary. Finally, alternative formulations of the problem based on physically more relevant statistical quantities, e.g. moments or probability distribution functions, that derive readily from experimental data, are also of practical interest. This is particularly relevant in the limited experimental data case that does not allow a complete probabilistic calibration.

Acknowledgments

References

[1] A.M. Osman, J.V. Beck, Nonlinear inverse problem for the estimation of time-and-space dependent heat transfer coefficients, J. Thermophys. Heat Transfer 3 (1989) 146–152.
[2] F.M. White, Heat and Mass Transfer, Addison-Wesley, Reading, 1988.
[3] T.J. Martin, G.S. Dulikravich, Inverse determination of steady heat convection coefficient distributions, J. Heat Transfer 120 (1998) 328–334.
[4] S. Chantasiriwan, Inverse determination of steady-state heat transfer coefficient, Int. Commun. Heat Mass Transfer 27 (2000) 1155–1164.
[5] P.G. Kaup, F. Santosa, Nondestructive evaluation of corrosion damage using electrostatic measurements, J. Nondestruct. Eval. 14 (1995) 127–136.

[6] G. Inglese, An inverse problem in corrosion detection, Inverse Probl. 13 (1997) 977–994.
[7] W. Fang, E. Cumberbatch, Inverse problems for metal oxide semiconductor field-effect transistor contact resistivity, SIAM J. Appl. Math. 52 (1992) 699–709.
[8] S. Chaabane, J. Ferchichi, K. Kunisch, Differentiability properties of the L^1-tracking functional and application to the Robin inverse problem, Inverse Probl. 20 (2004) 1083–1097.
[9] W. Fang, M. Lu, A fast collocation method for an inverse boundary value problem, Int. J. Numer. Methods Eng. 59 (2004) 1563–1585.
[10] F. Lin, W. Fang, A linear integral equation approach to the Robin inverse problem, Inverse Probl. 21 (2005) 1757–1772.
[11] B. Jin, Conjugate gradient method for the Robin inverse problem associated with the Laplace equation, Int. J. Numer. Methods Eng. 71 (2007) 433–453.
[12] B. Jin, J. Zou, Numerical estimation of the Robin coefficient in diffusion equation, submitted for publication.
[13] B. Jin, J. Zou, Inversion of Robin coefficient by a spectral stochastic finite element approach, Technical Report CUHK-2007-10 (351), Department of Mathematics, The Chinese University of Hong Kong.
[14] J.L. Xie, J. Zou, Numerical reconstruction of heat fluxes, SIAM J. Numer. Anal. 43 (2005) 1504–1535.
[15] M. Le Ravalec, Inverse Stochastic Modeling of Flow in Porous Media: Applications to Reservoir Characterization, Editions Technip, Paris, 2005.
[16] J. Glimm, D.H. Sharp, Prediction and the quantification of uncertainty, Physica D 133 (1999) 152–170.
[17] J. Kaipio, E. Somersalo, Statistical and Computational Inverse Problems, Springer, New York, 2005.
[18] I. Babuška, R. Tempone, G.E. Zouraris, Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation, Comput. Methods Appl. Mech. Eng. 194 (2005) 1251–1294.
[19] R.G. Ghanem, P.D. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer, New York, 1991.
[20] V.A. Badri Narayanan, N. Zabaras, Stochastic inverse heat conduction using a spectral approach, Int. J. Numer. Methods Eng. 60 (2004) 1569–1593.
[21] I. Babuška, K.M. Liu, R. Tempone, Solving stochastic partial differential equations based on the experimental data, Math. Models Methods Appl. Sci. 13 (2003) 415–444.
[22] D. Xiu, G.E. Karniadakis, Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos, Comput. Methods Appl. Mech. Eng. 191 (2002) 4927–4948.
[23] D. Xiu, G.E. Karniadakis, A new stochastic approach to transient heat conduction modeling with uncertainty, Int. J. Heat Mass Transfer 46 (2003) 4681–4693.
[24] D. Xiu, G.E. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2002) 619–644.
[25] I. Babuška, R. Tempone, G.E. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, SIAM J. Numer. Anal. 42 (2004) 800–825.
[26] L.C. Evans, R.F. Gariepy, Measure Theory and Fine Properties of Functions, CRC Press, Boca Raton, 1992.
[27] O.M. Alifanov, Inverse Heat Transfer Problems, Springer-Verlag, Berlin, 1994.
[28] D. Xiu, J.S. Hesthaven, High order collocation methods for differential equations with random inputs, SIAM J. Sci. Comput. 27 (2005) 1118–1139.
[29] S. Acharjee, N. Zabaras, A concurrent model reduction approach on spatial and random domains for the solution of stochastic PDEs, Int. J. Numer. Methods Eng. 66 (2006) 1934–1954.
[30] N. Wiener, The homogeneous chaos, Am. J. Math. 60 (1938) 897–936.
[31] R. Cameron, W. Martin, The orthogonal development of non-linear functionals in series of Fourier–Hermite functionals, Ann. Math. 48 (1947) 385–392.
[32] A. Kufner, Weighted Sobolev Spaces, Wiley, Chichester, New York, 1985.
[33] C. Desceliers, R. Ghanem, C. Soize, Maximum likelihood estimation of stochastic chaos representations from experimental data, Int. J. Numer. Methods Eng. 66 (2005) 978–1001.
[34] R. Ghanem, A. Doostan, On the construction and analysis of stochastic models: characterization and propagation of the errors associated with limited data, J. Comput. Phys. 217 (2006) 63–81.
[35] A. Jameson, J.C. Vassberg, Studies of alternative numerical optimization methods applied to the Brachistochrone problem, Comput. Fluid Dyn. J. 9 (2000) 281–296.
[36] P.C. Hansen, Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion, SIAM, Philadelphia, 1998.
[37] B. Ganapathysubramanian, N. Zabaras, Sparse grid collocation scheme for stochastic natural convection problems, J. Comput. Phys. 225 (2007) 652–685.
[38] A. Doostan, R.G. Ghanem, J. Red-Horse, Stochastic model reduction for chaos representations, Comput. Methods Appl. Mech. Eng. 196 (2007) 3951–3966.
[39] X. Wan, G.E. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, J. Comput. Phys. 209 (2005) 617–642.