Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 345 (2019) 263–282
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

# Sparsity-promoting elastic net method with rotations for high-dimensional nonlinear inverse problem

Yuepeng Wang[a,*], Lanlan Ren[a], Zongyuan Zhang[a], Guang Lin[b,c,**], Chao Xu[d]

[a] *School of Mathematics and Statistics, Nanjing University of Information Science and Technology (NUIST), Nanjing, 210044, China*
[b] *Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA*
[c] *School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA*
[d] *State Key Laboratory of Industrial Control Technology and Institute of Cyber-Systems & Control, Zhejiang University, Hangzhou, Zhejiang, 310027, China*

Received 7 June 2018; received in revised form 28 October 2018; accepted 30 October 2018
Available online 8 November 2018

## Highlights

- An elastic-net-based sparse polynomial chaos-ensemble Kalman filter is designed.
- Regularization parameters are selected with the information criterion.
- First work on employing the iterative rotations to the inverse problem.
- The selection of the optimal number of iterative rotations is studied.
- Gradient matrix is constructed in a multi-parameter response model.

## Abstract

An elastic-net (EN) based polynomial chaos (PC) ensemble Kalman filter (PC-EnKF) with iterative PC-basis rotations is developed for high-dimensional nonlinear inverse modeling. To avoid the huge computational cost of estimating PC expansion coefficients and the Kalman gain matrix in PC-EnKF, this paper focuses mainly on solving the minimization problem of the elastic-net (EN) cost function with the fast iterative shrinkage-thresholding algorithm (FISTA). To further enhance the sparsity and accuracy, an iterative PC-basis rotation method is employed. When performing the rotation technique, two key issues need to be addressed to accommodate the computation of the inverse problem. One is the derivation of a new multi-dimensional random variable. This can be realized by exploring the construction of the gradient matrix used in a multi-parameter and vector-valued response model. The other issue is the selection of the number of iterative rotations during the process of each data assimilation, which can be addressed by resorting to a curve of sparsity versus the number of iterations. As for the regularization parameters, they can be tuned by calculating the information criteria (IC). Through the numerical examples, we demonstrate that EN-based PC-EnKF combined with the iterative PC-basis rotation method is well suited in the high-dimensional nonlinear inverse modeling, and has great potential in the high-dimensional nonlinear inverse modeling of real-world complex systems.

* Corresponding author.
** Corresponding author at: School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA.
*E-mail addresses:* eduwyp@nuist.edu.cn (Y.P. Wang), Guanglin@purdue.edu (G. Lin), cxu@zju.edu.cn (C. Xu).

https://doi.org/10.1016/j.cma.2018.10.040

*Keywords:* Uncertainty quantification; Data assimilation; FISTA; Polynomial Chaos; Regularization; Ensemble Kalman filter

## 1. Introduction

Currently, the systematic uncertainty quantification (UQ) in models, simulations, experiments and the analysis of how they are propagated through complex models to affect the predicted outcomes constitutes an active area of research [1–3]. For the computational process from input (parameters) to output (simulation results), we refer it as the 'forward UQ'. Correspondingly, 'inverse UQ' means the inference of the input variables based on the output by gathering more observation information of the output variables to calibrate the simulation results and reduce the uncertainties of the input variables [1]. This process is also called 'data assimilation' which combines essentially the underlying dynamical system and the available data. In a comparison of the two UQ problems, the inverse UQ belongs to the inverse problem, which is generally ill-posed in the sense of Hadamard: it fails to satisfy at least one of the criteria for well-posedness — existence, uniqueness, and continuous dependence of the solution on data. A phenomenon occurs to us, that is, for a single output value, there may be multiple corresponding input values that all provide a match to the output value. This is because the information contained in the observed data is insufficient to determine all of the uncertain model parameters. So the parametric uncertainty usually cannot be totally eliminated, merely reduced [1]. For this reason, the inverse problem has received increasing attention in the UQ community [2].

The Kalman filter [3] is widely used as a data assimilation method. The original linear Kalman filter can be only successful if all the probability distributions involved are Gaussian, i.e., the system is linear, and all the random variables are normally distributed. As one of its variants, the extended Kalman filter (EKF) [4] can also be applied to deal with the moderately nonlinear problems by linearizing it via the Jacobian. However, it will suffer when a strong nonlinear inverse problem is involved, this is because the Jacobian provides only local information. In 1994, the ensemble Kalman filter (EnKF) was introduced in [5], which greatly alleviates the nonlinear challenge. It tracks the distribution evolution of the quantities of interest by using Monte Carlo sampling of the variable space that is evolved with time. Owing to the simplicity in its implementation, EnKF has gained great popularity within the past two decades. Since its invention, EnKF has been widely used in different fields such as oceanography, reservoir engineering, and meteorology [6–9]. Due to the slow convergence with the number of ensembles, a large ensemble size is required to get an accurate estimation of the system and an even larger size is needed for the estimation of the associated uncertainty. This is usually extremely computationally demanding due to the repeated analyses that have to take place, while one can afford only a small ensemble size owing to the fact that it is time-consuming to run each simulation for large-scale computational problems (say, with $O(10^4)$ samples in high-dimensional stochastic space). So in order to enhance the computational efficiency of the method, different techniques have been proposed to reduce the sampling errors in the EnKF with smaller-sized ensembles, which turns out to be promising for this purpose, e.g. [10,11].

Recently, an effort to minimize the number of simulations for a given required accuracy has led to the development of another variant of the Kalman filter: polynomial chaos (PC) based ensemble Kalman filter (PC-EnKF). The PC-EnKF developed in [2,12,13] can also be applied to the nonlinear inverse problems. In this method, the random process of interest is represented by the PC bases which are the orthogonal polynomials with respect to a set of independent random variables with known distributions [14–16]. Once the PC representation is obtained, the statistical moments of our interest (i.e. the mean and covariance of the random quantities) can be easily computed from the PC coefficients. The PC-EnKF resembles the traditional EnKF in every aspect except that it represents and propagates model uncertainty by PC expansion instead of an ensemble of model realizations. This method turns out to be a more efficient alternative to EnKF for many data assimilation problems [2,17–19].

In order to determine the unknown PC expansion coefficients of the solution, there commonly exist two types of methods to be resorted to, the intrusive [20–25] and the non-intrusive ones [26–28]. Historically, the intrusive Stochastic Galerkin (SG) method was used in [21,29]. However, it must solve a system of coupled equations which require robust and efficient solvers and the modification of an existing deterministic code. Often, the form of equations or code used to solve the deterministic equations is complex, which makes the implementation of

the intrusive SG method difficult, if not impossible. Compared with the 'intrusive' method, the advantage of 'non-intrusive' method is that there is no need to modify the deterministic solvers for the quantities of interest (QoI). The fundamental idea behind the 'non-intrusive' approach is essentially the repeated application of the existing or legacy deterministic solver. We consider non-intrusive sampling methods in our current study. However, the trade-off is frequently encountered, that is, keeping more PC basis functions in PC decomposition helps to capture uncertainty more accurately, but it increases the computational cost. An ideal PC expansion should accurately represent the model uncertainty but keep the number of basis functions as small as possible. It is challenging to identify such a subset of PC bases that have the strongest impact on the model uncertainty. Recent attempts at extracting only a subset of desired PC bases rely on the compressive sampling [30–34] and the formulation of convex optimization. In this paper, we are more interested in developing a sparsity-promoting PC-EnKF, where a $l_1$-solver is introduced to determine the coefficients of PC expansion in the prediction step of ensemble Kalman filter.

The $l_1$-solver essentially combines tools and ideas from convex optimization with compressive sensing. Among the most popular $l_1$ solvers, the Lasso (least absolute shrinkage and selection operator) [35] has been widely used in compressed sensing and image processing, where many regression coefficients are expected to be zero, and only a small subset of coefficients to be non-zero. The Lasso sparse solution is therefore obtained via a $l_1$ penalized least-squares criterion. Currently, there are many methods that can be implemented for solving this $l_1$ optimization problem. For instance, the least angle regression algorithm [36] (achieve the same time complexity as ordinary least-squares regression) and the even more efficient coordinate descent algorithm [37].

Recently, there is an important family of methods in connection with the iterative shrinkage-thresholding algorithm (ISTA). Initially, ISTA was introduced as an EM (expectation–maximization) algorithm for image deconvolution [38], later in [39], using a majorization–minimization approach. In [40], ISTA was placed on solid mathematical grounds, with a rigorous convergence proof in an infinite dimensional setting. According to [41], it can be used for Lasso problems and an obvious benefit is that each iteration of the ISTA only involves sum and relatively cheap matrix–vector multiplication followed by a shrinkage/soft-thresholding step in the computational procedure. However, ISTA has also been recognized as a slow method. In this case, a new fast iterative shrinkage-thresholding algorithm (FISTA) has been proposed [42], which preserves the computational simplicity of ISTA but with a global rate of convergence which is proven to be significantly better, both theoretically and practically, and shown to be faster than ISTA by several orders of magnitude.

However, there are still some of the weaknesses in Lasso estimation, e.g. the Lasso procedure is not stable enough when there exist high correlations among the variables, and Lasso tends to arbitrarily choose some important variables and ignore the other important variables when they have relatively high correlation or group structures [43]. All these often preclude the use and potential advantage of Lasso. To remedy this, as an improved version of the Lasso, the elastic net regularization [44] was proposed, and elastic net-based methods are widely used in statistics and machine learning, see, e.g., [45]. Elastic net is interesting in the sense that it combines both the $l_1$(Lasso) and the $l_2$(ridge) penalties. The elastic net (EN) keeps the model sparsity of Lasso, while the inclusion of $l_2$ penalization term stabilizes the estimation [45] and, hence, improves or maintain the prediction accuracy. Particularly, when there exist high correlations among variables, the elastic-net can significantly improve the prediction accuracy and even outperforms the Lasso. When the elastic net problem is solved, a common approach is often carried out by transforming the elastic net problem into an equivalent Lasso problem on an augmented data, based on which the least angle regression (LAR) procedure is applicable, referred to least angle regression elastic net (LAR-EN) [44]. Different from this optimization procedure, an iterative thresholding algorithm is designed and explored using convex analysis tool for computing the elastic-net solution [45], which is akin instead to the algorithm developed in [40], and the consistency properties of the elastic net scheme are further investigated within a suitable mathematical framework. In practice, the Lasso or elastic net regularization usually obtains a high-dimensional model which contains the true model with high probability [43]. So there is still room for making further improvement to increase the sparsity by introducing advanced techniques.

More recently, a promising novel iterative-rotation technique is proposed [46,47]. When the Lasso or elastic net is applied to the representation of QoI with Hermite chaos expansion to determine the coefficients, the original inputs are rotated such that a few of the new coordinates, i.e. linear combinations of original inputs, have significant impact on QoI, thereby increasing the sparsity of the solution and in turn, the accuracy of recovery [48].

The main task of current work is to explore the reformulation of the elastic net solution in the framework of PC-EnKF for the high-dimensional inverse problem when the iterative-rotation technique is employed. And the focus is placed on how the iterative-rotation approach can effectively improve the computational results given the limited

noisy measured data. For demonstrating the current algorithm, we solve an inverse problem from the 2D shallow water equations and attempt to recover the representation of accurate channel bed topography that is still a challenge [49,50] due to its ill-posed essence. Such a problem arises frequently in the field of hydraulic modeling of open channel flows in which the topography shape is an important parameter that needs to be identified prior to numerical modeling and flow simulation. So far, various methods have been developed for addressing this issue, please see [51] for details. To the best of our knowledge, this is the first application of the iterative-rotation approach to the high-dimensional inverse problem, especially in the setting of PC-EnKF. When the rotation technique is used, two key issues need to be addressed to accommodate the present computation of inverse problem, especially in the setting of PC-EnKF. One is the derivation of a new multi-dimensional random variable, which can be realized by exploring the construction of gradient matrix used in a multiparameter and vector-valued response model which is different from the case of literature [47], and the other is the selection of the number of iterative rotations during the process of each data assimilation, which can be addressed by resorting to a curve of the sparsity versus the number of iterations. For the determination of coefficients of PCE, the fast iterative soft-thresholding algorithm (FISTA) of the elastic net is explored, wherein the Bayesian information criterion (BIC) (Schwartz 1978) [52] is used for helping us to determine the optimal regularization parameters due to its consistency in selecting the true model (Shao 1997) [53]. We will concentrate mainly on the three points to evaluate the utility of the present algorithm: the quality of the recovery results of input topography, the sparsity of PC expansion coefficients and the reduction of uncertainty prediction. The experimental results show that the elastic net with iterative rotations is more effective in promoting the sparsity than elastic net for solving high-dimensional inverse UQ. We hope that the current algorithm developed for reducing the uncertainty of input random parameters will be of general interest and value.

The rest of this paper is structured as follows. First, in Section 2 we provide a formulation of rotation-based elastic net in the framework of PC-EnKF method; Secondly, in Section 3, numerical simulation experiments are carried out, and the recovery of topography in 2D shallow water equations is allowed to demonstrate the potential benefits and usefulness of the present algorithm, and then some conclusions and ideas in future research are provided in Section 4.

## 2. Rotation-based elastic net in the framework of PC-EnKF for inverse problem

### 2.1. Problem description

The model uncertainty prediction can be described using the following relation:

$$\mathbf{d} = g(\mathbf{m}) \tag{2.1.1}$$

where $\mathbf{m}$ denotes the parametric input variables or input parameters, and $\mathbf{d}$ is the model output. Generally, $g$ is the mathematical model and often involves partial differential equations (e.g. the 2D shallow water equations used in this paper). For the forward UQ, the mathematical model $g$ is used to predict the output $\mathbf{d}$ from the input $\mathbf{m}$. While for the inverse UQ, the mathematical model is used to make inferences about input $\mathbf{m}$ that would result in the given the observation data $\mathbf{d}^*$, $\mathbf{d}^* = \mathbf{d} + \epsilon$, where $\epsilon \sim p_\epsilon(\cdot)$ is a random variable with known pdf that represents uncertainty in the observation data. Assuming that prior information about the input $\mathbf{m}$ is available in the form of a pdf $p_0(\mathbf{m})$, the Bayes' rule allows us to solve the inverse UQ in a principled fashion, which states that

$$p(\mathbf{m}|\,\mathbf{d}^*) \propto p(\mathbf{d}^*\big|\,\mathbf{m})p_0(\mathbf{m}) \tag{2.1.2}$$

where $p(\mathbf{d}^*|\,\mathbf{m}) = p_\epsilon(\mathbf{d}^* - g(\mathbf{m}))$ stands for the likelihood. The formulation (2.1.2) tells us that the posterior $p(\mathbf{m}|\,\mathbf{d}^*)$, which is what we know about the unknown $\mathbf{m}$ given the data $\mathbf{d}^*$, is proportional to the likelihood $p(\mathbf{d}^*|\,\mathbf{m})$, which measures how likely the observed data is for given inputs $\mathbf{m}$, multiplied by the prior $p_0(\mathbf{m})$, which describes our knowledge of the unknown prior to the acquisition of data. Once the prior and forward model are specified, the posterior distribution is defined via Bayes' rule (2.1.2). Thus, the mean value and variance of $\mathbf{m}$ are derived theoretically. In the current study, we will apply the Bayesian parameter estimation method, called the EN-based PC-EnKF method with iterative rotations, to explore them as stated in Section 1. In order to see the benefit of this method, let us first introduce the polynomial chaos (PC) expansion in the following section.

### 2.2. PC representation of the uncertainty

To characterize uncertainty, We model the uncertain inputs as a $n$-dimensional vector of independent random variables $\xi := (\xi_1, \xi_2, \ldots, \xi_n)$, with probability density function $\rho(\xi)$. The QoI that we seek to approximate is $\mathbf{d}(\xi)$. Supposed that its variance is finite, then we can utilize PC expansions $\psi_j(\xi)$ to approximate $\mathbf{d}(\xi)$, which is of the form:

$$\mathbf{d}(\xi) = \sum_{j=0}^{P} c_j \psi_j(\xi) + \epsilon_t(\xi) \tag{2.2.1}$$

where $c_j$, $j = 0, 1, \ldots, P$, are the corresponding PC expansion coefficients. Denoting the maximum order of the truncated polynomials by $l$, the total number of terms is given by $P + 1 = \frac{(n+l)!}{n!l!}$. $\epsilon_t$ is truncation error associated with retaining $P + 1$ terms of PC bases. With the growth of the polynomial order and the number of random variables, the total number of terms in the expansion increases rapidly. We assume that $\psi_j(\xi)$ are normalized such that $E(\psi_j^2(\xi)) = 1$, where the operator $E$ denotes the mathematical expectation. In this work, the parametric input variables are assumed to be Gaussian, and therefore Hermite polynomials are chosen according to the Wiener–Askey scheme. The Hermite polynomials are normalized and the weight function is: $w(x) = \frac{1}{\sqrt{2\pi}} \exp(\frac{-x^2}{2})$. To identify PC expansion coefficients, we consider non-intrusive sampling methods, in which deterministic solvers for the QoI are not modified. Such methods include Monte Carlo simulation [54], pseudo-spectral stochastic collocation [55,56], least squares regression [57], and $l_1$-minimization [34,58–62]. Another method that has attracted a revived interest and considerable amount of attention in the signal processing literature is the elastic net. In this work, we will adopt the elastic net [44] to estimate the coefficients of PCE by minimizing the function

$$J(\mathbf{c}) = \frac{1}{2} \|\mathbf{d} - \mathbf{\Psi c}\|_2^2 + \frac{\lambda_2}{2} \|\mathbf{c}\|_2^2 + \lambda_1 \|\mathbf{c}\|_1 \tag{2.2.2}$$

where $\lambda_1$ and $\lambda_2$ are positive scalars, and the vector $\mathbf{c} := (c_0, \ldots, c_P)$ contains the PC expansion coefficients, while the vector $d$ and the so-called measurement matrix $\Psi$ contain function evaluations at realizations of the random input, $\xi$. Specifically, denoting the $i$th realization of $\xi$ as $\xi^{(i)}$, $\mathbf{d} := (d(\xi^{(1)}), \ldots, d(\xi^{(N)}))$, and $\mathbf{\Psi}(i, j) := \psi_j(\xi^{(i)})$. When $\lambda_2$ is equal to zero, then Eq. (2.2.2) is just the following form

$$J(\mathbf{c}) = \frac{1}{2} \|\mathbf{d} - \mathbf{\Psi c}\|_2^2 + \lambda_1 \|\mathbf{c}\|_1 \tag{2.2.3}$$

The problem (2.2.3) is known as Lasso, in which the residual is measured with the Euclidean norm and the regularization is done with $l_1$-norm. It intends to seek a balance between the sparsity and fitting, which is also named basis pursuit denoising (BPDN) by Chen et al. [63]. The sparsity-seeking property of (2.2.3) has been shown to have applications in geophysics, data compression, image processing, sensor networks, and more recently, in computer vision, and so on. The interested reader is referred to [64,65] for a comprehensive review of these applications.

### 2.3. Solving the elastic net using the fast iterative shrinkage-thresholding algorithm (FISTA) with rotations

The cost criterion appearing in (2.2.2) is known as the elastic net (EN). The elastic net criterion has some desirable properties, as it maintains the model sparsity of Lasso, but not as aggressive as Lasso in excluding correlated terms in the model. This is because these terms tend to be in or out of the model together as a result of the $l_2$ norm regularization by guaranteeing strong convexity of the cost. For the applications of the elastic-net regularization to statistics and learning theory, one may refer to De Mol et al. (2009) [45] and Hastie, Tibshirani, and Friedman (2009) [66]. Elastic net-based methods are amenable to very efficient large-scale solution algorithms (Friedman, Hastie, & Tibshirani, 2010) [37]. According to Eq. (2.2.2), the elastic net solution can be defined as

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \frac{1}{2} \|\mathbf{d} - \mathbf{\Psi c}\|_2^2 + \frac{\lambda_2}{2} \|\mathbf{c}\|_2^2 + \lambda_1 \|\mathbf{c}\|_1 , \tag{2.3.1}$$

So far, there are several numerical algorithms proposed for solving this optimization problem. The most commonly used method, as suggested by Zou and Hastie in [44], is frequently to transform the elastic net problem into an equivalent Lasso problem, i.e.,

$$\hat{\mathbf{c}}^* = \arg\min_{\mathbf{c}^*} \frac{1}{2} \left\|\mathbf{d}^* - \mathbf{\Psi}^* \mathbf{c}^*\right\|_2^2 + \gamma \left\|\mathbf{c}^*\right\|_1 , \tag{2.3.2}$$

where

$$\mathbf{\Psi}^* = (1+\lambda_2)^{-1/2}\begin{pmatrix}\mathbf{\Psi}\\ \sqrt{\lambda_2}\mathbf{I}\end{pmatrix}, \mathbf{d}^* = \begin{pmatrix}\mathbf{d}\\ 0\end{pmatrix} \text{ and } \gamma{=}\lambda_1/\sqrt{1{+}\lambda_2}$$

Then the $\hat{\mathbf{c}}$ can be derived using the following the relation

$$\hat{\mathbf{c}}{=}\frac{1}{\sqrt{1{+}\lambda_2}}\hat{\mathbf{c}}^*, \tag{2.3.3}$$

just due to the identity (2.3.2), the elastic net is often referred to as a particular version of the Lasso, based on which the LAR procedure is applicable. An efficient algorithm, called least angle regression elastic net (LAR-EN), is therefore designed to deal with the elastic net estimator (2.3.1), resulting in an improved performance [44]. More recently, another interesting contribution is the fast iterative shrinkage-thresholding algorithm (FISTA) that is developed for solving elastic net problems. The algorithm foundations are for instance thoroughly described by A. Beck and M. Teboulle [42]. Let us review very briefly its main idea and characteristics.

First consider a generalized form of the optimization problem:

$$\min F(x) := f(x) + g(x) \tag{2.3.4}$$

where, $g : R^n \to R$ a continuous convex function which may not be smooth. $f : R^n \to R$ a convex function with lower bound, continuously differentiable with Lipschitz continuous gradient $L(f) > 0$:

$$\|\nabla f(x) - \nabla f(y)\| \le L(f)\|x - y\|, \text{ for every } x,\ y \text{ in } R^n$$

where $\|\cdot\|$ denotes the standard Euclidean norm. The idea of iterative shrinkage-thresholding algorithm is to solve the original problem by solving an approximate model of the original problem. For any $L > 0$, consider the following quadratic approximation of $F(x) := f(x) + g(x)$ at a given point $y$:

$$Q_L(x, y) := f(y) + \langle x - y, \nabla f(y)\rangle + \frac{L}{2}\|x - y\|^2 + g(x) \tag{2.3.5}$$

Then it has a unique minimum point:

$$p_L(y) := \arg\min_x Q_L(x, y) \tag{2.3.6}$$

Ignoring the constant term we can get

$$p_L(y) := \arg\min_x \left\{ g(x) + \frac{L}{2}\|x - (y - \frac{1}{L}\nabla f(y))\|^2 \right\}$$

Thus, we can get the basic iteration of the generalized model $F(x)$

$$x_k = p_L(x_{k-1}) \tag{2.3.7}$$

Especially, when $g(x) = \lambda\|x\|_1$, the above iteration (2.3.7) can be then accomplished by the following iterative shrinkage thresholding algorithm (ISTA) (Daubechies et al., 2004) [40], which is a simple iterative scheme consisting of a gradient and a shrinkage step:

$$x_k = S_{\frac{\lambda}{L}}\left(x_{k-1} - \frac{1}{L}\nabla f(x_{k-1})\right) \tag{2.3.8}$$

where $S_\alpha : R^n \to R^n$ is the shrinkage thresholding operator defined by

$$S_\alpha(x) := \begin{cases} x - \frac{\alpha}{2} & x > \frac{\alpha}{2} \\ 0 & |x| \le \frac{\alpha}{2} \\ x + \frac{\alpha}{2} & x < -\frac{\alpha}{2} \end{cases}$$

As was pointed out in [42], when $f(\mathbf{x}) = \|\mathbf{d} - \mathbf{\Psi}^*\mathbf{x}\|^2$, the (smallest) Lipschitz constant of the $\nabla f$ can be provided by $L(f) = \lambda_{max}(\mathbf{\Psi}^{*\prime}\mathbf{\Psi}^*)$. The gradient step adds a non-sparse update to the current iterate. Subsequently, the shrinkage operator $S_\alpha$ shrinks the updated solution (i.e., PCE coefficients) componentwise by $\lambda$ towards zero. This step ensures that only dominant components can increase to non-zero values. The ISTA converges rather slowly.

In 1983, Y.E. Nesterov proposed a new gradient algorithm in the literature [67]. And Amir Beck and Marc Teboulle applied this idea to ISTA and proposed FISTA in 2009 [42]. The main difference between FISTA and the traditional algorithm is that the calculation of the new iteration point, say $x_{k+1}$, depends on a specific linear combination of the previous two points $\{x_k, x_{k-1}\}$ rather than $x_k$ only. This improvement does not increase the computational difficulty of the algorithm and speeds up the convergence of solution instead. The convergence rate is $O(\frac{1}{k^2})$ for FISTA, and in contrast, only $O(\frac{1}{k})$ for ISTA, where $k$ is the iteration counter. In the current study, we use FISTA to solve the reformulated Lasso problem (2.3.2), and for the related details and steps, see Algorithm 1.

**Algorithm 1** FISTA

**Input:** $\Psi, u, \lambda_1, \lambda_2$

**Initialize:** $y_1 = x_0 \in R^n$, $t_1 = 1$, max_iter, tol $= 1e-8$

$\mathbf{\Psi}^* = \mathbf{\Psi}'\mathbf{\Psi} + \lambda_2 \text{eye}(P+1)$

$L = \max(\text{eig}(\mathbf{\Psi}^{*'}\mathbf{\Psi}^*))$

**for** $k = 1 \rightarrow$ max_iter **do**

$x_k = S_{\frac{\lambda_1}{L}}\left(y_k - \frac{1}{L}\mathbf{\Psi}^{*'}(\mathbf{\Psi}^* y_k - u)\right)$

$t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$,

$y_{k+1} = x_k + \left(\frac{t_k - 1}{t_{k+1}}\right)(x_k - x_{k-1})$

error $= \|x_k - x_{k-1}\|_1/(P+1)$

**if** error $<$ tol **then**

Break

**Output:** $x_k$

The elastic net offers some advantages over the Lasso in certain situations, but these advantages come with the cost of needing to tune the model with respect to two regularization parameters $\lambda_1$ and $\lambda_2$. This can be done generally by a two-dimensional search via K-fold cross-validation (CV), but tremendously tedious and time-consuming, even impossible, particularly for the large-scale inverse problem. Another possibility is to use information criteria, for example, the Akaike information criterion (AIC) (Akaike, 1973) [68] and Bayesian information criterion (BIC) (Schwartz, 1978) [52] and so on, to select the regularization parameters. The AIC comes from approximately minimizing the difference between the true data distribution and the model distribution, known as the Kullback–Leibler information entropy. While Schwarz derived BIC to asymptotically approximate a transformation of the Bayesian posterior probability of a model. Using either AIC or BIC to pick a suitable model is typically indicated by the smallest value of each criterion. So we will face an optimization problem below to find the optimal elastic net model for the Eqs. ((2.2.2) or (2.3.1))

$$\lambda_{1,2}(\text{optimal}) = \arg\min_{\lambda_{1,2}} \frac{\|\mathbf{d} - \mathbf{\Psi} c_{\lambda_{1,2}}\|^2}{N_o \sigma^2} + \frac{w_{N_o}}{N_o} d\hat{f}(\lambda_{1,2}) \tag{2.3.9}$$

where $w_{N_o} = 2$ for AIC and $w_{N_o} = \log(N_o)$ for BIC. $N_0$ is the number of samples. $\sigma^2$ is the residual variance, which is here defined for the non-zero $\lambda_2$ as $\sigma^2 = \frac{\|\mathbf{d} - \mathbf{\Psi}^+\mathbf{d}\|^2}{N_0}$, $\mathbf{\Psi}^+ = \mathbf{\Psi}(\mathbf{\Psi}'\mathbf{\Psi})^{-1}\mathbf{\Psi}'$. As for $d\hat{f}(\lambda_{1,2})$, called the degree of freedom (DOF), it is often used to quantify the complexity of a model fit. Zou, Hastie & Tibshirani (2007) [69] and Ryan J. Tibshirani & Jonathan Taylor (2012) [70] show that an unbiased estimate of the number of degrees of freedom of elastic net solutions can be obtained by the following relation

$$d\hat{f}(\lambda_{1,2}) = \text{Trace}[\mathbf{\Psi}_A(\mathbf{\Psi}_A{}'\mathbf{\Psi}_A + \lambda_2 I)^{-1}\mathbf{\Psi}_A{}'] \tag{2.3.10}$$

where the prime ‘’ represents transpose, $A$ stands for the active set, which is defined as $A = \{i \in [1, 2, \ldots, P+1] : c_i \neq 0\}$, and $c_i$ is the component of the solution $\mathbf{c}$. $\mathbf{\Psi}_A$ is formed from the columns of $\mathbf{\Psi}$ associated with non-zero coefficients given $\lambda_2$. Accordingly, $I$ is an identity matrix of dimension equal to the cardinality of the active set. Note that for Eq. (2.3.10), inverting $\mathbf{\Psi}_A{}'\mathbf{\Psi}_A$ is computationally expensive: $O[(P+1)^3]$. Therefore, the singular value decomposition can be effectively utilized, i.e., $\mathbf{\Psi}_A = \check{U} D \check{V}'$. If so, Eq. (2.3.10) can be then written as

$$d\hat{f}(\lambda_{1,2}) = \text{Trace}[\check{U} D(D^2 + \lambda_2 I)^{-1} D\check{U}'] \tag{2.3.11}$$

For the selection of parameters $\lambda_1, \lambda_2$, we first create two sets of grid points using reasonable values for them as $\Lambda_1$ and $\Lambda_2$, respectively. Then we perform the Algorithm 2 over $\Lambda_1$ for each value of $\lambda_2$, and compare all the $\min_{\Lambda_1} \mathrm{BIC}(\lambda_1, \lambda_2)$ to offer the $\min_{\Lambda_2} \min_{\Lambda_1} \mathrm{BIC}(\lambda_1, \lambda_2)$, by which a specific parameter pair $(\lambda_1, \lambda_2)$ is eventually determined.

**Algorithm 2** The calculation of IC (BIC or AIC) and optimal solution

**Input:** $\boldsymbol{\Psi}, \mathbf{d}, \{\lambda_2^j\}, \{\lambda_1^i\}$

**for** $j = 1 \to n_j$ **do**

  Solve the corresponding ridge regression problem to get $\mathbf{c}_j^0$: $\mathbf{c}_j^0 = \arg\min_x \frac{1}{2}\|\mathbf{d} - \boldsymbol{\Psi}\mathbf{c}_j^0\|_2^2 + \frac{\lambda_2^j}{2}\|\mathbf{c}_j^0\|_2^2$;

  Compute the mean square error: $\sigma_j^2 = \|\mathbf{d} - \boldsymbol{\Psi}\mathbf{c}_j^0\|_2^2 / N_o$, $N_o$ is the number of samples.

  **for** $i = 1 \to n_i$ **do**

    Solve EN by FISTA: $\mathbf{c}^{i,j} = \arg\min_x \frac{1}{2}\|\mathbf{d} - \boldsymbol{\Psi}\mathbf{c}^{i,j}\|_2^2 + \lambda_1^i \|\mathbf{c}^{i,j}\|_1 + \frac{\lambda_2^j}{2}\|\mathbf{c}^{i,j}\|_2^2$

    Select $\mathbf{c}_A^{i,j}$ from $\mathbf{c}^{i,j}$, and $\boldsymbol{\Psi}_A$ from $\boldsymbol{\Psi}$.

    Make SVD decomposition: $\boldsymbol{\Psi}_A = \check{U} D \check{V}'$.

    Compute the Dof according to the equation (2.3.11);

    Obtain the residual: $r = \|\mathbf{d} - \boldsymbol{\Psi}_A \mathbf{c}_A^{i,j}\|_2^2$

    The $\mathrm{BIC}^{i,j}$ or $\mathrm{AIC}^{i,j}$ is derived by equation (2.3.9);

$\tilde{\lambda}_{1,2} = \underset{i,j}{\mathrm{argmin}}\{\mathrm{BIC}^{i,j}\}$

**Output:** $\mathbf{c}^{\mathbf{d}}_{\tilde{\lambda}_{1,2}}, \tilde{\lambda}_1, \tilde{\lambda}_2$

Despite its advantage in numerical implementation, the FISTA has also left significant room to further improve the resulting accuracy and performance. For this, we resort to the newly proposed iterative-rotation approach to address this challenge. The concept of the rotation plays a central role in this paper. The main idea behind this approach [46] shows that when determining the coefficients of Hermite chaos expansion for QoI of interest, a few of the newly derived coordinates from rotating the original random inputs can have a significant impact on QoI, thereby increasing the sparsity of solution and in turn, the accuracy of recovery. These excellent features exactly cater to the demand of PC-EnKF, therefore, are especially suitable for the current study. However, considering the application in the framework of PC-EnKF, two issues need to be addressed. One is the selection of an appropriate number of rotations, which can be realized through a curve of the sparsity versus the number of iterations. And the other is the construction of the gradient matrix, which has an important effect on the computational results. Here the gradient matrix must be constructed so as to devote to the uncertainty propagation through a multi-parameter input and multi-output model that occurs in the framework of PC-EnKF for the inverse solution, rather than a model exclusively focused on propagating a scalar uncertainty forward [46]. Next, let us review briefly the basic concepts and idea related to the iterative-rotation approach. More details please see [46].

As mentioned in Section 2.1, where the output variable $\mathbf{d}$ will produce large uncertainty due to the priori distribution of the input parameters $\mathbf{m}$. From the beginning of each data assimilation iteration, assume that we have two truncated series below for the input and output random vectors $\mathbf{m}$ and $\mathbf{d}$

$$\mathbf{m} \approx \sum_{i=0}^{P} \mathbf{c}_i^{\mathbf{m}} \Psi_i(\xi), \tag{2.3.12}$$

$$\mathbf{d} \approx \sum_{i=0}^{P} \mathbf{c}_i^{\mathbf{d}} \Psi_i(\xi), \tag{2.3.13}$$

where $\xi$, as stated in Section 2.1, is a random vector comprising a set of independent random variables with given normal distribution, correspondingly, the $\Psi(\xi)$ is selected as Hermite PC basis function. PC expansion representation of input parameters (2.3.12) is set according to the prior distribution of $\mathbf{m}$, whereas the coefficients in Eq. (2.3.13) are obtained by the process mentioned in Section 2.3. For more details about Eqs. (2.3.12) and (2.3.13), please refer to reference [1]. It is also noted that when the correlated input parameters are encountered, the Karhunen–Loève expansion (KLE) can be adopted to represent them with uncorrelated random variables. This technique is also known

as proper orthogonal decomposition (POD) or, in finite-dimensional setting, principal component analysis (PCA). We will see later this situation in the numerical experiment in Section 3.

The most important step of iterative-rotation approach is performed by finding a new set of random variables. This can be accomplished by searching for a linear mapping

$$\eta = g(\xi) = \hat{U}\xi, \tag{2.3.14}$$

where $\hat{U}$ is the rotation matrix given by the eigenvalue decomposition of a gradient matrix $G$. We now define the gradient matrix:

$$\begin{aligned} G &\stackrel{\text{def}}{=} E\{J_{\mathbf{m}}^T J_{\mathbf{m}} + J_{\mathbf{d}}^T J_{\mathbf{d}}\}, \\ &= \sum_{i=1}^{S_1} E\{\nabla m_i \cdot \nabla {m_i}^T\} + \sum_{j=1}^{S_2} E\{\nabla d_j \cdot \nabla {d_j}^T\} \end{aligned} \tag{2.3.15}$$

where $\mathbf{m} = (m_1, m_2, \ldots, m_{S_1})$, $\nabla m_i = (\frac{\partial m_i}{\partial \xi_1}, \frac{\partial m_i}{\partial \xi_2}, \ldots, \frac{\partial m_i}{\partial \xi_N})^T$, and the same expression and operator also hold for the vector $\mathbf{d}$. $J_{\mathbf{m}} \subset R^{S_1 \times N}$ and $J_{\mathbf{d}} \subset R^{S_2 \times N}$ are Jacobian matrix of input parameter vector $\mathbf{m}$ and output vector $\mathbf{d}$ with respect to the random vector $\xi$, respectively. Unless otherwise stated, the dot notation represents the outer product of vector that returns a matrix. And the choices of $S_1$ and $S_2$ depend the dimension of input parameter vector and the locations of spatial measure points, respectively. When we have Eqs. (2.3.12) and (2.3.13) at hand, the gradient matrix (2.3.15) can be approximated in real computation by the following relation

$$\begin{aligned} G \approx &\sum_{i=1}^{S_1} E\left\{\nabla\left(\sum_{n=0}^{P}(c^{\mathbf{m}})_n^i \Psi_n(\xi)\right)\cdot\nabla\left(\sum_{n'=0}^{P}(c^{\mathbf{m}})_{n'}^i \Psi_{n'}(\xi)\right)^T\right\} \\ &+ \sum_{j=1}^{S_2} E\left\{\nabla\left(\sum_{n=0}^{P}(c^{\mathbf{d}})_n^j \Psi_n(\xi)\right)\cdot\nabla\left(\sum_{n'=0}^{P}(c^{\mathbf{d}})_{n'}^j \Psi_{n'}(\xi)\right)^T\right\} \end{aligned} \tag{2.3.16}$$

The entries of $G$ can be approximated as

$$\begin{aligned} G_{st} \approx &\sum_{i=1}^{S_1} E\left\{\frac{\partial}{\partial\xi_s}\left(\sum_{n=0}^{P}(c^{\mathbf{m}})_n^i \Psi_n(\xi)\right)\frac{\partial}{\partial\xi_t}\left(\sum_{n'=0}^{P}(c^{\mathbf{m}})_{n'}^i \Psi_{n'}(\xi)\right)\right\} \\ &+ \sum_{j=1}^{S_2} E\left\{\frac{\partial}{\partial\xi_s}\left(\sum_{n=0}^{P}(c^{\mathbf{d}})_n^i \Psi_n(\xi)\right)\frac{\partial}{\partial\xi_t}\left(\sum_{n'=0}^{P}(c^{\mathbf{d}})_{n'}^i \Psi_{n'}(\xi)\right)\right\} \\ &= \sum_{i=1}^{S_1}\sum_{n=0}^{P}\sum_{n'=0}^{P}(c^{\mathbf{m}})_n^i (c^{\mathbf{m}})_{n'}^i E\left\{\frac{\partial\Psi_n(\xi)}{\partial\xi_s}\frac{\partial\Psi_{n'}(\xi)}{\partial\xi_t}\right\} \\ &+ \sum_{j=1}^{S_2}\sum_{n=0}^{P}\sum_{n'=0}^{P}(c^{\mathbf{d}})_n^j (c^{\mathbf{d}})_{n'}^j E\left\{\frac{\partial\Psi_n(\xi)}{\partial\xi_s}\frac{\partial\Psi_{n'}(\xi)}{\partial\xi_t}\right\} \\ &= \sum_{i=1}^{S_1}[(\mathbf{c}^{\mathbf{m}})^i]^T K_{st}(\mathbf{c}^{\mathbf{m}})^i + \sum_{j=1}^{S_2}[(\mathbf{c}^{\mathbf{d}})^j]^T K_{st}(\mathbf{c}^{\mathbf{d}})^j \end{aligned} \tag{2.3.17}$$

where $K_{st}$ is stiffness matrix with entries

$$(K_{st})_{kl} = E\left\{\frac{\partial\Psi_k(\xi)}{\partial\xi_s}\frac{\partial\Psi_l(\xi)}{\partial\xi_t}\right\}$$

Then rotation matrix $\hat{U}$ is therefore obtained by the following eigenvalue decomposition

$$G = \hat{U}\Sigma\hat{U}', \ \hat{U}\hat{U}' = I$$

where the prime ‘’ represents transpose. For more algorithmic details, please see Algorithm 3.

**Algorithm 3** Iterative-rotation algorithm

**Input:** $\{c^m\}, \{c_i^d\}$

1: **Initialize:** $l = 0$, $\eta^{(0)} = \xi$, $c^{(0)} = [c_1^d, c_2^d, \ldots, c_{S_2}^d]$
2: Calculate gradient matrix $G^{l+1} = \sum_{i=1}^{S_2} (c_i^d)' K c_i^d$ according to (2.3.17), where $K$ is the stiffness matrix.
3: Performing eigenvalue decomposition: $G^{l+1} = \hat{U}^{l+1} \Lambda^{l+1} (\hat{U}^{l+1})'$.
4: Make transform $\eta^{l+1} = \hat{U}^{l+1} \eta^l$, and obtain measure matrix $\Psi_{ij}^{l+1} = \psi_j(\eta_i^{l+1})$.
5: Use $\Psi_{ij}^{l+1}$ to calculate $\{c_i^d\}$ by elastic net algorithm.
6: Repeat 2-5, until a preset number of iterations is reached.
7: Use $\Psi_{ij}^{l+1}$ to recalculate $\{c^m\}$ by elastic net algorithm.

**Output:** Updated $\{c^m\}, \{c_i^d\}$

### 2.4. PC-EnKF inverse modeling

Now, let us go back again to the problem proposed in Section 2.1, where the output variable $\mathbf{d}$ will produce large uncertainty due to the priori distribution of the input parameters $\mathbf{m}$. We want to seek to reduce this uncertainty using PC-EnKF with the help of the observational data $\mathbf{d}^*$.

Implementation of the polynomial chaos ensemble Kalman filter (PC-EnKF) requires first estimating the predicted mean $\mu_{\mathbf{d}}$, as well as the covariance matrices $\mathbf{c_{dd}}$ and $\mathbf{c_{dm}}$. With the aid of PC expressions (2.3.12), (2.3.13), and using the orthogonal property of $\Psi_i(\xi)$, we can obtain these statistical moments from the obtained coefficients:

$$\mu_{\mathbf{d}} = \sum_{i=0}^{P} \mathbf{c_i^d} E(\Psi_i(\xi)) = \mathbf{c_0^d}, \tag{2.4.1}$$

$$\mathbf{C_{dd}} = E((\mathbf{d} - \mu_{\mathbf{d}})(\mathbf{d} - \mu_{\mathbf{d}})^T) = E((\sum_{i=1}^{P} \mathbf{c_i^d} \Psi_i(\xi))(\sum_{j=1}^{P} \mathbf{c_j^d} \Psi_j(\xi))^T) = \sum_{i=1}^{P} \mathbf{c_i^d c_i^{dT}}, \tag{2.4.2}$$

$$\mathbf{C_{dm}} = E((\mathbf{d} - \mu_{\mathbf{d}})(\mathbf{m} - \mu_{\mathbf{m}})^T) = E((\sum_{i=1}^{P} \mathbf{c_i^d} \Psi_i(\xi))(\sum_{j=1}^{P} \mathbf{c_j^m} \Psi_j(\xi))^T) = \sum_{i=1}^{P} \mathbf{c_i^d c_i^{mT}}, \tag{2.4.3}$$

From these above relations, we see that seeking the sparsity of expansion coefficients indeed facilitates the calculation of statistical moments.

Once the process above, called prediction stage, is finished, then the measurements update (analysis) stage is followed. That is to update the PC expansion representation of the parameter vector $\mathbf{m}$ using Kalman filter given the observation. The posterior mean of $\mathbf{m}$ is updated with Kalman gain $\mathbf{K} = \mathbf{C_{md}}(\mathbf{C_{dd}} + \mathbf{R})^{-1}$:

$$\mu_{\mathbf{m}|\mathbf{d}^*} := \mathbf{c_0^{mu}} = \mathbf{c_0^m} + \mathbf{K}(\mathbf{d}^* - \mathbf{c_0^d}), \tag{2.4.4}$$

where the superscript '$u$' denotes updated, and $\mathbf{R}$ is the covariance of observation error $\epsilon$. In addition, the update equation for the coefficients other than the mean (first) term:

$$\mathbf{c_j^{mu}} = \mathbf{c_j^m} - \tilde{\mathbf{K}} \mathbf{c_j^d}, \quad j = 1, 2, \ldots, P \tag{2.4.5}$$

where the $\tilde{\mathbf{K}}$ is updated using the following formulation [1]

$$\tilde{\mathbf{K}} = \mathbf{C_{md}}((\mathbf{C_{dd}} + \mathbf{R})^{-1})^T (\sqrt{\mathbf{C_{dd}} + \mathbf{R}} + \sqrt{\mathbf{R}})^{-1}, \tag{2.4.6}$$

So the posterior covariance of the parameters

$$\mathbf{C_{mm|d^*}} := \sum_{j=1}^{P} (\mathbf{c_j^{mu}})(\mathbf{c_j^{mu}})^T \tag{2.4.7}$$

## 3. The recovery of topography in 2D shallow water equations

### 3.1. Numerical scheme

Numerical experiments will be carried out to demonstrate the efficiency and exactness of the current algorithm, and 2D shallow water equations is selected for this purpose. The shallow water models are extensively used in numerical studies of large-scale atmospheric and oceanic motions. The shallow water equations are a nonlinear hyperbolic system of partial differential equations (conservation laws for depth and momentum) that describe a fluid layer of constant density in which the horizontal scale of the flow is much greater than the layer depth. The dynamics of the single layer model is of course less general than three-dimensional models, but is often preferred because of its mathematical and computational simplicity. The 2D shallow water equations can be written in conservative form as follows:

$$\frac{\partial w}{\partial t}+\frac{\partial(uw+U)}{\partial x}+\frac{\partial(vw+V)}{\partial y}=W, \tag{3.1.1}$$

where $0<x<L, 0<y<D, t\in(0,t_f]$, $w=(h,uh,vh)^T$ is a vector function, $h(x,y,t)$ is the fluid depth, $u(x,y,t)$ and $v(x,y,t)$ are velocity components in the $x$ and $y$ directions, respectively. $U=(0,gh^2/2,0)^T$, $V=(0,0,gh^2/2)^T$, $W=(0,-gh\frac{\partial H}{\partial x},-gh\frac{\partial H}{\partial y})^T$. "$T$" denotes transpose. $g$ is the gravitational constant, and $H(x,y)$ is the topography. We assume periodic solutions in the $x$-direction $(h,u,v)|_{x=0}=(h,u,v)|_{x=L}$, while in the $y$-direction, $\frac{\partial u}{\partial y}|_{y=0}=\frac{\partial u}{\partial y}|_{y=D}=0$, $v|_{y=0}=v|_{y=D}=0$. Initially, $(h,u,v)|_{t=0}=(0.05\exp(-\frac{(x+12)^2}{15}-\frac{y^2}{12})+0.2,0,0)$.

Now let $N_x$ and $N_y$ be positive integers. We subdivide $[0,L]$ and $[0,D]$ by equidistant mesh, with the corresponding meshsize defined by $\triangle x=L/(N_x-1)$, $\triangle y=D/(N_y-1)$, respectively. As for the time interval $[0,t_f]$, we discrete it using $N_t$ equally distributed points and $\triangle t=t_f/(N_t-1)$. Note that it is possible to use adaptive mesh refinement to compute accurate solutions. However, only the uniform mesh is used in this work to avoid coding complications that might be caused by adaptive meshes. Future work will consider an adaptive mesh. Various types of numerical methods have been designed to approximate the solution of the shallow water equations. Methods such as finite volume, finite difference and discontinuous Galerkin finite element schemes are widely used [71,72]. While in this present study we use the following Lax–Wendroff scheme to integrate the shallow water Eq. (3.1.1) forward in time

$$\begin{aligned}w_{i,j}^{n+1}=w_{i,j}^{n}+\triangle t[&-\frac{(uw+U)_{i+1/2,j}^{n+1/2}-(uw+U)_{i-1/2,j}^{n+1/2}}{\triangle x}\\&-\frac{(vw+V)_{i,j+1/2}^{n+1/2}-(vw+V)_{i,j-1/2}^{n+1/2}}{\triangle y}+W_{i,j}^{n}],\end{aligned} \tag{3.1.2}$$

here $w_{i,j}^n:=w(x_i,y_j,t_n)$, and $i=1,2\ldots,N_x$, $j=1,2,\ldots,N_y$, $n=1,2,\ldots,N_t$. The scheme is second-order accurate with respect to both space and time. Unlike the Euler methods which calculate each step of a function, the Lax–Wendroff method involves first calculating a half step and then using the results from the half step to produce the full step (3.1.2). The advantage is that it does not need any artificial diffusion in time or space to keep it stable.

### 3.2. Uncertainty of input topography versus observation

We consider the shallow water equation (3.1.1) with a stochastic topography over the domain $\Omega=[0,D]\times[0,L]$, and the stochastic topography is unknown and denoted as $H(x,y,\omega)$, $(x,y)\in\Omega$ and $\omega\in\tilde{\Omega}$, where $\tilde{\Omega}$ is a sample space in a probability space $(\tilde{\Omega},\tilde{U},\tilde{P})$ with sigma algebra $\tilde{U}$ over $\tilde{\Omega}$, $\tilde{P}$ is a probability measure on $\tilde{U}$. One of the commonly used stochastic descriptions of spatial fields is based on a two-point correlation function of the spatial field. For numerical simulation, the stochastic field can be characterized by its covariance function cov$[H]$, i.e.,

$$\mathrm{cov}[H](x_1,y_1;x_2,y_2)=\sigma^2\exp(-\frac{|x_1-x_2|^2}{\lambda_x^2}-\frac{|y_1-y_2|^2}{\lambda_y^2}), \tag{3.2.1}$$

where $(x_i,y_i)(i=1,2)$ is the spatial coordinate in 2D. Here $\sigma=0.1$ is the variance and $\lambda_x=\lambda_y=12$ are the correlation length in $x$ and $y$ directions. The truncated Karhunen–Loève expansion (KLE) is used to reparameterize

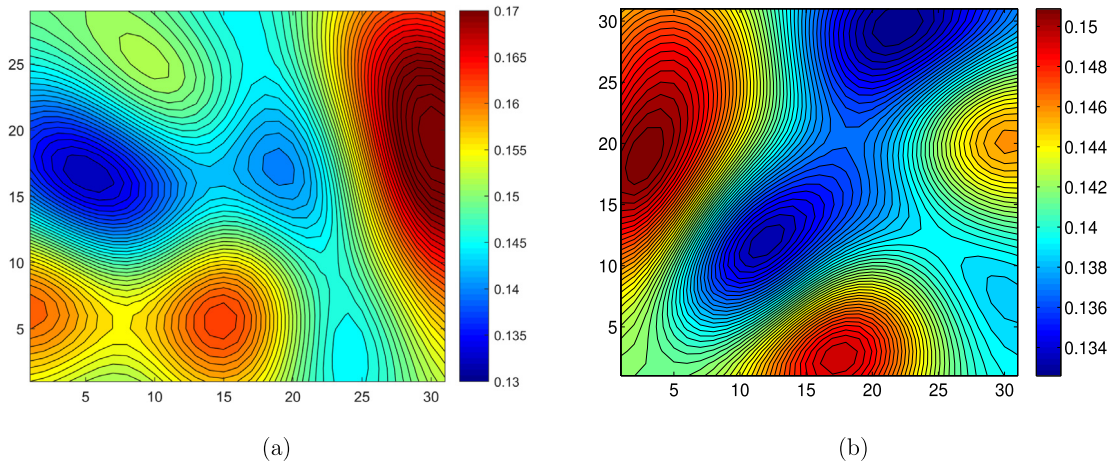


**Fig. 1.** The topography, (a) Reference, (b) initial guess.

the topography $H(x, y, \omega)$ and keep the leading $N$ terms in the KLE. For an $N$-term KLE approximation, we have the following decomposition

$$H := E[H] + \sum_{i=1}^{N} \sqrt{\lambda_i} b_i \theta_i \tag{3.2.2}$$

where $E[\cdot]$ refers to the expectation (i.e., average over all realizations). The $\lambda_i$ and $b_i$ are the eigenvalue and orthogonal eigenfunctions of the covariance function (3.2.1). The choice of $N$ is dictated by the decay rate of the eigenvalues $\lambda_i$. And the decay rate is directly related to the smoothness of cov$[H]$ and the correlation length $l_x$ and $l_y$ of the process. One typical choice of $N$ is that the sum of the neglected terms is sufficiently small compared with the sum of the first $N$ terms. The random vector $\Theta := (\theta_1, \theta_2, \ldots, \theta_N)$ is considered, where $\theta_i$ is mutually independent standard Gaussian random parameter. Thus the KLE helps to represent correlated parameters with uncorrelated random variables, which is important from the perspective of dimension reduction of input variables. Consequently, due to the input uncertainty induced by the random vector $\Theta$ contained in the topography, the output of the shallow water equations will be affected, which can be mitigated by a less uncertain estimation of the random parameters, that is, data assimilation. This can be done by gathering as more observe information of the output variables as possible to calibrate constantly the simulation results. In this numerical example, the mean of $H$ and the number of truncated terms are set as $E[H] = 0.15$ and $N = 20$, respectively. A realization of $H$ as a reference, $H_0$, is given in Fig. 1(a).

Furthermore, if the spatial domain $\Omega$ is uniformly discretized into $31 \times 31$ grid points, while the time domain $[0, t_f]$ is discretized into 2500 segments, ($t_f = 25$), a prediction result is then obtained at different times, see Fig. 2.

It is known generally that the assimilation is needed whenever significant flow behavior changes occur. In order to satisfy the assimilation requirement, a certain number of observation data will be considered. In our test, we assume that only the fluid height can be provided as an observation variable. The number of the observations threshold was attained only at $S_2 = 21$ observational locations and generated temporally at every four time-segments by running the numerical model that has been well-defined in Eqs. (3.1.2). The observational grid is as presented in Fig. 3. The observation error is set to as $\epsilon^o \sim N(\mathbf{0}, 0.001\mathbf{I})$, and $\mathbf{I}$ is a $21 \times 21$ identity matrix. The objective of the current inverse problem is to update the input topography given the limited number of observation data only for the fluid height, in the context of 2nd-order Hermite PC expansion, with total $P + 1 = \frac{(20+2)!}{20!2!} = 231$ terms. In order to estimate the PC expansion coefficients for the output variable, a set of samples will be taken. According to the suggestions (Hosder et al. 2007) [73], the number of sample points should be at least twice the number of terms in PCE. However, in our example, we conducted several experiments with different numbers of the standard Gaussian sample ensemble members, namely, $N_o = 100$, 200, 300 and 400, respectively. It is eventually shown that adopting only a total sample ensemble size of $N_o = 300$ members was sufficient to successfully perform the current data assimilation algorithm and that using a higher number of ensemble members seems to have no impact on the ensuing results.

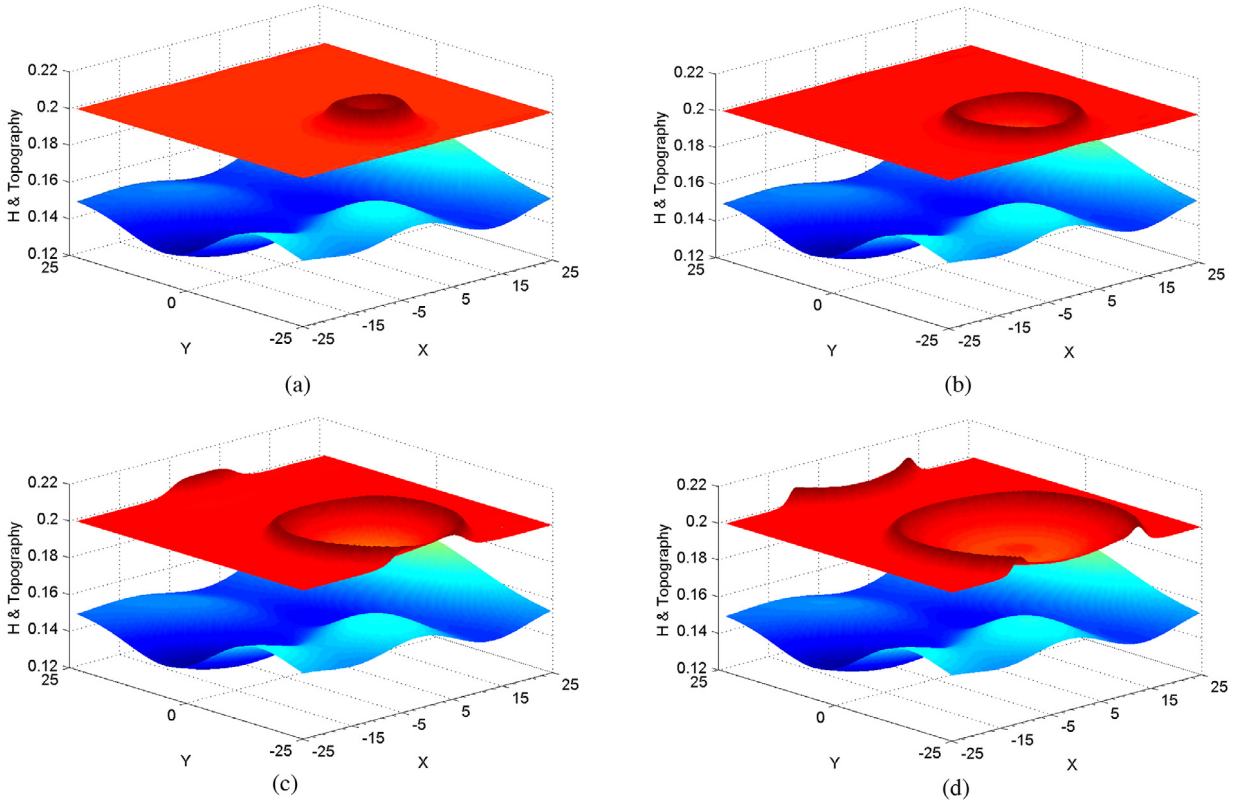


**Fig. 2.** The height $H(x, y, t, \xi)$ of modeling 2D shallow flow over the topography at time (a) $t = 4$, (b)$t = 9$, (c) $t = 14$, (d) $t = 19$, respectively. In each panel, the red is the flow height, and the blue represents the topography as depicted in Fig. 1(a). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

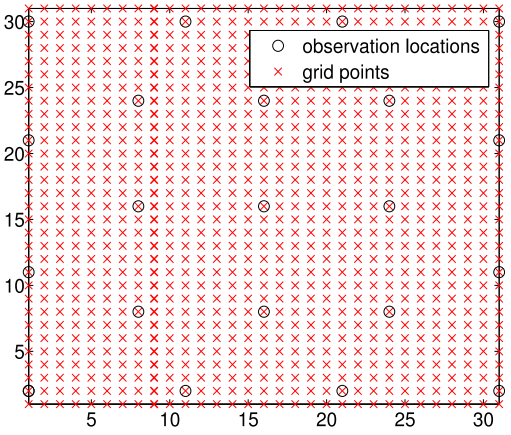


**Fig. 3.** Observational locations of the flow height.

### *3.3. Reduce the uncertainty of topography and its results*

We are concerned with the case in which the stochastic topography in 2D shallow water equation is unknown a priori to be identified. At this time, an initial guess then needs to be further designated. When the mean of $H$, according to Eq. (3.2.2), is assumed to be $E[H] = 0.14$, along with a realization of the independent standard Gaussian random

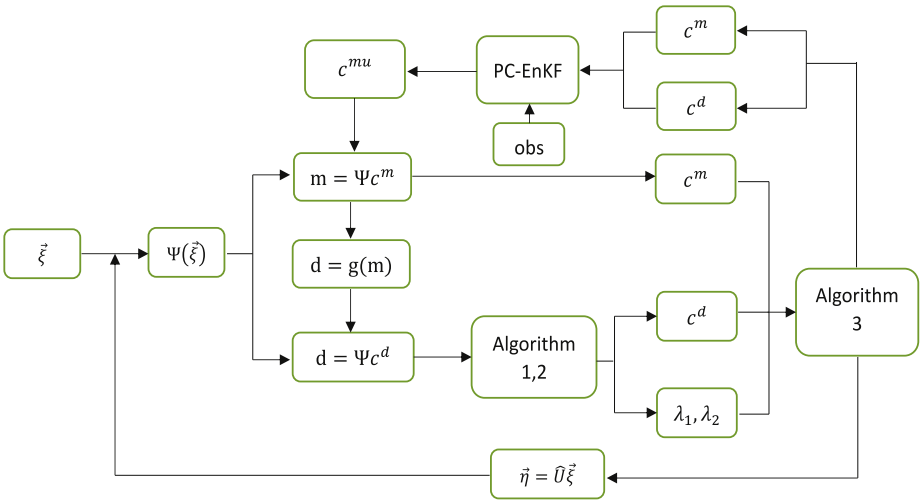


**Fig. 4.** Flow chart of the whole computational process used for the uncertainty reduction of topography, where $c^{mu}$ stands for the updated $c^m$.

variables, $\xi_i$, $(i = 1, 2, \ldots, 20)$, an initial topography is therefore derived as in Fig. 1(b). It is easy to find that there is a visually obvious difference between Fig. 1(a) and (b). And due to this difference associated with the initial guess of input topography, the model prediction will be affected. One path to uncertainty reduction of this initial topography is to calibrate the simulation results to available observations on the corresponding quantities (e.g., flow height only in this paper) of the 2D shallow water equations. This task can be completed via four consecutive assimilation loops through our method discussed in Section 2, for the whole process see Fig. 4. They are conducted sequentially (each time measurement data are available). During each of them, we start with an initial guess of topography to run the numerical model. The information is then propagated forward by use of elastic-net (EN) or elastic net with the iterative PC-basis rotation method (EN_RO) and followed by the update of topography. When these steps have been experienced, the assimilation results (updated topography) are therefore derived. These results can be used as the first guess for the next loop, more details see Fig. 5.

For further understanding of Fig. 5, an additional interpretation should be provided here. Fig. 5 shows us an illustration of the selection process of both regularization parameters and the determination of the number of the iterative rotations in each assimilation loop. $\sum_{i=1}^{21} \|c_i^{\mathbf{d}}\|_0$ represents the sparsity that concerns the total number of non-zero coefficients of polynomial chaos expansion for the model output (flow height only) at 21 space locations (see Fig. 3). Fig. 3 consists of 4 subplots that correspond to four different data assimilation loops, denoted by $L_i$, $(i = 1, 2, 3, 4)$, respectively. We take the second subplot as an example to illustrate the selection of regularization parameters and the determination of the number of the iterative rotations, which is just in the wake of the first data assimilation loop $L_1$ and is preparing for the second data assimilation $L_2$. The "black square" stands for the maximal sparsity that is selected by a minimum of a curve formed through different numbers of rotations, which indicates a desired number of rotations. And the *blue circle* represents the sparsity derived without the use of rotations. Another *blue curve* appears in a small subplot that exhibits the evolution of **BIC** with the regularization parameter $\lambda_1$ when determining the coefficients of polynomial chaos expansion during the process of performing the first rotation. The *blue square* implies the best selection of the regularization parameter $\lambda_1$.

However, it should be noted that there are several challenging problems that need to be treated. The **first** is the construction of gradient matrix used for making a rotational transformation of the random variable so that a new set of random variables are derived through the PC-basis rotation approach. This can be done by depending only on the PC expansion coefficients of the output uncertainty in a manner as described by Eq. (2.3.17). The **second** is to determine the number of iterative PC-basis rotations during each assimilation loop. This is accomplished by resorting to a curve of the sparsity versus the rotational iteration counts, for details see Fig. 4, where we select the optimal number of the iterative rotations at which the sparsest coefficient occur for the output variables. The **third** is to optimize two regularization parameters, $\lambda_1$ and $\lambda_2$, appearing in the elastic net cost function. When quantifying the output uncertainty of interest in the current study we can always deal with it as stated in Section 2.3 via **BIC**

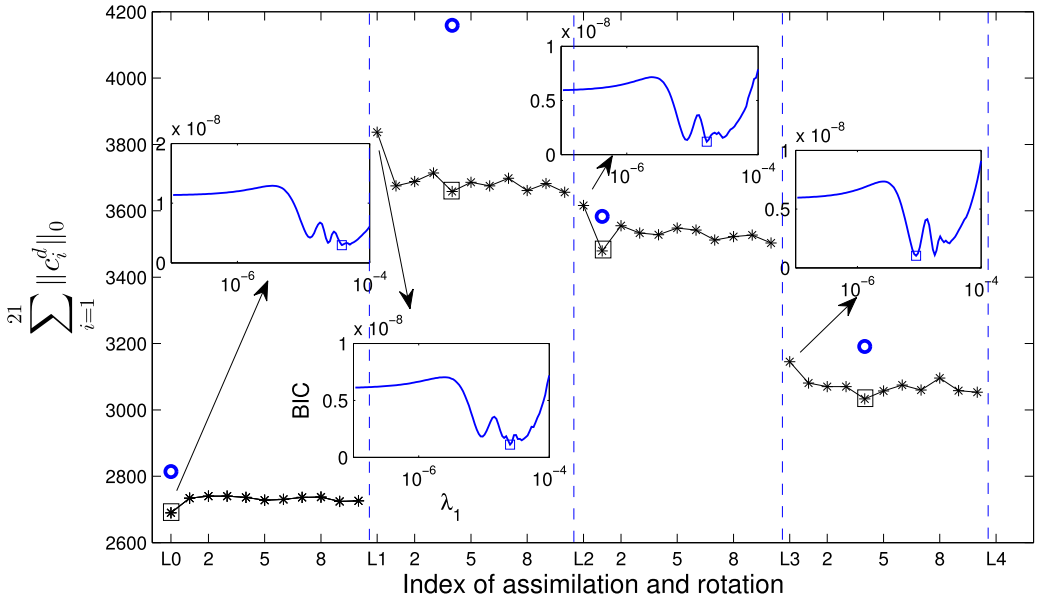


**Fig. 5.** Illustration of the selection process of both regularization parameter and the determination of the number of the iterative rotations in each assimilation loop. For further description and interpretation, please refer to Section 3.3, paragraph 2.

instead of CV to solve the coefficients of its polynomial chaos expansion. Unfortunately, the resulting computational cost is not substantially improved as expected. We attempt to ease this situation through an alternative way due to the nice property of the measurement matrix whose condition number is almost always indicated at a level of accuracy, $2.2107 \times 10^1$. In each data assimilation loop, we may develop a model such that its two regularization parameters are derived according to Algorithm 2 from the calculation of coefficients of polynomial chaos expansion for anyone of the output uncertainties of interest that requires being quantified. And the model selected like this is then applied to the rest of the output uncertainty of interest, and the subsequent iterative rotations as well, to carry out the calculation of coefficients of polynomial chaos expansion. It turns out that such a treatment causes no obvious effect on the computational results, instead, helps us to reduce considerable computational cost. In numerical test we perform a two-dimensional search over $\Lambda_1 \times \Lambda_2$ using Algorithm 2, where $\Lambda_1 = \log\text{space}(10^{-7}, 10^{-3}, 60)$ and $\Lambda_2 = [10^{-7}, 10^{-6}, 5 \times 10^{-6}, 10^{-5}, 5 \times 10^{-5}, 10^{-4}, 5 \times 10^{-4}, 10^{-3}, 5 \times 10^{-3}]$. Eventually, the model parameters pairs $(\lambda_1, \lambda_2)$ in four data assimilation loops are determined as $(2.5119 \times 10^{-5}, 5 \times 10^{-4})$, $(4.6416 \times 10^{-5}, 5 \times 10^{-5})$, $(6.3096 \times 10^{-5}, 5 \times 10^{-4})$, $(6.3096 \times 10^{-5}, 5 \times 10^{-4})$, respectively. The retrieval results of input topography are presented in Fig. 6, which are obtained by EN only and EN_RO, respectively. It is easy to see that the resulting two recoveries of the topography resemble closely the reference topography. This means that the uncertainty from the initial guess of the stochastic topography has a significant reduction, which confirms in our test example that an elastic-net based polynomial chaos ensemble Kalman filter (PC-EnKF) developed in this paper is feasible and effective.

If some alternative tools are introduced, such as the level of the sparsity denoted by $l_0$-norm ($\|\mathbf{c}\|_0$: the number of non-zero entries in the vector $\mathbf{c}$), the root mean square error (**RMSE**) and the correlation coefficients **Corr**, we then use them to further measure the difference between EN result or EN_RO result and the reference topography at the assimilation loop $n$. Thereby, we also assess how well our retrieval topography approximates the reference topography and see the benefit of the iterative PC-basis rotation method. The computational formulas of the RMSE and Corr are defined respectively as

$$\mathbf{RMSE^n} = \sqrt{\frac{\sum_{i=1}^{n_{xy}} (H_i^n - H_{0,i})^2}{n_{xy}}} \tag{3.3.1}$$

$$\mathbf{Corr(H, H_0)^n} = \frac{E((H^n - E(H))(H_0 - E(H_0)))}{\sigma_{H^n} \sigma_{H_0}} \tag{3.3.2}$$

where $\sigma_{H^n}$ is the standard deviation at the $n$th data assimilation loop, and $n_{xy}$ represents the number of spatial grids.

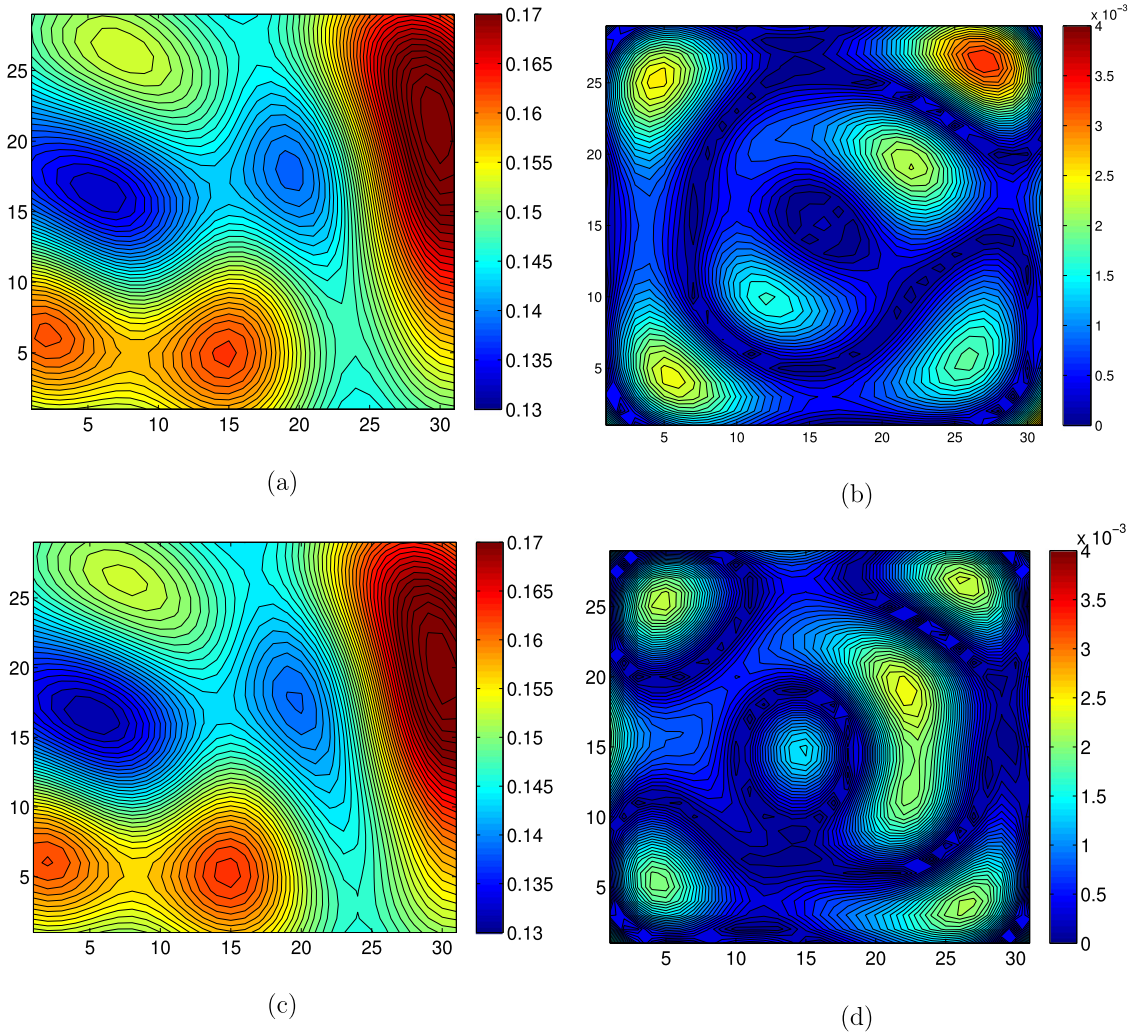


**Fig. 6.** The recovery results (a) and (c) are derived from the EN only and the EN_RO, respectively. For the sake of comparison, their absolute error between the retrieval result and the reference topography are provided in (b) and (d), respectively.

We see from Fig. 5 that the sparsity has greatly improved due to the iterative PC-basis rotation approach. The EN_RO is clearly superior to the EN in terms of the level of the sparsity of the PC coefficients for output variables. The sparsity resulting from the EN_RO not only facilitates the calculation of statistical moment used for the analysis step of PC-EnKF inverse modeling but also estimates the input topography with more accuracy compared to the EN. This fact is further confirmed by Fig. 7 from the perspective of the correlation and RMSE. So the iterative PC-basis rotation approach is advantageous in finding the sparsest possible approximation.

**Remark 1.** To further relieve the computational demands, it is possible to take into account the following two points: (1) $\lambda_2$ is fixed during the whole inverse solution; (2) At the beginning of each assimilation loop, collect each $\lambda_1$ when calculating coefficients of polynomial chaos expansion by the Algorithm 2 for the output uncertainty at every observational location, then take the mean value of them as a new regularization parameter to replace $\lambda_1$ used for the coming rotational computation. In the corresponding assimilation loop, as can be seen in Fig. 5, $\lambda_1$ is thus taken as $3.8019 \times 10^{-5}$, $2.5119 \times 10^{-5}$, $1.6596 \times 10^{-5}$, and $8.9125 \times 10^{-6}$, respectively. This way, a considerable speed-up in running time can be also gained with no potential effect on the computational results.

**Remark 2.** During the establishment of the iterative rotation method, the critical step is to construct the gradient matrix. In the current work, the construction of the gradient matrix depends only on the PC coefficients of output

**Table 1**
Comparison of CPU time required to run each algorithm.

| Algorithms | ADMM | ADMM_RO | TNIPM | TNIPM_RO | EN | EN_RO |
|---|---|---|---|---|---|---|
| CPU time (s) | 131.95 | 178.58 | 168.90 | 458.89 | 119.61 | 168.65 |

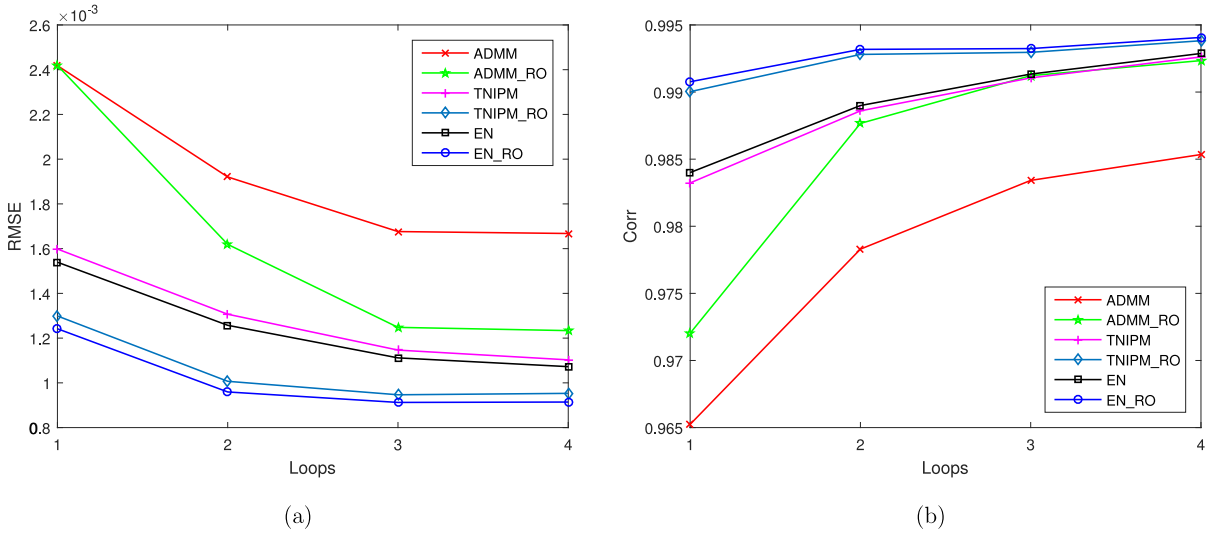


**Fig. 7.** The root mean square error (RMSE) and the correlation (Corr) between the retrieval topography and the reference topography, (a) RMSE, (b) Correlation.

variables simulated at the corresponding spatial measurement locations. Based on Eq. (2.3.15), we may have other possible ways to explore the construction of the gradient matrix. For example, the PC coefficients of input variables can also be allowed simultaneously, besides the PC coefficients of output variables. Maybe the construction of gradient matrix is problem dependent in practice. When the iterative PC-basis rotation approach is applied to a new circumstance, it is also possible to construct the gradient matrix in a way that is different from the present one. Particularly, when the uncertainty quantification (UQ) is involved, our work has shed a light on the application of the iterative rotations, which is different from the existing literature [46].

**Remark 3.** In order to further demonstrate the advantage of elastic net and the iterative rotation technique in improving the resulting accuracy, we made some comparison between our results and those from the Lasso (2.2.3) that is solved directly using the publicly available algorithms, which include the alternating direction method of multipliers (ADMM) [74] and the truncated Newton interior-point method (TNIPM) [75]. More specifically, the ADMM alternates between promoting the sparsity and minimizing the least-squares residual. It can decompose the underlying optimization problem into easily solvable modules where analytical updates can be obtained from the ridge regression and the soft-thresholding formula. And the TNIPM transforms to a Lasso problem (2.2.3) to a convex quadratic problem with linear inequality constraints. The resulting formulation is solved using a primal interior-point method with logarithmic barrier functions while invoking iterations of truncated Newton which gives a good trade-off of computational effort versus the convergence rate. These two algorithms rely on different mathematical principles, and they may perform differently in practice depending on the problem structure. It is thus valuable to investigate and compare them under problems and scenarios that we are interested in. Allowing for a fair comparison between these algorithms, we utilize the calculation of the BIC to determine the corresponding regularization parameters. The rotation technique is also considered for ADMM and TNIPM, called ADMM_RO and TNIPM_RO, respectively. We can see from Fig. 7 and Table 1 that the EN_RO method is competitive for gaining a relative good speedup for high accuracy.

## 4. Conclusions

Motivated by the novelty of the iterative PC-basis rotation approach developed recently for solving forward UQ with PC expansion approximation, we are especially interested in its feasibility in an alternative application related to the PC-EnKF high-dimensional inverse UQ. We first introduce the elastic-net cost function to improve the computational effectiveness of the polynomial chaos (PC) expansion coefficients, and adopt the fast iterative shrinkage-thresholding algorithm (FISTA) to solve the corresponding minimization problem. In order to further enhance the sparsity of PC coefficients, which may lead to a sufficiently accurate numerical result, we resort to the iterative PC-basis rotation approach. Thus an elastic-net based polynomial chaos ensemble Kalman filter (PC-EnKF) with iterative PC-basis rotations is developed for high-dimensional nonlinear inverse modeling. To guarantee the computational effectiveness, several key issues are addressed. These include the selection of regularization parameters via Bayesian information criteria (BIC), the determination of the number of iterative rotations during each assimilation loop and the construction of gradient matrix etc. When the elastic-net based polynomial chaos ensemble Kalman filter (PC-EnKF) is used for the uncertainty reduction of initial topography in 2D shallow water equations, the numerical results show that our the method is feasible and effective. Meanwhile, the better computation accuracy for the recovery of topography occurs in the case where the elastic net is coupled with the iterative PC-basis rotations. This can be seen from Fig. 7. To date, the method has only been developed and tested in the two-dimensional computational models, but the results of this study are extremely positive, which may be viewed as a first step towards solving practical inverse UQ problem by using more realistic observation data and models in order to assess the practical utility of the current method.

In the process of implementing the EN-based PC-EnKF with iterative PC-basis rotations, the non-intrusive sampling method is adopted, which leads to a repeated application of the existing or legacy deterministic solver. It will spend much more time on the large-scale problem, in turn, enhance the computational cost. This may not be an active factor for this method to make a further practical application. However, with the fast development of the reduced-order techniques [76,77], it is also possible to use the current method in the framework of the inverse problem of reduced-order model. The associated problem still needs to be further studied in depth as an interesting topic.

## Acknowledgments

We thank the anonymous reviewers who contributed in substantially improving the presentation of the manuscript. This work reported here is supported by the National Natural Science Foundation of China (Grant No. 41375115, and 61572015). The authors would also like to acknowledge the support of the Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (Grant No. ICT1600262). Lin would like to acknowledge the support from National Science Foundation (DMS-1555072, DMS-1736364, and DMS-1821233) and the National Natural Science Foundation of China (grant 51728601).

## References

[1] W.X. Li, G. Lin, D.X. Zhang, An adaptive ANOVA-based PCKF for high-dimensional nonlinear inverse modeling, J. Comput. Phys. 258 (C) (2014) 752–772.
[2] L.Z. Zeng, D.X. Zhang, A stochastic collocation based Kalman filter for data assimilation, Comput. Geosci. 14 (2010) 721–744.
[3] R.E. Kalman, A new approach to linear filtering and prediction problems, J. Basic Eng. 82 (1960) 35–45.
[4] L. Ljung, Asymptotic-behavior of the extended Kaman filter as a parameter estimator for linear-systems, IEEE Trans. Automat. Control 24 (1979) 36–50.
[5] G. Evensen, Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics, J. Geophys. Res. 99 (1994) 10143–10162.
[6] G. Evensen, P.J. Van Leeuwen, Assimilation of geosat altimeter data for the Agulhas current using the ensemble Kalman filter with a quasi-geostrophic model, Mon. Weather Rev. 124 (1) (1996) 85–96.
[7] S.I. Aanonsen, G. Naevdal, D.S. Oliver, A.C. Reynolds, B. Valles, The ensemble Kalman filter in reservoir engineering−a review, SPE J. 14 (2009) 393–412.
[8] P.L. Houtekamer, H.L. Mitchell, A sequential ensemble Kalman filter for atmospheric data assimilation, Mon. Weather Rev. 129 (1) (2001) 123–137.
[9] Y. Chen, D.X. Zhang, Data assimilation for transient flow in geologic formations via ensemble Kalman filter, Adv. Water Resour. 29 (2006) 1107–1122.
[10] R.V.D. Merwe, Sigma-point Kalman filters for probabilistic inference in dynamic state-space models, in: Proceedings of the Workshop on Advances in Machine Learning, 2004.

[11] E.A. Wan, R.V.D. Merwe, The unscented Kalman, in: Kalman Filtering and Neural Networks, John Wiley and Sons, 2002.
[12] B.V. Rosić, A. Litvinenko, O. Pajonk, H.G. Matthies, Sampling-free linear Bayesian update of polynomial chaos representations, J. Comput. Phys. 231 (17) (2012) 5761–5787.
[13] O. Pajonk, B.V. Rosić, A. Litvinenko, H.G. Matthies, A deterministic filter for non-gaussian bayesian estimation−applications to dynamical system estimation with noisy measurements, Phys. D 241 (7) (2012) 775–788.
[14] R. Ghanem, P. Spanos, Stochastic Finite Element: A Spectral Approach, Springer Verlag, 1991.
[15] D. Xiu, G.E. Karniadakis, Modeling uncertainty in flow simulations via generalized polynomial chaos, J. Comput. Phys. 187 (2003) 137–167.
[16] D. Xiu, G.E. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2002) 619–644.
[17] G. Saad, R. Ghanem, Characterization of reservoir simulation models using a polynomial chaos-based ensemble Kalman filter, Water Resour. Res. 45 (4) (2009) 546–550.
[18] J. Li, D. Xiu, A generalized polynomial chaos based ensemble Kalman filter with high accuracy, J. Comput. Phys. 228 (2009) 5454–5469.
[19] L. Zeng, H. Chang, D. Zhang, A probabilistic collocation-based Kalman filter for history matching, SPE J. 16 (2011) 294–306.
[20] T.Y. Hou, W. Luo, B. Rozovskii, H.M. Zhou, Wiener Chaos expansions and numerical solutions of randomly forced equations of fluid mechanics, J. Comput. Phys. 216 (2006) 687–706.
[21] X. Wan, G.E. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, J. Comput. Phys. 209 (2005) 617–642.
[22] Y.M. Marzouk, H.N. Najm, Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems, J. Comput. Phys. 228 (2009) 1862–1902.
[23] H.N. Najm, Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics, Annu. Rev. Fluid Mech. 41 (2009) 35–52.
[24] F. Augustin, A. Gilg, M. Paffrath, P. Rentrop, U. Wever, Polynomial chaos for the approximation of uncertainties: chances and limits, European J. Appl. Math. 19 (2008) 149–190.
[25] L. Ma, O.P. Tre, H.N. Najm, R.G. Ghanem, O.M. Knio, Multi-resolution analysis of Wiener-type uncertainty propagation schemes, J. Comput. Phys. 197 (2004) 502–531.
[26] L. Mathelin, M. Hussaini, T. Zang, Stochastic approaches to uncertainty quantification in CFD simulations, Numer. Algorithms 38 (2005) 209–236.
[27] S.S. Isukapalli, A. Roy, P.G. Georgopoulos, Efficient sensitivity/uncertainty analysis using the combined stochastic response surface method and automatic differentiation, Risk Anal. 20 (2000) 591–602.
[28] M.S. Eldred, J. Burkardt, Comparison of non-intrusive polynomial chaos and stochastic collocation methods for uncertainty quantification, in: AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, Vol. 0976, 2009, pp. 1–20.
[29] I. Babuška, R. Tempone, G. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, SIAM J. Numer. Anal. 42 (2) (2004) 800–825.
[30] A.M. Bruckstein, D.L. Donoho, M. Elad, From sparse solutions of systems of equations to sparse modeling of signals and images, SIAM Rev. 51 (1) (2009) 34–81.
[31] E.J. Candes, J. Romberg, Quantitative robust uncertainty principles and optimally sparse decompositions, Found. Comput. Math. 6 (2) (2006) 227–254.
[32] E.J. Candes, J. Romberg, Sparsity and incoherence in compressive sampling, Inverse Problems 23 (3) (2007) 969–985.
[33] M.K. Deb, I. Babuska, J.T. Oden, Solution of stochastic partial differential equations using Galerkin finite element techniques, Comput. Methods Appl. Mech. Engrg. 190 (2001) 6359–6372.
[34] A. Doostan, H. Owhadi, A non-adapted sparse approximation of PDEs with stochastic inputs, J. Comput. Phys. 230 (2011) 3015–3034.
[35] R. Tibshirani, Regression shrinkage and selection via the Lasso, J. Roy. Statist. Soc. 58 (1) (1996) 267–288.
[36] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, Least angle regression, Ann. Statist. 32 (2004) 407–499.
[37] J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for generalized linear models via coordinate descent, J. Statist. Softw. 33 (1) (2010) 1–22.
[38] R. Nowak, M. Figueiredo, Fast wavelet-based image deconvolution using the EM algorithm, in: IEEE Asilomar Conference, Vol. 1, 2001, pp. 371–375.
[39] M.A.T. Figueiredo, R.D. Nowak, A bound optimization approach to wavelet-based image deconvolution, in: IEEE International Conference on Image Processing, Vol. 2, 2005, pp. 782–785.
[40] I. Daubechies, M. Defriese, C.D. Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, Comm. Pure Appl. Math. 57 (2004) 1413–1457.
[41] M.A. Figueiredo, R.D. Nowak, S.J. Wright, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems, IEEE J. Sel. Top. Sign. Proces. 1 (4) (2007) 586–597.
[42] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problem, SIAM Imag. Sci. 2 (1) (2009) 183–202.
[43] N. Xiao, Q.S. Xu, Multi-step adaptive elastic-net: reducing false positives in high-dimensional variable selection, J. Stat. Comput. Simul. 85 (18) (2015) 3755–3765.
[44] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, J. Roy. Statist. Soc. 67 (2) (2005) 301–320.
[45] C.D. Mol, E.D. Vito, L. Rosasco, Elastic-net regularization in learning theory, J. Complexity 25 (2) (2009) 201–230.
[46] X. Yang, H. Lei, N.A. Baker, G. Lin, Enhancing sparsity of Hermite polynomial expansions by iterative rotations, J. Comput. Phys. 307 (C) (2016) 94–109.
[47] H. Lei, X. Yang, B. Zheng, G. Lin, N.A. Baker, Constructing surrogate models of complex systems with enhanced sparsity: Quantifying the influence of conformational uncertainty in biomolecular solvation, SIAM Multiscale Model. Simul. 13 (4) (2016) 1327–1353.

[48] N. Alemazkoor, H. Meidani, A near-optimal sampling strategy for sparse recovery of polynomial chaos expansions, J. Comput. Phys. 371 (2018) 137–151.
[49] R. Hilldale, D. Raff, Assessing the ability of airborne LiDAR to map river bathymetry, Earth Surf. Process. Landf. 33 (2008) 773–783.
[50] R.M. Westaway, S.N. Lane, D.M. Hicks, The development of an automated correction procedure for digital photogrammetry for the study of wide, shallow, gravel-bed rivers, Earth Surf. Process. Landf. 25 (2000) 209–225.
[51] A.F. Gessese, M. Sellier, E. Van, H. Smart, Reconstruction of river bed topography from free surface data using direct numerical approach in one-dimensional shallowwater flow, Inverse Problems 27 (2) (2011) 025001.
[52] G. Schwartz, Estimating the dimension of a model, Ann. Statist. 6 (2) (1978) 15–18.
[53] J. Shao, An asymptotic theory for linear model selection, Statist. Sinica 7 (2) (1997) 221–242.
[54] M.T. Reagan, H.N. Najm, R.G. Ghanem, O.M. Knio, Uncertainty quantification in reacting-flow simulations through non-intrusive spectral projection, Combust. Flame 132 (3) (2003) 545–555.
[55] D. Xiu, J.S. Hesthaven, High-order collocation methods for differential equations with random inputs, SIAM J. Sci. Comput. 27 (3) (2005) 1118–1139.
[56] P.G. Constantine, M. Eldred, E. Phipps, Sparse pseudospectral approximation method, Comput. Methods Appl. Mech. Engrg. 229 (2012) 1–12.
[57] G. Migliorati, F. Nobile, E.V. Schwerin, R. Tempone, Approximation of quantities of interest in stochastic pdes by the random discrete $l_2$ projection on polynomial spaces, SIAM J. Sci. Comput. 35 (3) (2013) 1440–1460.
[58] J. Peng, J. Hampton, A. Doostan, A weighted $l_1$-minimization approach for sparse polynomial chaos expansions, J. Comput. Phys. 267 (5) (2014) 92–111.
[59] J. Hampton, A. Doostan, Compressive sampling of polynomial chaos expansions: Convergence analysis and sampling strategies, J. Comput. Phys. 280 (2015) 363–386.
[60] L. Yan, L. Guo, D. Xiu, Stochastic collocation algorithms using $l_1$-minimization, Int. J. Uncertain. Quantif. 2 (3) (2012) 279–293.
[61] X. Yang G.E. Karniadakis, Reweighted $l_1$ minimization method for stochastic elliptic differential equations, J. Comput. Phys. 248 (2013) 87–108.
[62] J.D. Jakeman, M.S. Eldred, K. Sargsyan, Enhancing $l_1$-minimization estimates of polynomial chaos expansions using basis selection, J. Comput. Phys. 289 (2015) 18–34.
[63] S.S. Chen, D.L. Donoho, M. Saunders, Atomic decomposition by basis pursuit, SIAM J. Sci. Comput. 20 (1998) 33–61.
[64] A. Bruckstein, D. Donoho, M. Elad, From sparse solutions of systems of equations to sparse modeling of signals and images, SIAM Rev. 51 (2009) 34–81.
[65] A. Yang, M. Gastpar, R. Bajcsy, S. Sastry, Distributed sensor perception via sparse representation, Proc. IEEE 98 (2010) 1077–1088.
[66] T. Hastie, R. Tibshirani, J. Friedman, Elements of Statistical Learning: Data Mining, Inference and Prediction, Springer, Berlin, 2009.
[67] Y.E. Nesterov, A method for solving the convex programming problem with convergence rate $O(1/k^2)$, Dokl. Akad. Nauk SSSR 269 (1983) 543–547.
[68] H. Akaike, Information theory and an extension of the maximum likelihood principle, in: International Symposium on Information Theory, 1973, pp. 610–624.
[69] H. Zou, T. Hastie, R. Tibshirani, On the degrees of freedom of the Lasso, Ann. Statist. 35 (5) (2007) 2173–2192.
[70] R.J. Tibshirani, J. Taylor, Degrees of freedom in Lasso problems, Ann. Statist. 40 (2) (2012) 1198–1232.
[71] R. LeVeque, D. George, M. Berger, Tsunami modeling with adaptively refined finite volume methods, Acta Numer. 20 (2011) 211–289.
[72] A. Duran, F. Marche, Recent advances on the discontinuous Galerkin method for shallow water equations with topography source terms, Comput. & Fluids 101 (2014) 88–104.
[73] S. Hosder, R.W. Walters, M. Balch, Efficient sampling for non-intrusive polynomial chaos applications with multiple uncertain input variables, in: 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, 2007.
[74] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends Mach. Learn. 3 (2010) 1–122.
[75] S.J. Kim, K. Koh, M. Lustig, S. Boyd, D. Gorinevsky, An interior-point method for large-scale l1-regularized least squares, IEEE J. Sel. Top. Sign. Proces. 1 (4) (2007) 606–617.
[76] Y.P. Wang, I.M. Navon, X.Y. Wang, Y. Cheng, 2D Burgers equation with large Reynolds number using POD/DEIM and calibration, Internat. J. Numer. Methods Fluids 82 (12) (2016) 909–931.
[77] F. Fang, C.C. Pain, I.M. Navon, M.D. Piggott, G.J. Piggott, P. Allison, A.J.H. Goddard, A POD reduced order unstructured mesh ocean modelling method for moderate Reynolds number flows, Ocean Modell. 28 (2009) 127–136.