

## ACCELERATING THE BAYESIAN INFERENCE OF INVERSE PROBLEMS BY USING DATA-DRIVEN COMPRESSIVE SENSING METHOD BASED ON PROPER ORTHOGONAL DECOMPOSITION

MEIXIN XIONG, LIUHONG CHEN AND JU MING\*

Department of Mathematics and Statistics, Huazhong University of Science and Technology  
Wuhan 430074, China

JAEMIN SHIN

Department of Mathematical Sciences, Hanbat National University  
Daejeon, Republic of Korea

**ABSTRACT.** In Bayesian inverse problems, using the Markov Chain Monte Carlo method to sample from the posterior space of unknown parameters is a formidable challenge due to the requirement of evaluating the forward model a large number of times. For the purpose of accelerating the inference of the Bayesian inverse problems, in this work, we present a proper orthogonal decomposition (POD) based data-driven compressive sensing (DCS) method and construct a low dimensional approximation to the stochastic surrogate model on the prior support. Specifically, we first use POD to generate a reduced order model. Then we construct a compressed polynomial approximation by using a stochastic collocation method based on the generalized polynomial chaos expansion and solving an  $l_1$ -minimization problem. Rigorous error analysis and coefficient estimation was provided. Numerical experiments on stochastic elliptic inverse problem were performed to verify the effectiveness of our POD-DCS method.

**1. Introduction.** The inverse problems refer to exploring the inherent nature of things based on observable phenomena. There are many parameters of interest in scientific and engineering problems that cannot be directly observed. We often discuss these quantities indirectly through known data, that is, to formulate such problems in the form of inverse problems [14, 7]. It can be seen from this that inverse problems are the products of the rapid development of science and engineering, which have been widely used in various fields such as geological engineering [25, 10], medicine [4], environment [31], telemetry [28], control [2] and so on.

Due to the fact that parameters are sensitive to observation data, which is usually impure, this leads to the inverse problems are ill-posed, which can be addressed with regularization [9, 27]. However, regularization methods only provide point estimates of the unknown parameters and without quantifying the uncertainties of the solution. Then, the statistical inference method [14, 26] entered the researchers'

---

2020 *Mathematics Subject Classification.* Primary: 65D15, 65N21, 62F15; Secondary: 60H15.

*Key words and phrases.* Bayesian inference, inverse problems, data-driven method, compressive sensing, proper orthogonal decomposition.

\* Corresponding author: jming@hust.edu.cn.

vision. In Bayesian inverse problems, the unknown parameters are treated as random vectors with known prior distribution, and we need to infer their conditional distribution under observation data, namely the posterior distribution. In general, it is difficult to obtain the analytical form of the posterior distribution, so we often consider using the Markov Chain Monte Carlo (MCMC) method [3] to explore the posterior space. In these sampling algorithms, the forward model solution needs to be calculated for each candidate sample. We know that evaluating forward models is expensive in many practical problems, so direct sampling algorithms are prohibitive. To overcome this difficulty, many scholars have proposed a series of surrogate models and effective sampling algorithms to reduce the computational cost of Bayesian inverse problems. The former achieves the goal by reducing the cost of a single forward model evaluation. For example, the use of generalized polynomial chaos (gPC) basis functions to express the solutions of forward problems is proposed in [21, 22], the projection-based model reduction techniques can be seen in [19, 15], and adaptive methods to construct surrogate models are presented in [32, 17]. The latter achieves this purpose by reducing the number of samples required, which has been studied in [23, 20].

Here, we focus on the construction of effective surrogate models. Since unknown parameters are treated as random vector  $\boldsymbol{\xi}(\omega)$ , the forward problems can be regarded as the stochastic problems on prior support  $\Gamma$ . Although the stochastic problem is more complicated than the original one, we can calculate the approximate posterior probability corresponding to every candidate sample at a negligible cost once its solution expression obtained. In the physical domain  $D$  and image probability space of unknown parameters, we can use the deterministic and stochastic basis functions to represent the solutions of stochastic problems [13, 8]. For instance, with the stochastic finite element methods, an approximation of solution  $u(\mathbf{x}, \boldsymbol{\xi}) : D \times \Gamma \rightarrow \mathbb{R}$  has the form:

$$\check{u}(\mathbf{x}, \boldsymbol{\xi}) = \sum_{i=1}^n \sum_{j=1}^{n_u} \check{c}_{ij} \varphi_j(\mathbf{x}) \psi_i(\boldsymbol{\xi}), \quad (1)$$

where the coefficient  $\check{c}_{ij}$  is with respect to (w.r.t.) a finite element basis  $\varphi_j$  and a multivariate polynomial chaos (PC) basis  $\psi_i$ .

In the current work, we propose the data-driven compressive sensing method based on proper orthogonal decomposition for constructing the accurate approximate solutions of stochastic forward problems to accelerate the calculation of Bayesian inverse problems. The POD-DCS method is derived from the data-driven compressive sensing method proposed in [18]. But here, we use proper orthogonal decomposition (POD) basis functions instead of Karhunen-Loève basis, so as to generate a reduced order model (ROM) and avoid the recovery of covariance. According to the idea of compressive sensing (CS), an accurate approximate solution of the ROM can be obtained by calculating a basis pursuit (BP) problem [18, 30] with a small amount of data. The advantage of our method is that it constructs the reduced order model using snapshots first, which causes the degree of freedom (DoF) of the forward problem to be reduced from the cardinality of the finite element to the that of POD, and then we only need to reconstruct the low-dimensional ROM at a lower cost. From the numerical results we can see that when using the same number of fully discrete finite element solutions, our scheme can improve accuracy and sparsity compared to the conventional CS method based on finite element basis. Moreover, the cost of evaluating a forward model using our method is only a fraction of that with the finite element method (FEM), so it can speed up the

calculation of Bayesian inverse problems effectively. As we all know, many practical problems can be described by partial differential equations (PDE). Here, we concentrate only on elliptic PDE, which is widely used in the studies of oil reservoirs and groundwater [25], and its ill-posedness have been discussed in [16]. Of course, our method can be naturally extended to other partial differential equations. All the computations were performed using MATLAB R2014b on a personal computer with 1.60GHz CPU and 4GB RAM.

The rest of this paper is organized as follows. In Section 2, we describe the model problem used as the background of our study, then introduce the framework of Bayesian inference and the stochastic surrogate model. In Section 3, we discuss how to construct the POD-DCS approximate solution for a stochastic forward problem, and provide some direct simulation results and the corresponding algorithm. The error and sparsity analyses of our scheme are conducted in Section 4. In Section 5, we compare our POD-DCS scheme with other methods, and use it to solve elliptic inverse problem. Finally, we give some conclusions and indicate possible future work in Section 6.

**2. Problem setup and Bayesian framework.** Consider an underground steady state aquifers modelled by two-dimensional elliptic partial differential equation with Dirichlet boundary as

$$\begin{cases} -\nabla \cdot (a(\mathbf{x})\nabla u(\mathbf{x})) &= f, & \mathbf{x} \in D, \\ u(\mathbf{x}) &= 0, & \mathbf{x} \in \partial D, \end{cases} \tag{2}$$

where domain  $D = [0, 1]^2$ . Let  $u(\mathbf{x})$  be the hydraulic head,  $a(\mathbf{x})$  be the permeability field, and apply a constant source term in the whole domain and satisfy  $f \equiv 2$ . The goal of this elliptic inverse problem is to estimate the unknown permeability field  $a(\mathbf{x})$  from a set of observations  $\mathbf{z} = [u(\mathbf{x}_1, \boldsymbol{\xi}), \dots, u(\mathbf{x}_{n_z}, \boldsymbol{\xi})]^T \in \mathbb{R}^{n_z}$  of the solution  $u$  at points  $\{\mathbf{x}_s\}_{s=1}^{n_z}$ . In this work, we consider the permeability field defined by  $n_p = 4$  radial basis functions with centers  $\mathbf{x}_1, \dots, \mathbf{x}_{n_p}$  as:

$$a(\mathbf{x}) = 0.01 + \sum_{i=1}^{n_p} \xi_i \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_i\|_2^2}{0.02}\right), \tag{3}$$

where  $\|\cdot\|_2$  is Euclidean norm. These radial basis functions are shown in Figure 1, and what we need to do is determine these weights  $\{\xi_i\}_{i=1}^{n_p}$ .

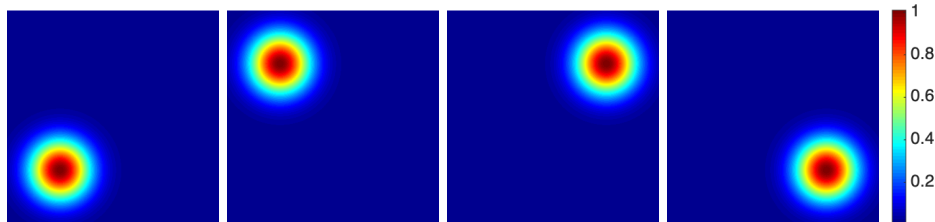


FIGURE 1. Radial basis functions used to define the permeability field.

Let  $\boldsymbol{\xi} = [\xi_1, \dots, \xi_{n_p}]^T$  be the weight vector, and the permeability field and model output depending on  $\boldsymbol{\xi}$  are denoted as  $a(\mathbf{x}; \boldsymbol{\xi})$  and  $u(\mathbf{x}; \boldsymbol{\xi})$ , respectively. Without loss of generality, here we consider the additive error, i.e.,

$$z_s = u(\mathbf{x}_s; \boldsymbol{\xi}) + e_s, \quad s = 1, \dots, n_z \tag{4}$$

where  $\{\mathbf{x}_s\}_{s=1}^{n_z}$  are sensor locations,  $\boldsymbol{\xi} \in \mathbb{R}^{n_p}$  is the weight vector corresponding to the data  $\mathbf{z} = [z_1, \dots, z_{n_z}]^T \in \mathbb{R}^{n_z}$ , and additive error  $\mathbf{e} = [e_1, \dots, e_{n_z}]^T \in \mathbb{R}^{n_z}$  comes from measurement error, etc. Then, the forward model can be written as

$$\mathbf{z} = \mathcal{G}(\boldsymbol{\xi}) + \mathbf{e}, \quad (5)$$

where  $\mathcal{G} : \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_z}$  is a mapping from input parameters  $\boldsymbol{\xi}$  to noise-free data. Let the noise  $\mathbf{e}$  and parameters  $\boldsymbol{\xi}$  are independent mutually. Suppose further that the components of  $\mathbf{e}$  are independent and identically distributed (i.i.d.), and satisfy  $e_i \sim N(0, \sigma_e^2)$  for  $i = 1, \dots, n_z$ .

**2.1. Bayesian framework.** In Bayesian inverse problems, the unknown parameters  $\boldsymbol{\xi}$  is regarded as random vector in a properly defined probability space  $(\Omega, \mathcal{F}, \mathcal{P})$ , whose components are i.i.d.. Moreover, the prior probability density function of  $\boldsymbol{\xi}$  is known. In this work, we only consider the uniform prior in  $[0, 1]^{n_p}$ .

By Bayes' rule, we can infer the posterior probability density function  $\pi(\boldsymbol{\xi}|\mathbf{z})$  of unknown parameters  $\boldsymbol{\xi}$  by combining the prior information  $\pi(\boldsymbol{\xi})$ , likelihood function  $\pi(\mathbf{z}|\boldsymbol{\xi})$  and observations  $\mathbf{z}$ , i.e.,

$$\pi(\boldsymbol{\xi}|\mathbf{z}) \propto \pi(\mathbf{z}|\boldsymbol{\xi})\pi(\boldsymbol{\xi}). \quad (6)$$

According the forward model (5) and the assumption of noise independence, likelihood function  $\pi(\mathbf{z}|\boldsymbol{\xi})$  has the form

$$\pi(\mathbf{z}|\boldsymbol{\xi}) = \prod_{i=1}^{n_z} \pi_{e_i}(z_i - \mathcal{G}_i(\boldsymbol{\xi})) \propto \exp\left(-\frac{\|\mathcal{G}(\boldsymbol{\xi}) - \mathbf{z}\|_2^2}{2\sigma_e^2}\right). \quad (7)$$

In many practical problems, the posterior distributions (6) are analytically intractable. Consequently, many sampling algorithms, e.g. MCMC, have been used to ascertain the posterior space. The framework of Bayesian inverse problems with direct sampling algorithm is shown in Figure 2. The form of the likelihood function (7) implies that for each candidate sample  $\boldsymbol{\xi}$ , we need to evaluate the corresponding forward model  $\mathcal{G}(\boldsymbol{\xi})$ , which is expensive generally. Therefore, the naive sampling methods are prohibitive. A series of reduced order models and effective sampling algorithms have been developed to deal with this problem. In order to reduce the total computational cost, the former is by reducing the cost of a single evaluation of the forward model, and the latter is by reducing the number of samples required. In current work, we concentrate on constructing a surrogate model to reduce the cost of a single forward model evaluation.

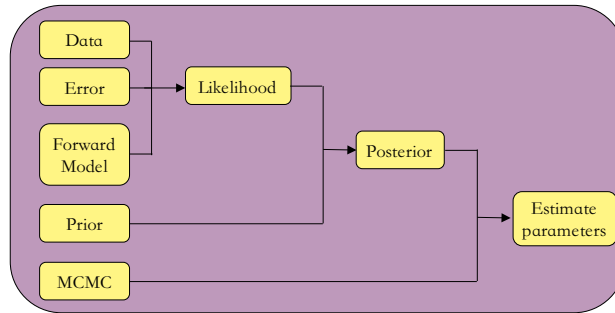


FIGURE 2. Bayesian inverse problems framework with direct sampling algorithm.

**2.2. Stochastic surrogate model.** In probability space  $(\Omega, \mathcal{F}, \mathcal{P})$ , we denote the probability density function of random variable  $\xi_i(\omega) : \Omega \rightarrow \mathbb{R}$  as  $\pi_i(\xi_i)$ , whose range is  $\Gamma_i = \xi_i(\Omega) \subset \mathbb{R}$ . Because the components of  $\boldsymbol{\xi}$  are assumed to be independent, the joint probability density function of random vector  $\boldsymbol{\xi}$  has the form

$$\pi(\boldsymbol{\xi}) = \prod_{i=1}^{n_p} \pi_i(\xi_i) \tag{8}$$

with support  $\Gamma = \prod_{i=1}^{n_p} \Gamma_i = [0, 1]^{n_p}$ . Here, the joint probability density function  $\pi(\boldsymbol{\xi})$  is equivalent to the prior density function in Bayes' formula (6).

For any  $\omega \in \Omega$ , the realization of random vector  $\boldsymbol{\xi}$  is belong to  $\Gamma$ . Therefore, in the Bayesian framework, we can restate the boundary value problem (2) as following stochastic forward problem in image probability space  $(\Gamma, \mathcal{B}(\Gamma), \pi(\boldsymbol{\xi})d\boldsymbol{\xi})$  rather than abstract probability space  $(\Omega, \mathcal{F}, \mathcal{P})$ :

$$\begin{cases} -\nabla \cdot (a(\mathbf{x}, \boldsymbol{\xi}) \nabla u(\mathbf{x}, \boldsymbol{\xi})) &= f, & (\mathbf{x}, \boldsymbol{\xi}) \in D \times \Gamma, \\ u(\mathbf{x}, \boldsymbol{\xi}) &= 0, & \mathbf{x} \in \partial D. \end{cases} \tag{9}$$

Here,  $\mathcal{B}(\Gamma)$  is the Borel  $\sigma$ -algebra on  $\Gamma$ , and  $\pi(\boldsymbol{\xi})d\boldsymbol{\xi}$  is the measure of random vector  $\boldsymbol{\xi}$ . In fact, we can construct explicit expression of the solution  $u(\mathbf{x}, \boldsymbol{\xi})$  to stochastic problem (9). Then for any  $\omega \in \Omega$ , substituting the realization  $\boldsymbol{\xi}(\omega)$  into  $u(\mathbf{x}, \boldsymbol{\xi})$ , the function value is consistent with the solution  $u(\mathbf{x}; \boldsymbol{\xi})$  of the deterministic problem (2) which has the same input  $\boldsymbol{\xi}$ . Using this solution expression, we can repeatedly evaluate the forward model (2) at a negligible cost. For this reason, the model (9) is called the stochastic surrogate model of problem (2).

**3. Data-driven compressive sensing method based on proper orthogonal decomposition.** In this section, we propose a data-driven compressive sensing method based on proper orthogonal decomposition, which can be used to construct efficient and sparse solution for a stochastic forward model.

It's well-known that the solution  $u(\mathbf{x}, \boldsymbol{\xi}) : D \times \Gamma \rightarrow \mathbb{R}$  of stochastic forward model (9) can be represented by stochastic and deterministic basis functions. Here, we propose the POD-DCS method, which chooses a set of data-driven POD basis functions  $\{\phi_j(\mathbf{x})\}_{j=1}^m$  for domain  $D$  and a set of gPC basis functions  $\{\psi_i(\boldsymbol{\xi})\}_{i=1}^n$  for prior support  $\Gamma$ , and the coefficients are determined by solving a BP problem. Next, we discuss how to construct the POD-DCS approximate solution for a stochastic forward problem.

**3.1. Proper orthogonal decomposition.** The corresponding degree of freedom of the finite element method is expressed by  $n_u$ , and let  $S \in \mathbb{R}^{n_u \times K_s}$  be the snapshot matrix:

$$S = [u^h(\mathbf{x}, \boldsymbol{\xi}_1), \dots, u^h(\mathbf{x}, \boldsymbol{\xi}_{K_s})] := [u_1^h(\mathbf{x}), \dots, u_{K_s}^h(\mathbf{x})], \tag{10}$$

where  $u^h(\mathbf{x}, \boldsymbol{\xi}_i) \in \mathbb{R}^{n_u}$  is the finite element solution of problem (9) associated with input parameter  $\boldsymbol{\xi}_i \in \Gamma$ . These random inputs  $\{\boldsymbol{\xi}_i\}_{i=1}^{K_s}$  are sampling from  $\Gamma$  randomly. By solving eigenvalue problem

$$RV = V\Lambda \tag{11}$$

with

$$R_{ij} = \frac{1}{K_s} (u_i^h(\mathbf{x}), u_j^h(\mathbf{x})), \quad i, j = 1, \dots, K_s, \tag{12}$$

we can obtain the diagonal singular value matrix  $\Lambda$  with  $\Lambda_{jj} = \lambda_j$ , and the matrix of singular vectors  $V = [\mathbf{v}_1, \dots, \mathbf{v}_{K_s}]$  with the  $j$ -th component

$$\mathbf{v}_j = [v_1^{(j)}, \dots, v_{K_s}^{(j)}]^T, \quad j = 1, \dots, K_s. \quad (13)$$

Since  $R$  is positive semidefinite and symmetric, we can make the eigenvectors  $\{\mathbf{v}_j\}_{j=1}^{K_s}$  to be orthogonal and sort the eigenvalues in a descending order. After normalization, the POD basis can be represented as

$$\phi_j(\mathbf{x}) = \frac{1}{\sqrt{K_s \lambda_j}} \sum_{i=1}^{K_s} v_i^{(j)} u_i^h(\mathbf{x}), \quad j = 1, \dots, m. \quad (14)$$

Here, the dimension of POD basis,  $m$ , is determined by the energy ratio  $\nu_m$  defined as

$$\nu_m = \sum_{j=1}^m \lambda_j / \sum_{j=1}^{K_s} \lambda_j. \quad (15)$$

Once the basis functions are computed, the solution  $u(\mathbf{x}, \boldsymbol{\xi})$  can be approximated as

$$\widehat{u}(\mathbf{x}, \boldsymbol{\xi}) = \sum_{j=1}^m \alpha_j(\boldsymbol{\xi}) \phi_j(\mathbf{x}). \quad (16)$$

The coefficients  $\{\alpha_j\}_{j=1}^m$  are called reduced states and satisfy the following algebraic system

$$A\boldsymbol{\alpha} = \mathbf{F}, \quad (17)$$

where

$$\begin{aligned} \boldsymbol{\alpha} &= [\alpha_1(\boldsymbol{\xi}), \dots, \alpha_m(\boldsymbol{\xi})]^T, \\ \mathbf{F}_i &= \int_D f(\mathbf{x}) \phi_i(\mathbf{x}) d\mathbf{x}, \quad i = 1, \dots, m, \\ A_{ij} &= \int_D a(\mathbf{x}, \boldsymbol{\xi}) \nabla \phi_j(\mathbf{x}) \cdot \nabla \phi_i(\mathbf{x}) d\mathbf{x}, \quad i, j = 1, \dots, m. \end{aligned}$$

From the definition of the POD basis functions we know that  $\phi_i(\mathbf{x})$  and  $\phi_j(\mathbf{x})$  are orthogonal for  $i, j = 1, \dots, m$ , and  $\{\phi_j(\mathbf{x})\}_{j=1}^m$  minimize the error measure

$$\mathcal{M} = \frac{1}{K_s} \sum_{i=1}^{K_s} \left\| u_i^h(\mathbf{x}) - \sum_{j=1}^m (u_i^h(\mathbf{x}), \phi_j(\mathbf{x})) \phi_j(\mathbf{x}) \right\|_{L^2(D)}^2 \quad (18)$$

with minimum  $\sum_{j=m+1}^{K_s} \lambda_j$  [12]. When  $K_s$  is large enough,  $\mathcal{M}$  can be regarded as an approximation of the mean square error in the  $L^2(D)$ -norm

$$\overline{\mathcal{M}} = \mathbb{E} \left[ \left\| u^h(\mathbf{x}, \boldsymbol{\xi}) - \sum_{j=1}^m (u^h(\mathbf{x}, \boldsymbol{\xi}), \phi_j(\mathbf{x})) \phi_j(\mathbf{x}) \right\|_{L^2(D)}^2 \right]. \quad (19)$$

**3.1.1. Simulation results using POD-based stochastic ROM.** For 4-dimensional stochastic problem (9), we draw  $K_s$  samples of input parameters  $\boldsymbol{\xi}$  from prior support  $\Gamma$  randomly, and evaluate the corresponding forward model (9) by using finite element method with mesh size  $h = 2^{-6}$  (i.e.  $n_u = 4225$ ). These solutions compose the snapshot matrix  $S$ . For  $K_s = 50, 100, 200$  and  $300$ , the first 15 eigenvalues of correlation matrix  $R$  are shown on the left of Figure 3. We can observe that the eigenvalues have very little differences when  $K_s \geq 100$ , but have large gaps between  $K_s = 50$  and  $K_s = 100$ . In order to explore the prior space fully and ensure the

accuracy of the POD-based reduced order model, we need use  $K_s \geq 100$  realizations to construct the snapshot matrix. The right side of the Figure 3 shows the cumulative energy ratios of the first 15 POD basis functions with  $K_s = 100, 200$  and  $300$ . We can see that the energy ratios of first 2 POD basis functions have slight differences for different  $K_s$ . For each case, the first 3 basis functions can capture 99% energy of snapshot set, and the cumulative energy ratio of the first 9 basis functions is very close to 1.

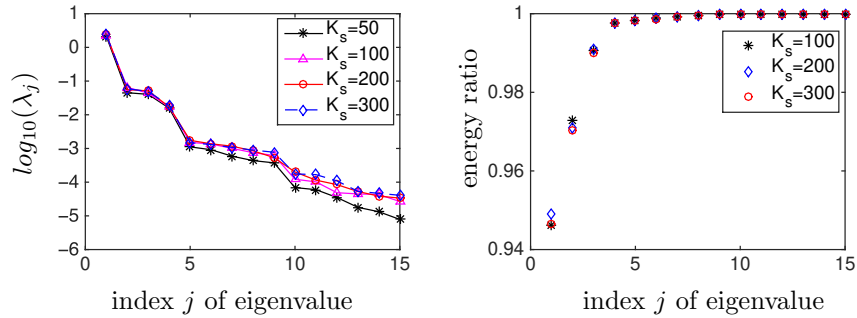


FIGURE 3. (Left): The first 15 eigenvalues of correlation matrix  $R$  for  $K_s = 50, 100, 200$  and  $300$ . (Right): The cumulative energy ratios of the first 15 POD basis functions with  $K_s = 100, 200$  and  $300$ .

We define the expectation and variance of  $L^2(D)$ -norm error between the full finite element solution  $u^h(\mathbf{x}, \boldsymbol{\xi})$  and the POD-based approximate solution  $\hat{u}(\mathbf{x}, \boldsymbol{\xi})$  as

$$\hat{\mathcal{E}} = \mathbb{E} \left[ \left\| u^h - \hat{u} \right\|_{L^2(D)}^2 \right], \tag{20}$$

and

$$\hat{\mathcal{V}} = \text{Var} \left[ \left\| u^h - \hat{u} \right\|_{L^2(D)}^2 \right], \tag{21}$$

where  $\mathbb{E}[\cdot]$  means the expectation and  $\text{Var}[\cdot]$  represents the variance. Here, we approximate these statistics with  $K_r = 500$  samples.

Figure 4 displays these error estimates of the POD-based approximate solution associated with different  $K_s$  and different degrees of freedom  $m$ . We observe that for a given  $m$ , the expectations and variances in these three cases only have a little differences. And when  $m \leq 4$ , the errors decrease very quickly, while they are very small variations with  $m \geq 9$ . The results for  $m = 3, 6, 9$  and  $12$  are shown in Table 1. It is clear that the change in energy from 9 to 12 POD basis functions is so small that it is almost negligible, and the improvements in the corresponding error statistics are also small. Moreover, for  $m = 9$ , the order of magnitude of the differences in  $\hat{\mathcal{E}}$  is -4, and in variance  $\hat{\mathcal{V}}$  is -5 between these three cases, while they are larger for  $m = 3$  and  $m = 6$ . Therefore, we consider using  $K_s = 100$  realizations to generate the POD basis functions and choose  $m$  to be 9 in the next experiment to ensure accuracy without consuming too much computing resources.

**3.2. Compressive sensing method.** In our POD-DCS algorithm, using the POD method first can greatly reduce the DoF of the problem in hand, and simplify the subsequent processing. Next, we utilize the idea of compressive sensing to express the reduced states  $\{\alpha_j(\boldsymbol{\xi})\}_{j=1}^m$  with gPC basis functions, and determine the

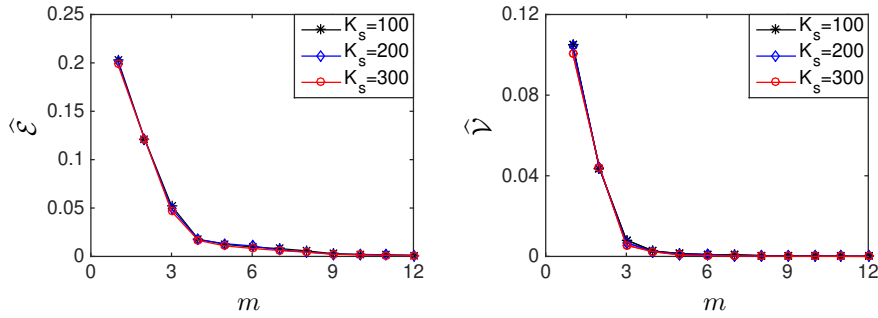


FIGURE 4. The statistics estimation of  $L^2(D)$ -norm error between the full finite element solution and the POD-based ROM solution with  $K_s = 100, 200$  and  $300$ . (Left): Expectation. (Right): Variance.

TABLE 1. The statistics estimation of  $L^2(D)$ -norm error between the finite element solution and the POD-based approximate solution with  $K_s = 100, 200, 300$  and  $m = 3, 6, 9, 12$ .

		$m$	3	6	9	12
$K_s = 100$	$\nu_m$		0.9906	0.9988	0.9998	0.9999
	$\widehat{\mathcal{E}} \times 10^{-2}$		5.2424	0.9679	0.2840	0.1267
	$\widehat{\mathcal{V}} \times 10^{-3}$		8.1367	0.8915	0.1475	0.0282
$K_s = 200$	$\nu_m$		0.9908	0.9988	0.9998	0.9999
	$\widehat{\mathcal{E}} \times 10^{-2}$		4.9558	1.0836	0.2422	0.0907
	$\widehat{\mathcal{V}} \times 10^{-3}$		6.7935	0.9507	0.1023	0.0126
$K_s = 300$	$\nu_m$		0.9900	0.9986	0.9997	0.9999
	$\widehat{\mathcal{E}} \times 10^{-2}$		4.6537	0.8260	0.2229	0.0622
	$\widehat{\mathcal{V}} \times 10^{-3}$		5.5292	0.4418	0.0898	0.0037

expression coefficients by using a small amount of data and solving a  $l_1$ -minimization problem.

Clearly, once the POD basis functions  $\{\phi_j(\mathbf{x})\}_{j=1}^m$  are obtained, we only need to compute the reduced states  $\alpha_j(\boldsymbol{\xi}) : \Gamma \rightarrow \mathbb{R}$  in the expression (16) for  $j = 1, \dots, m$ . Here, we use the  $N$ th-order generalized polynomial chaos to represent the reduced states as the form

$$\alpha_j(\boldsymbol{\xi}) \approx \sum_{i=1}^n c_{ij} \psi_i(\boldsymbol{\xi}), \quad j = 1, \dots, m, \tag{22}$$

where  $\{\psi_i(\boldsymbol{\xi})\}_{i=1}^n$  represents a set of multivariate Legendre polynomials with order  $N$  that are orthonormal w.r.t. the joint probability density function  $\pi(\boldsymbol{\xi})$ , and  $\psi_1(\boldsymbol{\xi}) = 1$ . The cardinality of the stochastic polynomial basis is

$$n = \frac{(N + n_p)!}{N!n_p!}. \tag{23}$$

The stochastic collocation method [29] based on polynomial interpolation can be utilized to determine the coefficients  $\{c_{ij}\}_{i=1, j=1}^{n, m}$ . With collocation points  $\{\boldsymbol{\xi}_i\}_{i=1}^{K_c}$ ,



we have a linear system

$$\begin{pmatrix} \alpha_1(\boldsymbol{\xi}_1) & \cdots & \alpha_m(\boldsymbol{\xi}_1) \\ \vdots & & \vdots \\ \alpha_1(\boldsymbol{\xi}_{K_c}) & \cdots & \alpha_m(\boldsymbol{\xi}_{K_c}) \end{pmatrix} = \begin{pmatrix} \psi_1(\boldsymbol{\xi}_1) & \cdots & \psi_n(\boldsymbol{\xi}_1) \\ \vdots & & \vdots \\ \psi_1(\boldsymbol{\xi}_{K_c}) & \cdots & \psi_n(\boldsymbol{\xi}_{K_c}) \end{pmatrix} \begin{pmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nm} \end{pmatrix}.$$

Simply expressed as

$$\widehat{A} = \Psi \mathbf{c}, \tag{24}$$

where  $\widehat{A} = (\alpha_j(\boldsymbol{\xi}_i)) \in \mathbb{R}^{K_c \times m}$  is the matrix of reduced state generated by solving system (17),  $\mathbf{c} = (c_{ij}) \in \mathbb{R}^{n \times m}$  is the coefficient matrix that needs to be determined,  $\Psi = (\psi_i(\boldsymbol{\xi}_j)) \in \mathbb{R}^{K_c \times n}$  represents the stochastic information matrix, and  $K_c$  is the number of stochastic collocation points. As shown in Equation (23), with the increase of random input dimension  $n_p$ , the cardinality  $n$  of polynomial basis functions grows very quickly, which will cause linear system (24) to be underdetermined, i.e.,  $K_c < n$ .

Based on the idea of compressive sensing, given a highly incomplete set of reduced states by solving algebraic system (17), an accurate approximate solution of linear system (24) can be obtained by solving the BP problem

$$\underset{\text{vec}(\mathbf{c})}{\text{vec}(\tilde{\mathbf{c}})} = \arg \min \|\text{vec}(\mathbf{c})\|_1, \quad \text{subject to} \quad \text{vec}(\widehat{A}) = \Theta \text{vec}(\mathbf{c}), \tag{25}$$

where dictionary matrix  $\Theta = I_m \otimes \Psi$  is the Kronecker product of an  $m \times m$  identity matrix  $I_m$  with stochastic information matrix  $\Psi$ ,  $\text{vec}(\widehat{A})$  denotes the vectorization of matrix  $\widehat{A}$  by column, and  $\tilde{\mathbf{c}} = (\tilde{c}_{ij})$  represents the coefficient matrix we calculated. Then the reduced states can be approximated by

$$\tilde{\alpha}_j(\boldsymbol{\xi}) = \sum_{i=1}^n \tilde{c}_{ij} \psi_i(\boldsymbol{\xi}), \quad j = 1, \dots, m. \tag{26}$$

Combining equations (16) and (26), we can get the POD-DCS approximate solution of stochastic elliptic problem (9) as

$$\tilde{u}(\mathbf{x}, \boldsymbol{\xi}) = \sum_{i=1}^n \sum_{j=1}^m \tilde{c}_{ij} \psi_i(\boldsymbol{\xi}) \phi_j(\mathbf{x}). \tag{27}$$

**Remark 1.** By using the orthogonality of multivariate Legendre polynomials, we have

$$\mathbb{E} [\tilde{u}(\mathbf{x}, \boldsymbol{\xi})] = \sum_{j=1}^m \tilde{c}_{1j} \phi_j(\mathbf{x}), \tag{28}$$

$$\mathbb{E} [\tilde{u}^2(\mathbf{x}, \boldsymbol{\xi})] = \sum_{i=1}^n \left( \sum_{j=1}^m \tilde{c}_{ij} \phi_j(\mathbf{x}) \right) \left( \sum_{j'=1}^m \tilde{c}_{ij'} \phi_{j'}(\mathbf{x}) \right). \tag{29}$$

The details of our POD-DCS method for stochastic forward problem is presented as following Algorithm 1.

3.2.1. *Simulation results of the POD-DCS method.* Table 1 illustrates that using 100 realizations to generate the POD-based approximate solution with dimension 9 can guarantee the accuracy of the model. Thus, we utilize 100 snapshots to construct ROM, and retain only the first 9 basis functions. In this case,  $\hat{\mathcal{E}} = 2.8400 \times 10^{-3}$  and  $\hat{\mathcal{V}} = 1.4750 \times 10^{-4}$ .

---

**Algorithm 1** The POD-DCS algorithm for stochastic forward problem

---

**Input:** Grid parameter  $h$ , positive integer  $K_s$ , energy ratio  $\nu$ , collocation points  $\{\boldsymbol{\xi}_i\}_{i=1}^{K_c}$ , the order of polynomial  $N$ .

**Output:** POD basis function  $\{\phi_j(\mathbf{x})\}_{j=1}^m$ , and the coefficient matrix  $\tilde{\mathbf{c}}$ .

- 1: Draw a set of random inputs  $\{\boldsymbol{\xi}_k\}_{k=1}^{K_s}$  from prior support  $\Gamma$  independently, and compute the corresponding solutions  $\{u^h(\mathbf{x}, \boldsymbol{\xi}_k)\}_{k=1}^{K_s}$  of problem (9) by using FEM with mesh size  $h$ , then compose a snapshot matrix  $S$ .
- 2: Generate the POD basis functions  $\{\phi_j(\mathbf{x})\}_{j=1}^m$  as (14) by solving eigenvalue problem (11), where  $m$  is the smallest positive integer such that  $\nu_m \geq \nu$  holds.
- 3: Construct the reduced state matrix  $\tilde{A}$  as (24) by solving algebraic system (17) with collocation points  $\{\boldsymbol{\xi}_i\}_{i=1}^{K_c}$ .
- 4: Select appropriate stochastic basis functions  $\{\psi_i(\boldsymbol{\xi})\}_{i=1}^n$  with order  $N$ .
- 5: Obtain coefficient matrix  $\tilde{\mathbf{c}}$  by solving BP problem (25).
- 6: Generate the approximate reduced states  $(\tilde{\alpha}_j(\boldsymbol{\xi}))$  and the POD-DCS approximate solution  $\tilde{u}(\mathbf{x}, \boldsymbol{\xi})$  of system (9) as

$$\tilde{\alpha}_j(\boldsymbol{\xi}) = \sum_{i=1}^n \tilde{c}_{ij} \psi_i(\boldsymbol{\xi}), \quad j = 1, \dots, m,$$

$$\tilde{u}(\mathbf{x}, \boldsymbol{\xi}) = \sum_{i=1}^n \sum_{j=1}^m \tilde{c}_{ij} \psi_i(\boldsymbol{\xi}) \phi_j(\mathbf{x}).$$


---

Now, we need to determine the coefficients  $\{\tilde{c}_{ij}\}_{i=1, j=1}^{n, m}$  in the expression (26) by solving  $l_1$ -minimization problem (25). Here, the 7th-order Legendre PC basis functions  $\psi(\boldsymbol{\xi})$  are used to represent the reduced states  $\alpha_j(\boldsymbol{\xi})$  for  $j = 1, \dots, m$ , and the total DoF of matrix  $\tilde{\mathbf{c}}$  is  $n \times m = 2970$ . Without loss of generality, we only consider the coefficients with absolute values larger than a fixed threshold  $\tau$ , i.e., set the entries of  $\tilde{\mathbf{c}}$  whose absolute value smaller than the threshold  $\tau$  to 0, and let  $\tilde{\mathcal{R}}_\tau$  be the proportion of non-zero coefficients in matrix  $\tilde{\mathbf{c}}$ , that is

$$\tilde{\mathcal{R}}_\tau = \frac{\#\{\tilde{\mathbf{c}} > \tau\}}{\#\{\tilde{\mathbf{c}}\}}. \quad (30)$$

Like the definitions of  $\hat{\mathcal{E}}$  and  $\hat{\mathcal{V}}$ , we denote the expectation and variance of  $L^2(D)$ -norm error between the full finite element solution  $u^h(\mathbf{x}, \boldsymbol{\xi})$  and the POD-DCS approximate solution  $\tilde{u}(\mathbf{x}, \boldsymbol{\xi})$  as  $\tilde{\mathcal{E}}$  and  $\tilde{\mathcal{V}}$  respectively, namely

$$\tilde{\mathcal{E}} = \mathbb{E} \left[ \left\| u^h - \tilde{u} \right\|_{L^2(D)}^2 \right], \quad (31)$$

and

$$\tilde{\mathcal{V}} = \text{Var} \left[ \left\| u^h - \tilde{u} \right\|_{L^2(D)}^2 \right]. \quad (32)$$

The sparsity and error estimation of the POD-DCS method w.r.t. different number of collocation points sampled from the prior space randomly and different thresholds are drawn in Figure 5. Compared with  $\tau = 0$ , for each  $K_c$ , the ratio of non-zero coefficients is less than 5% with  $\tau = 0.1$ , but the mean of the  $L^2(D)$ -norm error is much larger. This is due to the fact that the threshold is so large that some influential basis functions are ignored. For  $\tau = 0.01$ , the highest proportion of non-zero coefficients in matrix  $\tilde{\mathbf{c}}$  is about 33%, and the error statistics are almost

consistent with those of  $\tau = 0$ . Therefore, we can take the threshold  $\tau = 0.01$  to ensure the accuracy. Through regression, the expectation of our approach converges with order 1.58 and variance converges with order 2.97, that is,  $\tilde{\mathcal{E}} \approx \mathcal{O}(K_c^{-3/2})$  and  $\tilde{\mathcal{V}} \approx \mathcal{O}(K_c^{-3})$ , which can be seen in Figure 6. Such a highly “incomplete” set of reduced states does contain enough information to accurately recover the reduced order model. When  $K_c \geq 200$ , the expectations  $\tilde{\mathcal{E}}$  and variances  $\tilde{\mathcal{V}}$  of  $L^2(D)$ -norm error do not change much as  $K_c$  increase. So here we use 200 collocation points to determine the coefficients  $(\tilde{c}_{ij})$  of expression (27), that is, we need to calculate 200 algebraic systems (17). At this time,  $\tilde{\mathcal{R}}_{0.01} = 20.24\%$ ,  $\tilde{\mathcal{E}} = 4.1758 \times 10^{-3}$ , and  $\tilde{\mathcal{V}} = 2.4931 \times 10^{-4}$ .

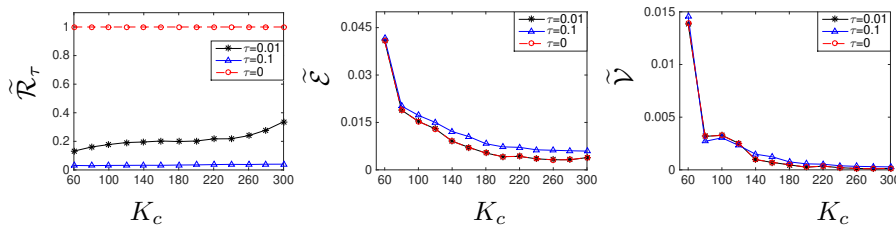


FIGURE 5. Sparsity and error estimations of the POD-DCS method with different  $K_c$  and different threshold  $\tau$ . (Left): The proportion  $\tilde{\mathcal{R}}_\tau$  of non-zero coefficients in matrix  $\tilde{\mathbf{c}}$ ; (Middle): The expectation  $\tilde{\mathcal{E}}$  of  $L^2(D)$ -norm error; (Right): The variance  $\tilde{\mathcal{V}}$  of  $L^2(D)$ -norm error.

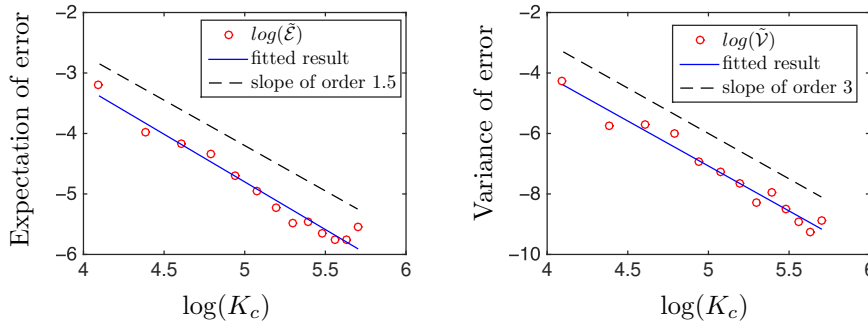


FIGURE 6. Error estimations of the POD-DCS method w.r.t. different collocation point sizes and  $\tau = 0.01$ .

**4. Analysis.** The POD-DCS method has been described in the previous section. In order to introduce the conclusions of error analysis and coefficients estimation of our scheme, we first introduce several relevant properties of the compressive sensing method in this section.

**4.1. Error analysis.**

**Definition 4.1** (see [11]). A vector  $\mathbf{b} \in \mathbb{R}^n$  is called  $k$ -sparse if  $\|\mathbf{b}\|_0 \leq k$  holds, and its best  $k$ -term approximation error in the  $l_p$ -norm is defined as

$$\sigma_{k,p}(\mathbf{b}) = \inf_{\|\mathbf{x}\|_0 \leq k} \|\mathbf{x} - \mathbf{b}\|_p, \tag{33}$$

where  $\|\mathbf{b}\|_0$  is the number of non-zero terms in vector  $\mathbf{b}$ .

Note that for  $0 < q < p \leq \infty$  and set  $s = 1/q - 1/p > 0$ , the prior estimation holds:

$$\sigma_{k,p}(\mathbf{b}) \leq k^{-s} \|\mathbf{b}\|_q. \quad (34)$$

In order to ensure that the matrix  $\mathbf{c}$  can be reconstructed exactly by using  $l_1$  minimization, we introduce the restricted isometry property (RIP).

**Lemma 4.2** (see [5, 24]). *If for any  $k$ -sparse vector  $\mathbf{b} \in \mathbb{R}^n$ , there exists a constant  $\delta \in (0, 1)$  such that the inequality*

$$(1 - \delta) \|\mathbf{b}\|_2^2 \leq \|\mathbf{A}\mathbf{b}\|_2^2 \leq (1 + \delta) \|\mathbf{b}\|_2^2 \quad (35)$$

holds. Then  $\delta_k(A) := \min(\delta)$  is called the restricted isometry constant (RIC) of matrix  $A \in \mathbb{R}^{m \times n}$ , and the matrix  $A$  is said to satisfy the RIP of order  $k$  with RIC  $\delta_k(A)$ . Similarly, denoting the RIC of matrix  $B \in \mathbb{R}^{p \times q}$  by  $\delta_k(B)$ , we can obtain the inequality

$$\max\{\delta_k(A), \delta_k(B)\} \leq \delta_k(A \otimes B) \leq \delta_k(A) + \delta_k(B) + \delta_k(A)\delta_k(B). \quad (36)$$

Based on Lemma 4.2 we know that the RIC of dictionary matrix  $\Theta$  in (25) satisfies  $\delta_k(\Theta) = \delta_k(\Psi)$ .

**Lemma 4.3** (see [11]). *Assuming that  $\delta_{3k}(\Theta) < 1/3$ , then the solution of  $l_1$ -minimization problem (25) satisfies*

$$\|\text{vec}(\mathbf{c}) - \text{vec}(\tilde{\mathbf{c}})\|_2 \leq C_\delta \frac{\sigma_{k,1}(\text{vec}(\mathbf{c}))}{\sqrt{k}} \quad (37)$$

where the constant  $C_\delta > 0$  depends only on  $\delta_{3k}(\Theta)$ .

By Lemma 4.3, we give the error estimation of our POD-DCS method as following.

**Theorem 4.4.** *There exist constants  $C_1, C_2, C_3, C_4, \theta > 0$  and  $0 < q < 1$ , assuming that the dictionary matrix  $\Theta$  in (25) satisfies  $\delta_{3k}(\Theta) < 1/3$ . Then with probability close to one, the expectation (31) of  $L^2(D)$ -norm error between the finite element solution and the POD-DCS approximate solution satisfies*

$$\tilde{\mathcal{E}} \leq C_1 \sqrt{\frac{\hat{\mathcal{V}}}{K_s}} + C_2 \sum_{j=m+1}^{K_s} \lambda_j + C_3 N^{-\theta} + C_4 k^{1-\frac{2}{q}} \|\text{vec}(\mathbf{c})\|_q^2, \quad (38)$$

where  $C_3, C_4$  depend on the smoothness of the reduced states and  $\delta_k(\Theta)$  respectively, while constants  $C_1, C_2$  are universal.

*Proof.* By using the inequality  $2ab \leq a^2 + b^2$ , the error (31) follows that

$$\tilde{\mathcal{E}} = \mathbb{E} \left[ \|u^h - \hat{u} + \hat{u} - \tilde{u}\|_{L^2(D)}^2 \right] \leq 2 \left\{ \mathbb{E} \left[ \|u^h - \hat{u}\|_{L^2(D)}^2 \right] + \mathbb{E} \left[ \|\hat{u} - \tilde{u}\|_{L^2(D)}^2 \right] \right\}.$$

Denote

$$I_1 = \mathbb{E} \left[ \|u^h - \hat{u}\|_{L^2(D)}^2 \right], \quad I_2 = \mathbb{E} \left[ \|\hat{u} - \tilde{u}\|_{L^2(D)}^2 \right].$$

We first estimate  $I_1$ . The standard Monte Carlo Finite Element Method (MCFEM) approximates the expectation function  $I_1$  by the average of samples. Note that the random inputs  $\{\xi_j\}_{j=1}^{K_s}$  used to generate the snapshot set are sampling

from prior support  $\Gamma$  randomly. Thus, we can use these samples to approximate  $I_1$  as

$$\bar{I}_1 = \frac{1}{K_s} \sum_{j=1}^{K_s} \|u_j^h - \hat{u}_j\|_{L^2(D)}^2.$$

where  $u_j^h$  and  $\hat{u}_j$  represent the finite element solution and POD-based solution with random input  $\boldsymbol{\xi}_j$  for  $j = 1, \dots, K_s$ , respectively. The number of realizations,  $K_s$ , controls the statistical error  $\mathcal{E}_S = I_1 - \bar{I}_1$  [1], and

$$\begin{aligned} \mathcal{E}_S &= \mathbb{E} \left[ \|u^h - \hat{u}\|_{L^2(D)}^2 \right] - \frac{1}{K_s} \sum_{j=1}^{K_s} \|u_j^h - \hat{u}_j\|_{L^2(D)}^2 \\ &= \sqrt{\frac{\text{Var} \left[ \|u^h - \hat{u}\|_{L^2(D)}^2 \right]}{K_s}} \times \frac{K_s \mathbb{E} \left[ \|u^h - \hat{u}\|_{L^2(D)}^2 \right] - \sum_{j=1}^{K_s} \|u_j^h - \hat{u}_j\|_{L^2(D)}^2}{\sqrt{K_s \text{Var} \left[ \|u^h - \hat{u}\|_{L^2(D)}^2 \right]}}. \end{aligned}$$

According to the central limit theorem we have

$$\mathcal{E}_S \sim N \left( 0, \hat{V}/K_s \right),$$

then choose a constant  $C_q \geq 1.65$ , such that

$$|\mathcal{E}_S| \leq C_q \sqrt{\hat{V}/K_s},$$

holds with probability close to one, where  $C_q$  is called quantile. With the error measure (18),  $\bar{I}_1$  satisfies

$$\bar{I}_1 = \frac{1}{K_s} \sum_{j=1}^{K_s} \|u_j^h - \hat{u}_j\|_{L^2(D)}^2 = \sum_{j=m+1}^{K_s} \lambda_j.$$

As (22), we use multivariate polynomials with order  $N$  to approximate reduced state  $\alpha_j(\boldsymbol{\xi})$  for  $j = 1, \dots, m$ , then there exist constants  $C_N, \theta > 0$  such that

$$\int_{\Gamma} \left( \alpha_j(\boldsymbol{\xi}) - \sum_{i=1}^n c_{ij} \psi_i(\boldsymbol{\xi}) \right)^2 \pi(\boldsymbol{\xi}) d\boldsymbol{\xi} \leq C_N N^{-\theta}, \quad j = 1, \dots, m,$$

holds. The constant  $C_N$  depends on the smoothness of  $\{\alpha_j(\boldsymbol{\xi})\}_{j=1}^m$ . Using the biorthogonality of the POD-DCS solution and Lemma 4.3, we have

$$\begin{aligned} I_2 &= \int_{\Gamma} \int_D \left( \sum_{j=1}^m (\alpha_j(\boldsymbol{\xi}) - \tilde{\alpha}_j(\boldsymbol{\xi})) \phi_j(\mathbf{x}) \right)^2 \pi(\boldsymbol{\xi}) dx d\boldsymbol{\xi} \\ &= \sum_{j=1}^m \int_{\Gamma} (\alpha_j(\boldsymbol{\xi}) - \tilde{\alpha}_j(\boldsymbol{\xi}))^2 \pi(\boldsymbol{\xi}) d\boldsymbol{\xi} \\ &= \sum_{j=1}^m \int_{\Gamma} \left( \alpha_j(\boldsymbol{\xi}) - \sum_{i=1}^n c_{ij} \psi_i(\boldsymbol{\xi}) + \sum_{i=1}^n (c_{ij} - \tilde{c}_{ij}) \psi_i(\boldsymbol{\xi}) \right)^2 \pi(\boldsymbol{\xi}) d\boldsymbol{\xi} \\ &\leq 2 \sum_{j=1}^m \int_{\Gamma} \left( \alpha_j(\boldsymbol{\xi}) - \sum_{i=1}^n c_{ij} \psi_i(\boldsymbol{\xi}) \right)^2 \pi(\boldsymbol{\xi}) d\boldsymbol{\xi} + 2 \sum_{j=1}^m \sum_{i=1}^n (c_{ij} - \tilde{c}_{ij})^2 \end{aligned}$$

$$\leq 2mC_N N^{-\theta} + 2C_\delta^2 \frac{\sigma_{k,1}^2(\text{vec}(\mathbf{c}))}{k}.$$

By the prior estimation (34), we can arrive the error estimate with  $0 < q < 1$

$$\begin{aligned} \tilde{\mathcal{E}} &\leq 2 \left( C_q \sqrt{\frac{\hat{\mathcal{V}}}{K_s}} + \sum_{j=m+1}^{K_s} \lambda_j + 2mC_N N^{-\theta} + 2C_\delta^2 \frac{\sigma_{k,1}^2(\text{vec}(\mathbf{c}))}{k} \right) \\ &\leq C_1 \sqrt{\frac{\hat{\mathcal{V}}}{K_s}} + C_2 \sum_{j=m+1}^{K_s} \lambda_j + C_3 N^{-\theta} + C_4 k^{1-\frac{2}{q}} \|\text{vec}(\mathbf{c})\|_q^2, \end{aligned}$$

which completes the proof. □

This theorem implies that the mean square error  $\tilde{\mathcal{E}}$  consists of four parts, including statistical error, truncation error, polynomial approximation error and sparse reconstruction error. Among them, the first two items are derived from the POD-based reduced order model, and the last two items are caused by the sparse reconstruction of the corresponding reduced order model. Therefore, we can balance these four errors and discrete error to save computing resources for a desired accuracy.

**4.2. Sparsity.** The error analysis has been completed and we are now ready to show the sparsity of our POD-DCS solution.

**Theorem 4.5.** *The coefficient matrix  $\tilde{\mathbf{c}}$  in the POD-DCS solution expression (27) satisfies the estimation*

$$\sum_{i=1}^n \tilde{c}_{ij}^2 = \lambda_j, \quad j = 1, \dots, m. \tag{39}$$

*Proof.* According to the eigenvalue problem (11) and the POD basis function (14) we obtain

$$\begin{aligned} \lambda_j \phi_j(\mathbf{x}) &= \frac{1}{\sqrt{K_s \lambda_j}} \sum_{i=1}^{K_s} \lambda_j v_i^{(j)} u_i^h(\mathbf{x}) \\ &= \frac{1}{\sqrt{K_s \lambda_j}} \sum_{i=1}^{K_s} \left( \frac{1}{K_s} \sum_{r=1}^{K_s} (u_i^h(\mathbf{x}), u_r^h(\mathbf{x})) v_r^{(j)} u_i^h(\mathbf{x}) \right) \\ &= \frac{1}{K_s} \sum_{i=1}^{K_s} (u_i^h(\mathbf{x}), \phi_j(\mathbf{x})) u_i^h(\mathbf{x}). \end{aligned}$$

Thus, we know that  $\lambda_j \phi_j(\mathbf{x})$  can be regarded as an approximation of the expectation of  $(u^h(\mathbf{x}, \boldsymbol{\xi}), \phi_j(\mathbf{x})) u^h(\mathbf{x}, \boldsymbol{\xi})$ . Using the POD-DCS solution to approximate finite element solution, and combining the biorthogonality of the POD-DCS solution,  $\lambda_j \phi_j(\mathbf{x})$  can be approximated by

$$\begin{aligned} &\mathbb{E}[(\tilde{u}(\mathbf{x}, \boldsymbol{\xi}), \phi_j(\mathbf{x})) \tilde{u}(\mathbf{x}, \boldsymbol{\xi})] \\ &= \int_{\Gamma} \left( \int_D \sum_{i=1}^n \sum_{k=1}^m \tilde{c}_{ik} \psi_i(\boldsymbol{\xi}) \phi_k(\mathbf{x}) \phi_j(\mathbf{x}) d\mathbf{x} \right) \sum_{i'=1}^n \sum_{k'=1}^m \tilde{c}_{i'k'} \psi_{i'}(\boldsymbol{\xi}) \phi_{k'}(\mathbf{x}) \pi(\boldsymbol{\xi}) d\boldsymbol{\xi} \end{aligned}$$

$$\begin{aligned}
 &= \int_{\Gamma} \left( \sum_{i=1}^n \tilde{c}_{ij} \psi_i(\boldsymbol{\xi}) \right) \sum_{i'=1}^n \sum_{k'=1}^m \tilde{c}_{i'k'} \psi_{i'}(\boldsymbol{\xi}) \phi_{k'}(\mathbf{x}) \pi(\boldsymbol{\xi}) d\boldsymbol{\xi} \\
 &= \sum_{k'=1}^m \left( \sum_{i=1}^n \tilde{c}_{ij} \tilde{c}_{ik'} \right) \phi_{k'}(\mathbf{x}),
 \end{aligned}$$

which implies that (39) holds true. □

Figure 7 confirms the conclusion of Theorem 4.5 numerically. The error in this figure is due to the fact that we only use the average of 100 samples to approximate the expectation, and the error of POD-DCS solution. It is well-known that the eigenvalues decay rapidly in many practical problems, so Theorem 4.5 implies that the coefficients are compressible and illustrates the feasibility of using the POD-based ROM first.

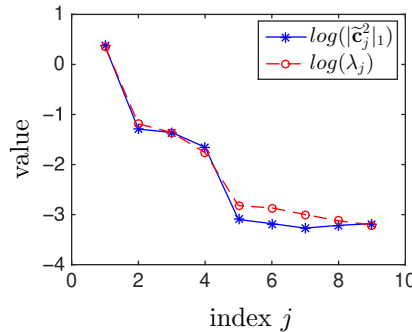


FIGURE 7. The eigenvalues and coefficients in the POD-DCS method with  $K_s = 100$ ,  $m = 9$  and  $K_c = 200$ .

**5. Numerical experiments.** In this section, we compare the POD-DCS scheme with POD-based ROM and the conventional CS method, and use our method to solve the 4-dimensional elliptic inverse problem (2)-(3) to further describe its feasibility and advantages.

**5.1. Comparison of accuracy.** The error estimates and sparsity of our POD-DCS scheme and other two different methods are shown in Table 2. All three methods are constructed with 100 full discrete finite element solutions. Among them, the POD-DCS method is constructed as above discussion with  $K_s = 100$ ,  $m = 9$ ,  $K_c = 200$ , POD-based ROM has same dimension  $m$ , and CS represents the conventional compressive sensing method.

Obviously, the POD method has highest accuracy, while the reconstruction error of reduced states makes the accuracy of our POD-DCS method to be inferior to POD, and relatively little information leads to the worst accuracy of the conventional CS method. In terms of sparsity, the proposed POD-DCS scheme is slightly better than the CS method. However, the total DoF of our method is  $0.2024 \times n \times m = 601$ , which is only 0.2% of that in CS method, and its DoF is  $0.2182 \times n \times n_u \approx 3.04 \times 10^5$ . The reason for the sharply decrease in the DoF is due to the fact that we use the reduced order model, which leads to the DoF in physical space decline from  $n_u$  to  $m$ . Therefore, compared to the CS method, the

proposed method can not only improve the accuracy and sparsity, but also greatly reduce the degree of freedom.

From the discussions in Section 3 and Theorem 4.4 we know that the accuracy of the POD-DCS method can be improved by increasing the number of snapshots, the dimension of POD basis, the number of collocation points and the order of polynomials. These quantities can be determined with the required accuracy.

TABLE 2. Error estimates and sparsity for different methods.

	POD-DCS	POD	CS
$\mathcal{E} \times 10^{-3}$	4.1758	2.8399	7.3896
$\mathcal{V} \times 10^{-4}$	2.4931	1.4750	11.1484
$\mathcal{R}_{0.01}$	0.2024	-	0.2182

**5.2. Comparison of efficiency.** Here we compare the efficiency of different methods. Table 3 summarizes the cost of each stage in the construction of different models. From the perspective of model construction, the number of full finite element solutions used in the three methods is the same. For POD-based ROM, although it has high accuracy and does not need to solve the BP problem, the time required to evaluate the model once is about 5 times that of the other two methods, which is not conducive to the implementation of the sampling algorithm. It takes only 0.2095s to evaluate the solution expression obtained by CS method, but the size of the dictionary matrix  $\Theta$  in the corresponding BP problem (25) is  $(100 \times n_u) \times (n \times n_u)$ , which is usually so large that it will exceed the computing resources or the calculation speed will be very slow. According to the nature of the finite element basis functions, here we can decompose the BP problem into solving the each column of the coefficient matrix  $\mathbf{c}$ , that is, transform the problem into solving  $n_u$   $l_1$ -minimization problems with size of  $100 \times n$ . This process takes about 1.4 hours. Since the sparse reconstruction in our POD-DCS scheme is just for reduced states, and the size of dictionary matrix is  $(m \times K_c) \times (m \times n)$ . It takes only 21s to solve this BP problem, which is 0.42% of the time required by the CS method. For the POD-DCS method, we also need to evaluate the reduced order models w.r.t.  $K_c$  collocation points, but it is usually much cheaper than the full finite element method.

Therefore, from an efficiency perspective, the offline cost of POD-based ROM is relatively small, but the online cost is larger than the other two methods. While compared with the CS method, the online time of the POD-DCS method only has a little difference, but the offline time has a obvious advantage. Moreover, our method achieves 11 times acceleration when evaluating a forward model, and the time required to construct the solution expression of stochastic surrogate model can be offset by the repeated calculation of the forward model. It is well known that both POD and stochastic collocation methods can deal with highly nonlinear problems, so for such complex problems, our method will be more attractive due to its high efficiency.

**5.3. Elliptic inverse problem.** The accuracy and efficiency comparison of the POD-DCS method with other two methods has been completed. Now, we utilize this method to deal with the elliptic inverse problem (2)-(3). We use finite element method with mesh size  $h = 2^{-8}$  to generated noise-free data associated with input



TABLE 3. Computational times, in seconds, given by different methods.

	# of full FE solution	Time for per FE solution	# of POD solution	Time for per POD solution	Time for BP	Time for per model output
POD-DCS	100	2.1970	200	1.0085	21	0.2002
POD	100	2.1970	-	-	-	1.0085
CS	100	2.1970	-	-	5045	0.2095

$\xi_o = [0.3, 0.3, 0.3, 0.3]^T$ . Note that the mesh size is finer than that used in inversion in order to prevent the “inverse crime”. The true permeability field  $a(\mathbf{x}; \xi_o)$  and noise-free output are shown in Figure 8. The  $n_z = 49$  measurement sensors are uniformly distributed over  $D$  with grid spacing  $2^{-3}$ .

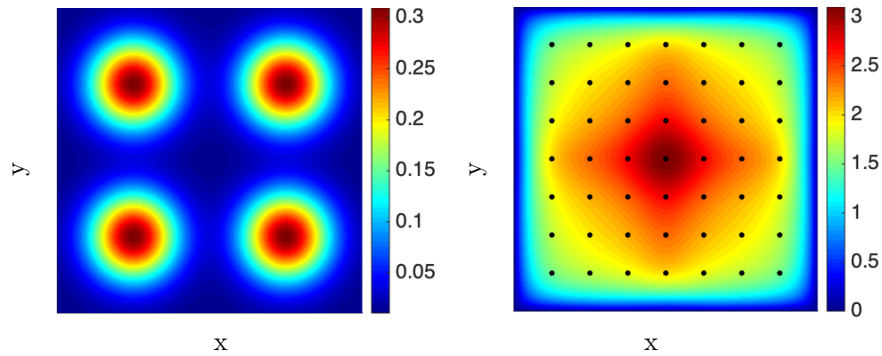


FIGURE 8. (Left): The true permeability field  $a(\mathbf{x})$  used for generating the synthesize data  $\mathbf{z}$ ; (Right): The model outputs associated with true permeability field, where black dots are the measurement sensors.

In this Bayesian inverse problem, the components of weight vector  $\xi$  are assumed to be i.i.d. and have a prior  $\xi_i \sim U(0, 1)$  for  $i = 1, \dots, n_p$ , where  $n_p = 4$ . As described in section 3, we first construct the POD-DCS approximate solution  $\tilde{u}(\mathbf{x}, \xi)$  of the stochastic problem (9) in prior support  $\Gamma = [0, 1]^{n_p}$  with  $K_s = 100, m = 9, K_c = 200$  and  $\tau = 0.01$ . Then substituting this approximate solution expression into (7) to get the approximate likelihood function, thereby obtaining the approximate posterior probability density function from Bayes’ rule (6). The framework for solving the Bayesian inverse problems using the POD-DCS approximate solution of the stochastic surrogate model is shown in Figure 9. Denote the approximate likelihood function and the approximate posterior probability density function associated with POD-DCS solution as  $\tilde{\pi}(\mathbf{z}|\xi)$  and  $\tilde{\pi}(\xi|\mathbf{z})$  respectively. The Hellinger distance and Kullback–Leibler distance between the exact posterior distribution  $\pi(\xi|\mathbf{z})$  and its approximation  $\tilde{\pi}(\xi|\mathbf{z})$  have been shown in [26] and [30]. Here, we use delayed-rejection adaptive Metropolis (DRAM) algorithm to explore the approximate posterior distribution  $\tilde{\pi}(\xi|\mathbf{z})$ , whose computational cost is only a fraction of that in exploring the true posterior distribution  $\pi(\xi|\mathbf{z})$ . This is because the cost of evaluating a finite element solution is about 11 times that of a POD-DCS solution, which can be observed from Table 3.

Here we consider that the observation data  $\mathbf{z}$  contains noise  $\mathbf{e}$ , which satisfies zero-mean Gaussian distribution  $N(\mathbf{0}, \sigma_e^2 I_{n_z})$  with  $\sigma_e = 0.05$ . For the finite element

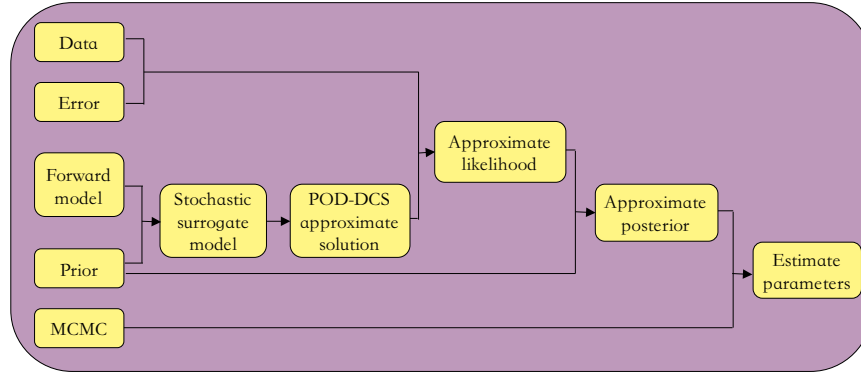


FIGURE 9. Bayesian inverse problems framework with POD-DCS approximate solution.

method, conventional CS method and POD-DCS method, we use the DRAM algorithm to draw  $3 \times 10^4$  samples respectively. Among them, the first 5000 samples are discarded as burn-in periods and the last 25000 samples are used to estimate the posterior densities. The 95% confidence interval and posterior marginal probability density of each component of unknown input parameters  $\xi = [\xi_1, \xi_2, \xi_3, \xi_4]$  are shown in Table 4 and Figures 10. Obviously, the 95% confidence intervals w.r.t. all these three methods contain true input parameters  $\xi_o$ , and the length of these confidence intervals does not exceed 0.03. But compared with the conventional CS method, the posterior marginal densities obtained by our scheme are closer to that obtained by FEM. In Figure 11, we display the estimated permeability field and its error, which correspond to the maximum a posterior probability (MAP) parameters generated by the POD-DCS method. We can see that the error between the estimated permeability field and the true permeability field is small enough. Therefore, the proposed method can deal with Bayesian inverse problems efficiently.

TABLE 4. The 95% confidence intervals of each input parameter obtained by different methods.

	FEM	POD-DCS	CS
$\xi_1$	[0.2898, 0.3173]	[0.2883, 0.3164]	[0.2963, 0.3248]
$\xi_2$	[0.2858, 0.3134]	[0.2844, 0.3126]	[0.2913, 0.3175]
$\xi_3$	[0.2913, 0.3191]	[0.2927, 0.3200]	[0.2906, 0.3208]
$\xi_4$	[0.2750, 0.3016]	[0.2762, 0.3047]	[0.2736, 0.3022]

From the previous discussion, it is clear that the POD-DCS surrogate model speeds up the evaluation speed of model while ensuring the accuracy. Therefore, this method can be considered for Bayesian inverse problem, optimal control and other problems requiring repeated evaluation of forward model. Note that the BP problem (25) is solvable in polynomial time [6]. While the dictionary matrix  $\Theta$  we used depends on the numbers of POD bases and stochastic bases, i.e.  $m$  and  $n$ . Therefore, for a specific problem, the dimension of POD-based ROM and the degree of the polynomial interpolation should be selected appropriately. It is further known that this method has limitations for the problem whose eigenvalue decays slowly, which is the inherent shortcoming of the POD method.

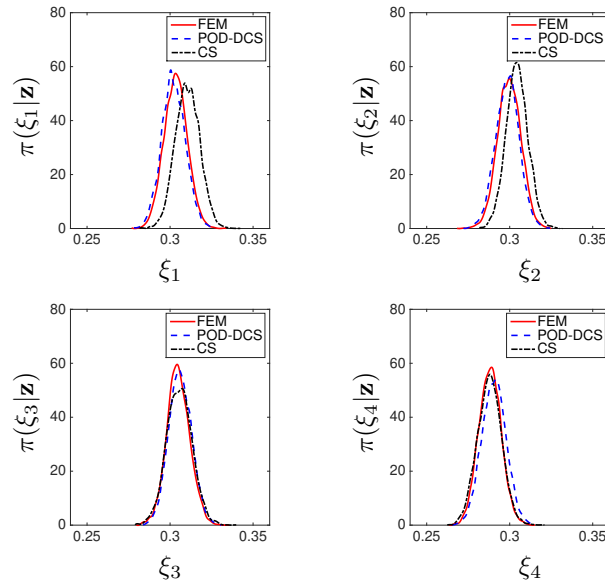


FIGURE 10. The posterior marginal densities of unknown parameters with  $\sigma_e = 0.05$  noise in the observations.

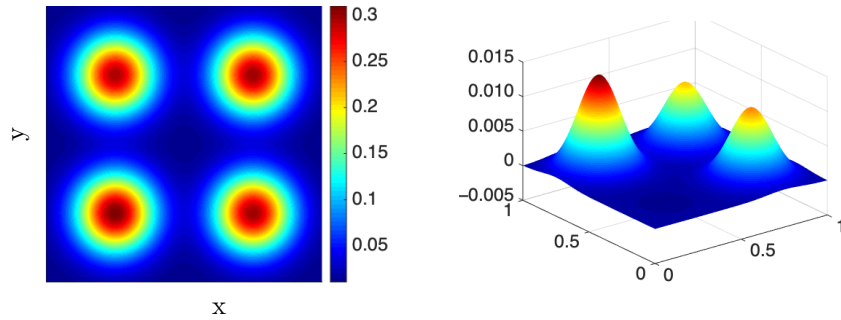


FIGURE 11. The estimated permeability field and error corresponding to the MAP parameters generated by the POD-DCS method. (Left): Estimated permeability field; (Right): Error.

**6. Conclusion.** In summary, for statistical inverse problems, we can regard the deterministic forward problem as a stochastic forward problem on the prior support of unknown parameters, and the solutions of these two problems with the same input are equal. Therefore, in this work, we propose a data-driven compressive sensing method based on proper orthogonal decomposition to construct the solution expression of stochastic surrogate model to accelerate the Bayesian inference of an inverse problem. The snapshot-based POD method is first used to construct the ROM of stochastic problem, then the stochastic collocation method based on gPC basis functions is adopted to represent the reduced states, and the coefficients are determined by solving an  $l_1$ -minimization problem. Substituting this approximate solution into the likelihood function can obtain an approximate likelihood function, thereby obtaining an approximate posterior. The approximate posterior space can be explored faster than the original one. We make error analysis and coefficient

estimation for this method, and prove that as the number of snapshots, POD cardinality, polynomial order, and number of collocation points increase, the mean square error of approximate solution will decrease. In addition, the expression coefficients are related to the eigenvalues. A series of numerical experiments show that when using the same number of full finite element solutions, the POD-DCS method has better accuracy and sparsity than the conventional CS method. Although our method needs to solve the ROM w.r.t. the collocation points, the sparse reconstruction of low-dimensional reduced states is much cheaper than the CS method. Moreover, in inferring the permeability field for elliptic PDE, our method can not only get an accurate result, but also be much faster than the direct calculation.

Accelerating the decline of eigenvalue in POD method is a problem worthy of consideration. And for complex problems, evaluating the forward model is costly. In order to construct an accurate reduced order model, we usually need many realizations, which can be prohibitive. Therefore, we plan to use a multi-fidelity scheme or select appropriate snapshots to overcome this difficulty. This is the subject of future work.

#### REFERENCES

- [1] I. Babuška, R. Tempone and G. E. Zouraris, [Solving elliptic boundary value problems with uncertain coefficients by the finite element method: The stochastic formulation](#), *Comput. Methods Appl. Mech. Engrg.* **194** (2005), 1251–1294.
- [2] K. S. Bormotin, [Inverse problems of optimal control in creep theory](#), *J. Appl. Ind. Math.*, **6** (2012), 421–430.
- [3] S. Brooks, [Markov Chain Monte Carlo method and its application](#), *Journal of the Royal Statistical Society*, **47** (1998), 69–100.
- [4] O. Bruno and J. Chaubell, [One-dimensional inverse scattering problem for optical coherence tomography](#), *Inverse Problems*, **21** (2005), 499–524.
- [5] E. J. Candès, [The restricted isometry property and its implications for compressed sensing](#), *C. R. Math. Acad. Sci. Paris*, **346** (2008), 589–592.
- [6] S. S. Chen, D. L. Donoho and M. A. Saunders, [Atomic decomposition by basis pursuit](#), *SIAM Rev.*, **43** (2001), 129–159.
- [7] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer-Verlag, Berlin, 1998.
- [8] A. Doostan and H. Owhadi, [A non-adapted sparse approximation of PDEs with stochastic inputs](#), *J. Comput. Phys.*, **230** (2011), 3015–3034.
- [9] H. W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problem*, Kluwer Academic Publishers, 1996.
- [10] E. M. Fernández-Berdaguer and G. B. Savioli, [An inverse problem arising from the displacement of oil by water in porous media](#), *Appl. Numer. Math.*, **59** (2009), 2452–2466.
- [11] M. Fornasier and H. Rauhut, [Compressive sensing](#), in: O. Scherzer (Ed.), *Handbook of mathematical methods in imaging*, Springer New York, (2015), 205–256.
- [12] M. Gunzburger and J. Ming, [Optimal control of stochastic flow over a backward-facing step using reduced-order modeling](#), *SIAM J. Sci. Comput.*, **33** (2011), 2641–2663.
- [13] M. D. Gunzburger, C. G. Webster and G. Zhang, [Stochastic finite element methods for partial differential equations with random input data](#), *Acta Numer.*, **23** (2014), 521–650.
- [14] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, Springer-Verlag, New York, 2005.
- [15] J. Kaipio and E. Somersalo, [Statistical inverse problems: Discretization, model reduction and inverse crimes](#), *J. Comput. Appl. Math.*, **198** (2007), 493–504.
- [16] I. Knowles and M. A. LaRussa, [Conditional well-posedness for an elliptic inverse problem](#), *SIAM J. Appl. Math.*, **71** (2011), 952–971.
- [17] J. Li and Y. M. Marzouk, [Adaptive construction of surrogates for the Bayesian solution of inverse problems](#), *SIAM J. Sci. Comput.*, **36** (2014), A1163–A1186.
- [18] H. Liang, Q. Sun and Q. Du, [Data-driven compressive sensing and applications in uncertainty quantification](#), *J. Comput. Phys.*, **374** (2018), 787–802.

- [19] C. Lieberman, K. Willcox and O. Ghattas, [Parameter and state model reduction for large-scale statistical inverse problems](#), *SIAM J. Sci. Comput.*, **32** (2010), 2523–2542.
- [20] J. Martin, L. C. Wilcox, C. Burstedde and O. Ghattas, [A stochastic newton MCMC method for large-scale statistical inverse problems with application to seismic inversion](#), *SIAM J. Sci. Comput.*, **34** (2012), A1460–A1487.
- [21] Y. M. Marzouk, H. N. Najm and L. A. Rahn, [Stochastic spectral methods for efficient Bayesian solution of inverse problems](#), *J. Comput. Phys.*, **224** (2007), 560–586.
- [22] Y. M. Marzouk and H. N. Najm, [Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems](#), *J. Comput. Phys.*, **228** (2009), 1862–1902.
- [23] N. Petra, J. Martin, G. Stadler and O. Ghattas, [A computational framework for infinite-dimensional Bayesian inverse problems, Part II: Stochastic Newton MCMC with application to ice sheet flow inverse problems](#), *SIAM J. Sci. Comput.*, **36** (2014), A1525–A1555.
- [24] H. Rauhut and R. Ward, [Sparse Legendre expansions via  \$l\_1\$ -minimization](#), *J. Approx. Theory*, **164** (2012), 517–533.
- [25] G. R. Richter, [An inverse problem for the steady state diffusion equation](#), *SIAM J. Appl. Math.*, **41** (1981), 210–221.
- [26] A. M. Stuart, [Inverse problems: A Bayesian perspective](#), *Acta Numer.*, **19** (2010), 451–559.
- [27] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, Halsted Press, 1977.
- [28] A. Voutilainen and J. P. Kaipio, [Model reduction and pollution source identification from remote sensing data](#), *Inverse Probl. Imaging*, **3** (2009), 711–730.
- [29] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, 2010.
- [30] L. Yan and L. Guo, [Stochastic collocation algorithms using  \$l\_1\$ -minimization for Bayesian solution of inverse problems](#), *SIAM J. Sci. Comput.*, **37** (2015), A1410–A1435.
- [31] W. W. Yeh, [Review of parameter identification procedures in groundwater hydrology: The inverse problem](#), *Water Resources Research*, **22** (1986), 95–108.
- [32] H. Ying and B. Han, [A wavelet adaptive-homotopy method for inverse problem in the fluid-saturated porous media](#), *Appl. Math. Comput.*, **208** (2009), 189–196.

Received November 2020; revised March 2021.

*E-mail address:* [xmx2018@hust.edu.cn](mailto:xmx2018@hust.edu.cn)

*E-mail address:* [liuhong\\_c@hust.edu.cn](mailto:liuhong_c@hust.edu.cn)

*E-mail address:* [jming@hust.edu.cn](mailto:jming@hust.edu.cn)

*E-mail address:* [jaemin.shin@hanbat.ac.kr](mailto:jaemin.shin@hanbat.ac.kr)