Int. J. Appl. Math. Comput. Sci., 2017, Vol. 27, No. 2, 229–243
DOI: 10.1515/amcs-2017-0016

A DYNAMIC BI–ORTHOGONAL FIELD EQUATION APPROACH TO EFFICIENT BAYESIAN INVERSION

PIYUSH M. TAGADE [a], HAN-LIM CHOI [b,*]

[a] Samsung Advanced Institute of Technology
Samsung R&D Institute India, Bangalore 560 037, India
e-mail: piyush.tagade@gmail.com

[b] Department of Aerospace Engineering
KAIST, 291 Daehak-ro, Yuseong, Deajeon 34141, Republic of Korea
e-mail: hanilmc@kaist.ac.kr

This paper proposes a novel computationally efficient stochastic spectral projection based approach to Bayesian inversion of a computer simulator with high dimensional parametric and model structure uncertainty. The proposed method is based on the decomposition of the solution into its mean and a random field using a generic Karhunen–Loève expansion. The random field is represented as a convolution of separable Hilbert spaces in stochastic and spatial dimensions that are spectrally represented using respective orthogonal bases. In particular, the present paper investigates generalized polynomial chaos bases for the stochastic dimension and eigenfunction bases for the spatial dimension. Dynamic orthogonality is used to derive closed-form equations for the time evolution of mean, spatial and the stochastic fields. The resultant system of equations consists of a partial differential equation (PDE) that defines the dynamic evolution of the mean, a set of PDEs to define the time evolution of eigenfunction bases, while a set of ordinary differential equations (ODEs) define dynamics of the stochastic field. This system of dynamic evolution equations efficiently propagates the prior parametric uncertainty to the system response. The resulting bi-orthogonal expansion of the system response is used to reformulate the Bayesian inference for efficient exploration of the posterior distribution. The efficacy of the proposed method is investigated for calibration of a 2D transient diffusion simulator with an uncertain source location and diffusivity. The computational efficiency of the method is demonstrated against a Monte Carlo method and a generalized polynomial chaos approach.

Keywords: Bayesian framework, stochastic partial differential equation, Karhunen–Loève expansion, generalized polynomial chaos, dynamically biorthogonal field equations.

1. Introduction

Recent advances in digital technologies have facilitated the use of computer simulators for investigation of large scale systems. However, computer simulators are fraught with uncertainties due to poorly known/unknown models, parameters, initial and boundary conditions etc. (Oreskes *et al.*, 1994). Various researchers have investigated the effect of these uncertainties on the credibility of a computer simulator and established uncertainty quantification and calibration as an integral aspect of a modeling and simulation process (Mehta, 1991; 1996; Oberkampf *et al.*, 2002, Trucano *et al.*, 2006; Janiszowski and Wnuk, 2016). This paper focuses on the Bayesian approach that provide a formal framework to identify, characterize and quantify the uncertainties, and provides a generic inference method for calibration of a computer simulator using limited and noisy experimental data (Kennedy and O'Hagan, 2001; Higdon *et al.*, 2005; Goldstein and Rougier, 2005; Bayarri *et al.*, 2007; Kelly and Smith, 2009; Zaidi *et al.*, 2012; Tagade and Sudhakar, 2011; Tagade *et al.*, 2009).

The Bayesian inversion framework is preferred to more traditional calibration methods due to its ability to provide complete posterior statistics of the parameters of interest. Sampling techniques such as the Markov chain Monte Carlo (MCMC) method (Besag *et al.*, 1995; Gamerman and Lopes, 2006) are used for exploration of the posterior statistics, especially for calibration of

*Corresponding author

nonlinear dynamical simulators in non-Gaussian settings. Satisfactory approximation of the posterior distribution and associated statistics using MCMC requires evaluation of the simulator at a large number of input settings, often in the range of 10^3 to 10^6. A collection of a large number of samples becomes computationally prohibitive for simulation of a large scale system, imposing a key challenge for implementation of the Bayesian framework. To make the Bayesian inversion framework accessible to large-scale problems, it is necessary to develop computationally efficient uncertainty propagation and calibration techniques.

An efficient implementation of the Bayesian inversion can be achieved by reducing the computation cost of MCMC sampling or forward nonlinear dynamical simulators (Kamiński, 2015). The methods proposed in the literature for efficient MCMC sampling include rejectionless MCMC sampling (Bortz *et al.*, 1975), adaptive Metropolis–Hastings algorithm (Gilks *et al.*, 1998), and Metropolis within Gibbs sampling (Cai *et al.*, 2008). See the work of Cotter *et al.* (2013) for a discussion on some of the recent modified MCMC sampling methods for dynamical simulators, while application of one of the methods to the variational data assimilation is discussed elsewhere (Cotter *et al.*, 2012).

Schwab and Stuart (2012) have proposed a sparse deterministic approximation (Bieri and Schwab, 2009; Schwab and Gittelson, 2011) of the posterior distribution for efficient implementation of the Bayesian inference in the context of elliptic partial differential equations. In particular, Schwab and Stuart (2012) investigate generalized polynomial chaos (gPC) expansion for approximations of the posterior. Hoang *et al.* (2013) have theoretically investigated the computational complexity of MCMC, sparse MCMC using gPC expansion and proposed a multi-level MCMC sampling. These works (Schwab and Stuart, 2012; Hoang *et al.*, 2013) demonstrate the possibility of significant computational complexity reduction for Bayesian inference using spectral projection based approximation of the forward problem.

This work concerns computationally efficient implementation of the forward dynamical simulators. In particular, the paper focuses on the spectral projection based methods. Marzouk and Najm (2007) have proposed a computationally efficient implementation of the Bayesian framework using stochastic spectral methods. Stochastic spectral projection (SSP) based methods are extensively used for uncertainty propagation as a computationally efficient alternative to Monte Carlo methods with comparable accuracy. Homogeneous chaos theory introduced by Wiener (1938; 1958) is the earliest exposition of the SSP method, where random variables are represented as an expansion series in orthogonal Hermite polynomials that converges in the mean square sense (Cameron and Martin, 1947). The present state of the art in the field of SSP based methods for uncertainty propagation is based on the generalized polynomial chaos (gPC) method. The method has been successfully implemented for solution of stochastic finite element methods (Ghanem and Spanos, 1991; 2003, Ghanem and Red-Horse, 1999) and stochastic fluid flow problems (Knio and Maitre, 2006; Xiu and Karniadakis, 2002). Xiu and Karniadakis (2003) have extended the method to a set of an Askey scheme of orthogonal polynomials. Subsequently, the method has been applied by various researchers for uncertainty propagation through simulators of systems of engineering importance (Lucor *et al.*, 2003; Mathelin *et al.*, 2004; Narayanan and Zabaras, 2004; Poette *et al.*, 2009). The Bayesian inversion formulation proposed by Marzouk and Najm (2007) uses the gPC method to propagate the prior parametric uncertainty to the simulator prediction. The resultant gPC expansion of the prediction is used in the Bayes theorem to define the likelihood. The methodology is further extended by Marzouk and Najm (2009) for inference of spatially/temporally varying uncertain parameters.

Although the gPC method offers a computationally efficient estimation of the uncertainty, the computational cost of the implementation grows significantly as the number of stochastic dimensions increases (Sapsis and Lermusiaux, 2012). Such a high dimensional uncertainty typically arises for a simulator with a large number of uncertain parameters, and more predominantly in the case of a spatially/temporally varying uncertain parameter with rapidly decaying covariance functions. The research work presented in this paper addresses the Bayesian framework for calibration of a simulator with high dimensional uncertainty.

Sapsis and Lermusiaux (2009) have proposed the dynamically orthogonal field equations (DOFEs) method for efficient propagation of high-dimensional uncertainty. The method features decomposition of the system response into a mean and stochastic dynamical component using a truncated generalized Karhunen-Loève expansion. The stochastic component is spectrally represented in terms of an orthogonal eigenfunction basis in the spatial dimension, while the respective coefficients define the time-varying stochastic dimension. The dynamic orthogonality (DO) condition (Sapsis and Lermusiaux, 2009) is used to derive the closed-form evolution equations for the mean, eigenfunction basis, and the stochastic coefficients.

However, the DOFE method requires Monte Carlo sampling for the solution of the evolution equation for stochastic coefficients. Due to Monte Carlo sampling, a parametric form of the PDF is not available. This makes it hard to directly apply the DOFE methodology for Bayesian inversion problems. This paper proposes

a dynamic bi-orthogonality based approach that extends the DOFE method for the Bayesian inversion of a computer simulator with high dimensional uncertainty. The proposed bi-orthogonal method uses the spectral expansion of the stochastic field in the gPC basis. A parametric form of the PDF of this gPC basis is available; this makes the Bayesian inference tractable using the proposed formulation.

The random coefficients of the truncated generalized Karhunen–Loève expansion, obtained using dynamically orthogonal field equations, are projected on the gPC basis using Galerkin projection (Ghanem and Spanos, 2003; Dumbser and Munz, 2007; Karczewska *et al.*, 2016). The resultant field equations are termed here as dynamically bi-orthogonal field equations (DBFEs).[1] The Bayesian inversion approach proposed in this paper uses DBFEs to project the prior parametric uncertainty to the system response. The resultant bi-orthogonal expansion is used in the likelihood during MCMC sampling for Bayesian inference of the uncertain parameters. The proposed DBFE method provides a substantial computational speedup over the gPC based method for Bayesian inversion. The efficacy of the proposed method is demonstrated for calibration of a 2D transient diffusion equation simulator with uncertain source location and the diffusivity field. Note that a preliminary version of this work was reported by Tagade and Choi (2012), while this article includes (a) extension of the method to take into account model structural uncertainty; (b) substantially refined and expanded theoretical analysis; and (c) additional numerical results demonstrating computational efficiency of the proposed methodology.

The rest of the paper is organized as follows. In Section 2, the statistical formulation is discussed in detail. The proposed DBFE-based Bayesian method is presented in Section 3. Section 4 provides numerical results for a 2D transient diffusion equation. Finally, the paper is summarized and concluded in Section 5.

2. Statistical formulation

The proposed Bayesian framework is developed for a simulator $T(\boldsymbol{x}, t, \boldsymbol{\theta}(\omega))$, where $t \in \mathbb{R}_{\geq 0}$ is time, $\boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^d$ is a spatial dimension and $\boldsymbol{\theta}(\omega)$ is a set of uncertain parameters that induce uncertainty in the predictions. Note that the simulator $T(\boldsymbol{x}, t, \boldsymbol{\theta}(\omega))$ is defined over a probability space $(\Omega, \mathcal{F}, \mathcal{P})$, where $\omega \in \Omega$, a set of elementary events, $\mathcal{F}$ is the associated σ-algebra and $\mathcal{P}$ is a probability measure defined over $\mathcal{F}$. In this paper, the proposed method is particularly developed for simulators with a model given by the partial differential equation

$$\frac{\partial u(\boldsymbol{x}, t; \omega)}{\partial t} = \mathcal{L}\left[u(\boldsymbol{x}, t; \omega); \boldsymbol{\theta}(\omega)\right], \qquad \text{(SPDE)}$$

where $u(\boldsymbol{x}, t; \omega)$ is the system response and $\mathcal{L}$ is an arbitrary differential operator. Thus, note that the simulator is

$$T(\boldsymbol{x}, t, \boldsymbol{\theta}(\omega)) = u(\boldsymbol{x}, t; \omega). \tag{1}$$

Equation (SPDE) is a stochastic partial differential equation. The stochasticity in (SPDE) emanates from the uncertainty in the parameters. (SPDE) is initialized using a random field $u(\boldsymbol{x}, 0; \omega)$, while the boundary condition is given by

$$\mathcal{B}(\boldsymbol{\beta}, t; \omega) = h(\boldsymbol{\beta}, t; \omega), \quad \boldsymbol{\beta} \in \partial\mathcal{X}, \quad \omega \in \Omega, \tag{2}$$

where $\mathcal{B}$ is a linear differential operator.

The simulator $T(\boldsymbol{x}, t, \boldsymbol{\theta}(\omega))$ approximates the physical phenomena within the limits of available knowledge. Let $\zeta(\boldsymbol{x}, t)$ be the 'true' but unknown model that perfectly represents the physical phenomena. Since $T(\boldsymbol{x}, t, \boldsymbol{\theta}(\omega))$ is an approximate representation of the physical phenomena, the simulator predictions deviate from $\zeta(\boldsymbol{x}, t)$ by $\delta(\boldsymbol{x}, t)$, where $\delta(\boldsymbol{x}, t)$ is known as a discrepancy function.[2] The relationship between $\zeta(\boldsymbol{x}, t)$ and $T(\boldsymbol{x}, t, \boldsymbol{\theta}(\omega))$ is given by (Kennedy and O'Hagan, 2001)

$$\zeta(\boldsymbol{x}, t) = T(\boldsymbol{x}, t, \hat{\boldsymbol{\theta}}(\omega)) + \delta(\boldsymbol{x}, t), \tag{3}$$

where $\hat{\boldsymbol{\theta}}(\omega)$ denotes the 'true' value of the uncertain parameters.

The proposed Bayesian framework uses experimental observations at finite locations for inference of the uncertain parameters. At time t, let the system be experimentally observed at M spatial locations $\{\boldsymbol{x}_i; i = 1, \ldots, M\}$. The measurement at $\boldsymbol{x}_i$, denoted as $y_e(\boldsymbol{x}_i, t)$, is given by

$$y_e(\boldsymbol{x}_i, t) = \zeta(\boldsymbol{x}_i, t) + \epsilon(\boldsymbol{x}_i, t), \tag{4}$$

where $\epsilon(\boldsymbol{x}_i, t)$ is the measurement uncertainty. Let $\boldsymbol{y}_e = \{y_e(\boldsymbol{x}_i, t);\ i = 1, \ldots, M\}$ be the set of available experimental observations. Using $\boldsymbol{y}_e$, the uncertain parameters $\boldsymbol{\theta}$ and the discrepancy function $\delta(\boldsymbol{x}, t)$ can be inferred through the Bayes theorem as

$$f\left(\hat{\boldsymbol{\theta}}(\omega), \delta(\boldsymbol{x}, t) \mid \boldsymbol{y}_e\right) \propto f\left(\boldsymbol{y}_e \mid T(\boldsymbol{x}, t, \hat{\boldsymbol{\theta}}(\omega)), \delta(\boldsymbol{x}, t)\right) \times f\left(\hat{\boldsymbol{\theta}}(\omega), \delta(\boldsymbol{x}, t)\right), \tag{5}$$

[1] Note that the term *dynamic bi-orthogonality* is used by other authors in related, but different, contexts (Venturi, 2011; Cheng *et al.*, 2013a; 2013b). Venturi (2011) defines dynamic bi-orthogonality in terms of dynamically evolving inner products. Cheng *et al.* (2013a; 2013b) define dynamic bi-orthogonality by decomposing the spatial dimension on an eigenfunction basis, while maintaining the independence of random coefficients of the Karhunen–Loève expansion over time evolution.

[2] In the literature $\delta(\boldsymbol{x}, t)$ is also termed model misspecification, model form uncertainty, and model structural uncertainty.

where $f(\hat{\boldsymbol{\theta}}(\omega), \delta(\boldsymbol{x},t))$ stands for the prior, $f(\boldsymbol{y}_e \mid T(\boldsymbol{x},t,\hat{\boldsymbol{\theta}}(\omega)), \delta(\boldsymbol{x},t))$ signifies the likelihood and $f(\hat{\boldsymbol{\theta}}(\omega), \delta(\boldsymbol{x},t) \mid \boldsymbol{y}_e)$ is the posterior probability distribution.

Uncertainty in the experimental observations, $\epsilon(\boldsymbol{x},t)$, is assumed to be specified using a zero mean Gaussian distribution with covariance matrix

$$\Sigma_e = \sigma_e^2 I_M, \tag{6}$$

where I is the $M \times M$ identity matrix. In this paper, the proposed Bayesian framework is developed for independent prior uncertainties in $\hat{\boldsymbol{\theta}}(\omega)$ and $\delta(\boldsymbol{x},t)$, with the prior for $\delta(\boldsymbol{x},t)$ given by a zero-mean Gaussian process with a covariance function of the form

$$\Sigma_\delta(\boldsymbol{x}_1,\boldsymbol{x}_2) = \sigma_\delta^2 \exp\left(-\sum_{k=1}^{d} \lambda_k^\delta \left(x_1^k - x_2^k\right)^2\right), \tag{7}$$

where σ_δ^2 is the variance and λ_k^δ is the correlation strength of the covariance function, which are treated as uncertain hyper-parameters. In the present paper, the inverse Gamma prior $IG(\alpha_\sigma, \beta_\sigma)$ is used for σ_δ^2, while, the Gamma prior $G(\alpha_{\lambda_k}, \beta_{\lambda_k})$ is used for λ_k^δ (Kennedy and O'Hagan, 2001; Paulo, 2005; O'Hagan, 2006). The uncertainty in $\delta(\boldsymbol{x},t)$ is specified using the hierarchical zero-mean Gaussian process prior as

$$\begin{aligned} &f(\delta(\boldsymbol{x},t), \sigma_\delta^2, \boldsymbol{\lambda^\delta}) \\ &\propto \frac{\exp\left(-\frac{1}{2}\boldsymbol{\delta}^T \Sigma_\delta^{-1} \boldsymbol{\delta}\right)}{\sqrt{|\Sigma_\delta|}} \times \frac{\exp\left(-\beta_\sigma/\sigma_\delta^2\right)}{(\sigma_\delta^2)^{\alpha_\sigma+1}} \\ &\quad \times \prod_{k=1}^{d} (\lambda_k i)^{\alpha_{\lambda_k i}-1} \exp\left(-\beta_{\lambda_k}\lambda_k^\delta\right) \end{aligned} \tag{8}$$

where $\boldsymbol{\delta} = \{\delta(\boldsymbol{x}_i,t);\ i = 1,\ldots,N\}$ and $\boldsymbol{\lambda^\delta} = \{\lambda_k^\delta;\ k = 1,\ldots,d\}$. Use (8) and (6) in the Bayes theorem (5) to obtain

$$\begin{aligned} &f\left(\hat{\boldsymbol{\theta}}(\omega), \boldsymbol{\delta}, \sigma_\delta^2, \boldsymbol{\lambda_\delta} \mid \boldsymbol{y_e}\right) \\ &\propto \frac{\exp\left(-\frac{1}{2}\boldsymbol{d}^T \Sigma_e^{-1} \boldsymbol{d}\right)}{\sqrt{|\Sigma_e|}} \frac{\exp\left(-\frac{1}{2}\boldsymbol{\delta}^T \Sigma_\delta^{-1} \boldsymbol{\delta}\right)}{\sqrt{|\Sigma_\delta|}} \\ &\quad \times \frac{\exp\left(-\beta_\sigma/\sigma_\delta^2\right)}{(\sigma_\delta^2)^{\alpha_\sigma+1}} \\ &\quad \times \left[\prod_{k=1}^{d} (\lambda_k i)^{\alpha_{\lambda_k i}-1} \exp\left(-\beta_{\lambda_k}\lambda_k^\delta\right)\right] \times f(\hat{\boldsymbol{\theta}}(\omega)), \end{aligned} \tag{9}$$

where $\boldsymbol{d} = \{y_e(\boldsymbol{x}_i,t) - \left(T(\boldsymbol{x}_i,t,\hat{\boldsymbol{\theta}}(\omega)) + \delta(\boldsymbol{x}_i,t)\right);\ i = 1,\ldots,M\}$. Marginalization of $\delta(\boldsymbol{x},t)$ from (9) gives

$$\begin{aligned} &f(\hat{\boldsymbol{\theta}}(\omega), \sigma_\delta^2, \boldsymbol{\lambda_\delta} \mid \boldsymbol{y_e}) \\ &\propto \frac{\exp\left(-\frac{1}{2}\boldsymbol{\eta}^T \Sigma^{-1} \boldsymbol{\eta}\right)}{\sqrt{|\Sigma|}} \times \frac{\exp\left(-\beta_\sigma/\sigma_\delta^2\right)}{(\sigma_\delta^2)^{\alpha_\sigma+1}} \\ &\quad \times \prod_{k=1}^{d} (\lambda_k^\delta)^{\alpha_{\lambda_k}-1} \exp\left(-\beta_{\lambda_k}\lambda_k^\delta\right) \times f(\hat{\boldsymbol{\theta}}(\omega)), \end{aligned} \tag{10}$$

where $\Sigma = \Sigma_\delta + \Sigma_e$ and we set $\boldsymbol{\eta} = \{y_e(\boldsymbol{x}_i,t) - T(\boldsymbol{x}_i,t,\hat{\boldsymbol{\theta}}(\omega));\ i = 1,\ldots,M\}$.

3. Proposed methodology

Equation (10) can be solved by sampling from the posterior distribution using MCMC, which requires evaluation of $T(\boldsymbol{x}_i,t,\hat{\boldsymbol{\theta}}(\omega))$ for each sample, which is computationally prohibitive for large-scale system simulators. The approach proposed in this paper requires a single evaluation of (SPDE) using dynamically bi-orthogonal field equations. The DBFE is used for propagating the prior uncertainty in $\hat{\boldsymbol{\theta}}(\omega)$ to the system response. The resultant bi-orthogonal expansion of the system response is used in (10) to define the posterior distribution, which is explored using MCMC. The proposed method is described in detail in this section.

3.1. Dynamically bi-orthogonal field equations. The proposed DBFE method is based on the dynamically orthogonal field equations (DOFEs) proposed by Sapsis and Lermusiaux (Sapsis and Lermusiaux, 2009). Consider a generic Karhunen–Loève expansion of $u(\boldsymbol{x},t;\omega)$ truncated at N terms as

$$u(\boldsymbol{x},t;\omega) = \overline{u}(\boldsymbol{x},t) + \sum_{i=1}^{N} Y_i(t;\omega) u_i(\boldsymbol{x},t), \tag{11}$$

where $\overline{u}(\boldsymbol{x},t)$ is the mean, $u_i(\boldsymbol{x},t)$ are the functions that form a complete orthonormal basis on $L^2(\mathcal{X})$, while $Y_i(t;\omega)$ are the zero-mean independent random variables. Note that throughout this paper the equality sign, $=$, is used to represent the approximate equality if no confusion is expected. Substituting the expansion (11) in (SPDE) gives

$$\begin{aligned} &\frac{\partial \overline{u}(\boldsymbol{x},t)}{\partial t} + \sum_{i=1}^{N} u_i(\boldsymbol{x},t) \frac{\mathrm{d}Y_i(t;\omega)}{\mathrm{d}t} \\ &\qquad + \sum_{i=1}^{N} Y_i(t;\omega) \frac{\partial u_i(\boldsymbol{x},t)}{\partial t} \\ &\qquad = \mathcal{L}[u(\boldsymbol{x},t;\omega);\boldsymbol{\theta}(\omega)]. \end{aligned} \tag{12}$$

Equation (12) is a set of coupled partial differential equations. Due to the coupled nature of the equations,

solution of (12) is numerically intractable. This necessitates the imposition of additional constraints to ensure numerical tractability. Sapsis and Lermusiaux (2009) proposed imposition of a dynamic orthogonality (DO) condition to derive the independent evolution equations for $\overline{u}(\boldsymbol{x},t)$, $Y_i(t;\omega)$ and $u_i(\boldsymbol{x},t)$. The DO condition constrains the time evolution of $u_i(\boldsymbol{x},t)$ such that

$$\left\langle \frac{\partial u_i(\boldsymbol{x},t)}{\partial t}, u_j(\boldsymbol{x},t) \right\rangle_{\boldsymbol{X}} = 0, \quad \forall i,j = 1,\ldots,N, \tag{13}$$

where $\langle\cdot,\cdot\rangle$ is the inner product. Note that the DO condition ensures that $u_i(\boldsymbol{x},t)$ preserves orthonormality over the time evolution of (SPDE).

Remark 1. In this paper, the inner product is defined over spatial and stochastic dimensions. The inner product over the spatial dimension is defined as

$$\langle u(\boldsymbol{x},t;\omega), v(\boldsymbol{x},t;\omega)\rangle_{\boldsymbol{X}} = \int_{\boldsymbol{X}} u(\boldsymbol{x},t;\omega)v(\boldsymbol{x},t;\omega)\,\mathrm{d}\boldsymbol{x}, \tag{14}$$

while the inner product over the stochastic dimension is defined as

$$\langle u(\boldsymbol{x},t;\omega), v(\boldsymbol{x},t;\omega)\rangle_{\Omega} = \int_{\Omega} u(\boldsymbol{x},t;\omega)v(\boldsymbol{x},t;\omega)\,\mathrm{d}\mathcal{P}(\omega). \tag{15}$$

Using the DO condition, the independent evolution equations for $\overline{u}(\boldsymbol{x},t)$, $u_i(\boldsymbol{x},t)$ and $Y_i(t;\omega)$ are derived as follows (Sapsis and Lermusiaux, 2009).

3.1.1. Dynamically orthogonal field equations. Apply the expectation operator to (12) to obtain the evolution equations for $\overline{u}(\boldsymbol{x},t)$ as

$$\frac{\partial \overline{u}(\boldsymbol{x},t)}{\partial t} = E^{\omega}\left[\mathcal{L}[u(\boldsymbol{x},t;\omega);\boldsymbol{\theta}(\omega)]\right]. \tag{16}$$

Multiply (12) by $Y_j(t;\omega)$ and apply the expectation operator to have

$$\sum_{i=1}^{N} C_{Y_i(t)Y_j(t)} \frac{\partial u_i(\boldsymbol{x},t)}{\partial t} + \sum_{i=1}^{N} C_{\frac{\mathrm{d}Y_i(t;\omega)}{\mathrm{d}t}Y_j(t;\omega)} u_i(\boldsymbol{x},t) = E^{\omega}\left[\mathcal{L}\left[u(x,t;\omega);\boldsymbol{\theta}(\omega)\right]Y_j(t;\omega)\right], \tag{17}$$

where $C_{Y_i(t)Y_j(t)}$ denotes the covariance between $Y_i(t;\omega)$ and $Y_j(t;\omega)$. Multiplying (17) by $u_k(\boldsymbol{x},t)$, taking the inner product and applying the expectation operator gives

$$C_{\frac{\mathrm{d}Y_k(t)}{\mathrm{d}t}Y_j(t)} = \left\langle E^{\omega}\left[\mathcal{L}\left(u(x,t;\omega);\omega\right)Y_j(t;\omega)\right], u_k(x,t)\right\rangle_{\boldsymbol{X}}, \tag{18}$$

which on substitution in (17) yields

$$\begin{aligned}\sum_{i=1}^{N} C_{Y_i(t)Y_j(t)} \frac{\partial u_i(\boldsymbol{x},t)}{\partial t} &= E^{\omega}\left[\mathcal{L}[u(\boldsymbol{x},t,\omega);\boldsymbol{\theta}(\omega)]Y_j(t;\omega)\right] \\ &\quad - \sum_{k=1}^{N} \left\langle E^{\omega}\left[\mathcal{L}[u(\boldsymbol{x},t,\omega);\boldsymbol{\theta}(\omega)]Y_j(t;\omega)\right], u_k(\boldsymbol{x},t)\right\rangle_{\boldsymbol{X}} \\ &\quad \times u_k(\boldsymbol{x},t).\end{aligned} \tag{19}$$

Equation (19) can be written in matrix form

$$\mathbf{U} = \boldsymbol{\Gamma}^{-1}\mathbf{D}, \tag{20}$$

where $\boldsymbol{\Gamma}$ is the covariance matrix with the (i,j)-th element $\Sigma_{ij} = C_{Y_i(t)Y_j(t)}$.

To derive the evolution equation for $Y_j(t;\omega)$, multiply both the sides of (12) by $u_j(\boldsymbol{x},t)$ and take the inner product to obtain

$$\begin{aligned}&\left\langle \frac{\partial \overline{u}(\boldsymbol{x},t)}{\partial t}, u_j(\boldsymbol{x},t)\right\rangle_{\boldsymbol{X}} + \sum_{i=1}^{N} \langle u_i(\boldsymbol{x},t), u_j(\boldsymbol{x},t)\rangle_{\boldsymbol{X}} \frac{\mathrm{d}Y_i(t;\omega)}{\mathrm{d}t} \\ &\quad + \sum_{i=1}^{N} Y_i(t;\omega)\left\langle \frac{\partial u_i(\boldsymbol{x},t)}{\partial t}, u_j(\boldsymbol{x},t)\right\rangle_{\boldsymbol{X}} \\ &\qquad = \left\langle \mathcal{L}[u(\boldsymbol{x},t;\omega);\boldsymbol{\theta}(\omega)], u_j(\boldsymbol{x},t)\right\rangle_{\boldsymbol{X}}.\end{aligned} \tag{21}$$

Note that the third term on the left-hand side in (21) vanishes–completely due to the DO condition (13), while the second term vanishes for all $i \neq j$ owing to the orthonormality of $u_i(\boldsymbol{x},t)$. Thus

$$\frac{\mathrm{d}Y_i(t;\omega)}{\mathrm{d}t} + \left\langle \frac{\partial \boldsymbol{u}(\boldsymbol{x},t)}{\partial t}, u_i(\boldsymbol{x},t)\right\rangle_{\boldsymbol{X}} = \left\langle \mathcal{L}[u(\boldsymbol{x},t;\omega);\boldsymbol{\theta}(\omega)], u_i(\boldsymbol{x},t)\right\rangle_{\boldsymbol{X}}. \tag{22}$$

Note that multiplying (16) by $u_i(\boldsymbol{x},t)$ and taking the inner product gives

$$\left\langle \frac{\partial \overline{u}(\boldsymbol{x},t)}{\partial t}, u_i(\boldsymbol{x},t)\right\rangle_{\boldsymbol{X}} = \left\langle E^{\omega}\left[\mathcal{L}[u(\boldsymbol{x},t;\omega);\boldsymbol{\theta}(\omega)]\right], u_i(\boldsymbol{x},t)\right\rangle_{\boldsymbol{X}}. \tag{23}$$

Using (23) in (22) gives the evolution equation for $Y_i(t;\omega)$ as

$$\begin{aligned}\frac{\mathrm{d}Y_i(t;\omega)}{\mathrm{d}t} &= \Big\langle \mathcal{L}\left[u(\boldsymbol{x},t;\omega);\boldsymbol{\theta}(\omega)\right] - E^{\omega}\left[\mathcal{L}\left[u(\boldsymbol{x},t;\omega);\boldsymbol{\theta}(\omega)\right]\right], \\ &\qquad u_i(\boldsymbol{x},t)\Big\rangle_{\boldsymbol{X}}.\end{aligned} \tag{24}$$

3.1.2. Bi-orthogonal expansion. Note that the numerical solution of (SPDE) using the DOFE method provide the samples of $u(\boldsymbol{x},t;\omega)$ through the coefficients $Y_i(t;\omega)$, whereas the Bayesian inference requires an analytic form of the probability distribution of the system response. Thus, the DOFE method cannot directly be used for Bayesian inference. In the present paper, a bi-orthogonal expansion approach is proposed to impose the orthogonality based geometric structure on the stochastic dimension. Consider a gPC expansion of $Y_i(t;\omega)$ truncated at P terms as

$$Y_i(t;\omega)=\sum_{p=1}^{P} Y_p^i(t)\psi_p(\boldsymbol{\xi}(\omega)), \tag{25}$$

where $\psi_p(\boldsymbol{\xi}(\omega))$ are the orthogonal polynomials from the Askey scheme, while $\boldsymbol{\xi}(\omega)\in L^2(\Xi)$ are the random variables with the appropriate probability density function (Xiu and Karniadakis, 2003). Use (25) in (11) to get

$$u(\boldsymbol{x},t;\omega)=\overline{u}(\boldsymbol{x},t)+\sum_{i=1}^{N}\sum_{p=1}^{P} Y_p^i(t)\psi_p(\boldsymbol{\xi}(\omega))u_i(\boldsymbol{x},t). \tag{26}$$

Equation (26) is termed here the bi-orthogonal expansion. Differentiate (25) with respect to time and use the Galerkin projection to obtain

$$\begin{aligned}\frac{\mathrm{d}Y_p^i(t)}{\mathrm{d}t}=\frac{1}{\left\langle\psi_q^2\right\rangle_\Omega}\Big\langle\Big\langle&\mathcal{L}\left[u(\boldsymbol{x},t;\omega);\boldsymbol{\theta}(\omega)\right]\\&-E^\omega\left[\mathcal{L}\left[u(x,t;\omega);\boldsymbol{\theta}(\omega)\right]\right],\Big\rangle_{\boldsymbol{X}},\\&u_i(\boldsymbol{x},t)\psi_p(\boldsymbol{\xi}(\omega))\Big\rangle_\Omega.\end{aligned} \tag{27}$$

Note that (27) is a key innovation of this paper, which differentiates the proposed approach from the state-of-the-art DOFE method. Equation (27) allows time evolution of the gPC expansion coefficients, while retaining the gPC basis $\psi_p(\boldsymbol{\xi}(\omega))$ invariant. This allows efficient data assimilation by updating the PDF of $\boldsymbol{\xi}(\omega)$ using the Bayes theorem. Equations (16), (19) and (27) form dynamically bi-orthogonal field equations (DBFEs) that define the dynamic evolution of the mean $\overline{u}(\boldsymbol{x},t)$, the eigenfield $u_i(\boldsymbol{x},t)$ and the associated coefficients $Y_p^i(t)$. The resultant bi-orthogonal expansion (26) approximates the system response $u(\boldsymbol{x},t;\omega)$ to an arbitrary accuracy depending on the number of eigenfunctions used, N, and the number of expansion coefficients for each eigenfunction, P. Equations (16) and (19) are numerically solved using a finite difference scheme in spatial dimension and a fourth-order Runge–Kutta scheme in the temporal dimension. Equation (27) is numerically solved using a fourth-order Runge–Kutta scheme while the inner product is evaluated using a Gaussian quadrature.

3.1.3. Boundary conditions. To define boundary conditions for the DBFE, consider a generic Karhunen–Loève expansion of $h(\boldsymbol{\beta},t;\omega)$

$$h(\boldsymbol{\beta},t;\omega)=\overline{h}(\boldsymbol{\beta},t)+\sum_{i=1}^{N} Y_i(t;\omega)u_i(\boldsymbol{\beta},t). \tag{28}$$

Applying the expectation operator to (28), the boundary condition for the mean is given by

$$\mathcal{B}(\overline{u}(\boldsymbol{x},t))=\overline{h}(\boldsymbol{\beta},t). \tag{29}$$

Multiplying (28) by $Y_j(t;\omega)$ and applying the expectation operator, we obtain the boundary condition for $u_i(\boldsymbol{x},t)$

$$\mathcal{B}(u_i(\boldsymbol{\beta},t))=\sum_{j=1}^{N} C^{-1}_{Y_i(t)Y_j(t)}E^\omega\left[h(\boldsymbol{\beta},t;\omega)Y_j(t;\omega)\right]. \tag{30}$$

3.2. Bayesian inference. With no loss of generality, the proposed method is described here for a spatially varying uncertain parameter with the prior given by a scalar stochastic process $v(\boldsymbol{x};\omega)$, i.e., $\boldsymbol{\theta}(\omega)=\{v(\boldsymbol{x};\omega)\}$. Use a KL expansion of $v(\boldsymbol{x};\omega)$ as

$$v(\boldsymbol{x};\omega)=\overline{v}(\boldsymbol{x})+\sum_{i=1}^{N}\sqrt{\lambda_i^v}v_i(\boldsymbol{x})\chi_i, \tag{31}$$

where χ_i are independent identically distributed zero-mean random variables, while λ_i^v and $v_i(\boldsymbol{x})$ are respectively the eigenvalues and eigenfunctions of the covariance function of $v(\boldsymbol{x};\omega)$. For the covariance function $C_v(\boldsymbol{x}_1,\boldsymbol{x}_2)$, λ_i^v and $v_i(\boldsymbol{x})$ are a solution to the eigenvalue problem

$$\int_{\mathcal{X}} C_v(\boldsymbol{x}_1,\boldsymbol{x}_2)\nu_i(\boldsymbol{x}_1)\,\mathrm{d}\boldsymbol{x}_1=\lambda_i^v v_i(\boldsymbol{x}_2). \tag{32}$$

For a Gaussian process prior, χ_i are standard normal random variables, whereas, for a generic stochastic process prior, χ_i are given by

$$\chi_i=\frac{1}{\sqrt{\lambda_i^v}}\int_{\mathcal{X}}\left(v(\boldsymbol{x};\omega)-\overline{v}(\boldsymbol{x})\right)v_i(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{33}$$

Use the gPC expansion of χ_i

$$\chi_i=\sum_{p=1}^{P}\hat{\chi}_p^i\psi_p(\boldsymbol{\xi}(\omega)), \tag{34}$$

where $\hat{\chi}_p^i$ are the gPC expansion coefficients, in (31) to get the bi-orthogonal expansion of $v(\boldsymbol{x};\omega)$ as

$$v(\boldsymbol{x};\omega)=\overline{v}(\boldsymbol{x})+\sum_{i=1}^{N}\sum_{p=1}^{P}\sqrt{\lambda_i^v}v_i(\boldsymbol{x})\hat{\chi}_p^i\psi_p(\boldsymbol{\xi}(\omega)). \tag{35}$$

The expansion (35) is used in $\mathcal{L}\left[u(\boldsymbol{x},t;\omega);\boldsymbol{\theta}(\omega)\right]$ to define the RHS of the DBFE governing equations (16), (19) and (27). The numerical solution of (16), (19) and (27) produces the bi-orthogonal expansion (26) of the system response $u(\boldsymbol{x},t;\omega)$. Note that numerical solution of the DOFEs requires use of a Monte Carlo method, whereas the DBFEs are numerically solved using a deterministic numerical integration method. Since the DBFEs do not require Monte-Carlo sampling, the proposed method is computationally efficient compared with the DOFE method.

Remark 2. The numerical solution is initiated with the initial condition for the mean $\overline{u}(\boldsymbol{x},t)$ given by

$$\overline{u}(\boldsymbol{x},t) = E^{\omega}\left[F[u(\boldsymbol{x},0;\omega);\boldsymbol{\theta}(\omega)]\right], \tag{36}$$

while the initial conditions for the eigenfield are given by

$$u_i(\boldsymbol{x},t) = v_i(\boldsymbol{x}). \tag{37}$$

Since the stochasticity in (SPDE) emanates due to the uncertainty in $\nu(\boldsymbol{x};\omega)$, (SPDE) is initialized with a deterministic initial condition. Thus, the initial condition for the expansion coefficients $Y_p^i(t)$ is given by

$$Y_p^i(0) = 0, \quad \forall i = 1,\ldots,N, \quad p = 1,\ldots,P. \tag{38}$$

The bi-orthogonal expansion of $u(\boldsymbol{x},t;\omega)$ is used in (10) to define the likelihood

$$f(\boldsymbol{y}_e \mid \boldsymbol{\xi},\sigma_\delta^2,\boldsymbol{\lambda^\delta}) \propto \mid \Sigma \mid^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\boldsymbol{\eta}^T\Sigma^{-1}\boldsymbol{\eta}\right), \tag{39}$$

where $\boldsymbol{\eta} = \{\eta_k, \quad k = 1,\ldots,M\}$ with

$$\begin{aligned}\eta_k = y_e(\boldsymbol{x}_k,t) - \Big(&\overline{u}(\boldsymbol{x}_k,t) \\ &+ \sum_{i=1}^{N}\sum_{p=1}^{P} Y_p^i(t)u_i(\boldsymbol{x}_k,t)\psi_p(\boldsymbol{\xi}(\omega))\Big).\end{aligned} \tag{40}$$

Note that conditional on the hyper-parameters of the discrepancy function, σ_δ^2 and $\boldsymbol{\lambda^\delta}$, $\boldsymbol{\xi}$ are the only uncertain parameters in (39). Thus, the Bayesian inversion problem is reformulated in the space $L^2(\Xi)$ as the inference of $\boldsymbol{\xi}$. In the present paper, the proposed method is demonstrated for Hermite polynomials as the gPC basis, where $\boldsymbol{\xi}$ are independent identically distributed standard normal random variables. Thus, the prior for $\boldsymbol{\xi}$ is given by

$$f(\boldsymbol{\xi}) \propto \prod_{k=1}^{N_z} \exp\left(-\frac{\xi_k^2}{2}\right), \tag{41}$$

where N_z is the dimension of stochasticity. Using (41) and (39) in (10), the proposed formulation for the Bayesian inference is

$$\begin{aligned}f(\boldsymbol{\xi},\sigma_\delta^2,\boldsymbol{\lambda_\delta} \mid \boldsymbol{y_e}) &\propto \frac{\exp\left(-\frac{1}{2}\boldsymbol{\eta}^T\Sigma^{-1}\boldsymbol{\eta}\right)}{\sqrt{|\Sigma|}} \times \frac{\exp\left(-\beta_\sigma/\sigma_\delta^2\right)}{(\sigma_\delta^2)^{\alpha_\sigma+1}} \\ &\times \prod_{k=1}^{d}(\lambda_k^\delta)^{\alpha_{\lambda_k}-1}\exp\left(-\beta_{\lambda_k}\lambda_k^\delta\right) \times \prod_{k=1}^{N_z} \\ &\times \exp\left(-\frac{\xi_k^2}{2}\right).\end{aligned} \tag{42}$$

Note that (42) does not involve the solution of $T(\boldsymbol{x},t,\boldsymbol{\theta}(\omega))$. Thus, the posterior distribution can be explored efficiently using MCMC. In the present paper, the Metropolis–Hastings algorithm (Metropolis *et al.*, 1953; Hastings, 1970) is used to sample from the posterior distribution.

4. Numerical example: A 2D transient diffusion equation

The efficiency of the proposed method is investigated for calibration of a two-dimensional transient diffusion simulator with an uncertain source location and a diffusivity field. The present paper considers a stochastic transient diffusion equation defined over the two-dimensional domain $\mathcal{X} = [-1,1]\times[-1,1]$

$$\begin{aligned}\frac{\partial u(\boldsymbol{x},t;\omega)}{\partial t} &= \nabla\cdot(\nu(\boldsymbol{x};\omega)\nabla u(\boldsymbol{x},t;\omega)) \\ &+ \sum_{l=1}^{N_s}\frac{s_l}{2\pi\sigma^2}\exp\left(-\frac{(\boldsymbol{z}_l-\boldsymbol{x})^2}{2\sigma_l^2}\right)\delta_{t\in[0,T_i]},\end{aligned} \tag{43}$$

where $\nu(\boldsymbol{x};\omega)$ is the spatially varying diffusivity field, while total N_s source terms are active during time $[0,T_i]$ at locations $\boldsymbol{z}_l$ with source strengths s_l. Here $\delta_{t\in[0,T_i]}$ is a Kronecker delta function, which is defined such that $\delta_{t\in[0,T_i]} = 1$ if $t \in [0,T_i]$ otherwise $\delta_{t\in[0,T_i]} = 0$. The diffusivity field, $\nu(\boldsymbol{x};\omega)$, and the location of the source, $\boldsymbol{z}$, are assumed to be uncertain. The transient diffusion equation (43) satisfies adiabatic boundary conditions, i.e.

$$\nabla u(\boldsymbol{x},t;\omega)\cdot\hat{n} = 0, \tag{44}$$

$$u(\boldsymbol{x},t;\omega) = 0, \tag{45}$$

at the boundary.

The efficacy of the proposed method is demonstrated using the 'hypothetical test bed' data, which is defined using the numerical solution of (43) for the completely known source locations and the diffusivity field. In the present paper, the proposed method is demonstrated for a single source located at $(0.2,-0.2)$, which is active during

the time interval $[0, 0.01]$. The spatial variation of the mean diffusivity is assumed to take the form

$$\overline{\nu(\boldsymbol{x})} = 0.05(\nu_0 + 10.0 + 0.25x + 0.65y + x^3 + y^3), \quad (46)$$

where ν_0 is a user defined constant. Figure 1(a) shows the spatial variation in the diffusivity. The deterministic numerical solution is obtained using a second-order central difference scheme in the spatial dimension with the uniform grid spacing $h = 0.02$, while the explicit fourth order Runge–Kutta scheme is used for time integration with the time step $\Delta t = 0.0001$. Figure 1(b) shows the numerical solution at $t = 0.05$ s, while the solution at $t = 0.1$ s is shown in Fig. 1(c). Note that the source has peak strength at $\boldsymbol{z}_l$ and reduces exponentially with the distance, resulting in the peak value of u at the source location and the subsequent diffusion with time to other locations. Upon removal of the source, diffusion of u is non-uniform owing to the non-linear diffusivity.

4.1. DBFE formulation. For notational convenience, define

$$S(\boldsymbol{x};\omega) = \frac{s}{2\pi\sigma^2}\exp\left(-\frac{(\boldsymbol{z}-\boldsymbol{x})^2}{2\sigma^2}\right), \quad (47)$$

which is uncertain owing to the uncertainty in the source location $\boldsymbol{z}$. The prior uncertainty in $\boldsymbol{z}$ is expanded in a gPC basis, while the Galerkin projection is used to obtain the resultant gPC coefficients, $\hat{S}(\boldsymbol{x})$, of

$$S(\boldsymbol{x};\omega) = \sum_{p=1}^{P}\hat{S}(\boldsymbol{x})\psi_p(\omega). \quad (48)$$

The prior uncertainty in $\nu(\boldsymbol{x};\omega)$ is represented using a Gaussian process, which is spectrally represented using the bi-orthogonal expansion as

$$\nu(\boldsymbol{x};\omega) = \overline{\nu}(\boldsymbol{x}) + \sum_{i=1}^{N}\sum_{p=1}^{P} V_p^i \nu_i(\boldsymbol{x})\psi_p(\omega), \quad (49)$$

where $\overline{\nu}(\boldsymbol{x})$ is the mean, $\nu_i(\boldsymbol{x})$ are the eigenfunctions of the covariance function of $\nu(\boldsymbol{x};\omega)$ and V_p^i are the respective expansion coefficients. Use (48) and (49) in (43) to obtain the differential operator in (SPDE) as

$$\begin{aligned}
&\mathcal{L}\left[u(\boldsymbol{x},t;\omega);\boldsymbol{\theta}(\omega)\right] \\
&\quad = \nabla[\overline{\nu}(\boldsymbol{x})\nabla\overline{u}(\boldsymbol{x},t) \\
&\qquad + \overline{\nu}(\boldsymbol{x})\sum_{i=1}^{N}\sum_{p=1}^{P} Y_p^i(t)\psi_p(\boldsymbol{\xi}(\omega))\nabla u_i(\boldsymbol{x},t) \\
&\qquad + \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{p=1}^{P}\sum_{q=1}^{P} V_p^i Y_q^j(t)\nu_i(\boldsymbol{x})\psi_p(\boldsymbol{\xi}(\omega)) \\
&\qquad \times \psi_q(\boldsymbol{\xi}(\omega))\nabla u_j(\boldsymbol{x},t) \\
&\qquad + \sum_{i=1}^{N}\sum_{p=1}^{P} V_p^i \nu_i(\boldsymbol{x})\psi_p(\boldsymbol{\xi}(\omega))\nabla\overline{u}(\boldsymbol{x},t)] \\
&\qquad + \sum_{p=1}^{P}\hat{S}(\boldsymbol{x})\psi_p(\boldsymbol{\xi}(\omega)).
\end{aligned} \quad (50)$$

Use (50) in (16), (19) and (27) to obtain the DBFE governing equations for the two-dimensional transient-diffusion equation.

4.2. Solution of the forward problem. The prior uncertainty in the source location is specified using independent Gaussian distributions for the x and y co-ordinates with zero mean and a standard deviation of 0.3. The prior for diffusivity $\nu(\boldsymbol{x};\omega)$ is specified using a Gaussian process with the mean

$$\nu(\boldsymbol{x}) = 0.05(\nu_0 + 10.0 + 0.25x + 0.65y) \quad (51)$$

and the squared exponential covariance function

$$\begin{aligned}
&C(\boldsymbol{x}_1,\boldsymbol{x}_2) \\
&\quad = \sigma_\nu^2\exp\left(-\lambda_1^v(x_1-x_2)^2 - \lambda_2^v(y_1-y_2)^2\right),
\end{aligned} \quad (52)$$

where σ_ν^2 is the variance of the Gaussian process and λ_i^v is the correlation length. Note that though these specific priors are chosen in this paper for demonstration, the method is insensitive to the choice of the prior. The proposed method can be implemented for an arbitrary prior using an appropriate orthogonal gPC basis (Xiu and Karniadakis, 2003).

The efficacy and the computational efficiency of the proposed Bayesian inference depend on the ability of the DBFE method to accurately solve the forward problem. In this section, the accuracy and the computational cost for the numerical implementation of the DBFE are compared against the Monte Carlo and the generalized polynomial chaos method (see the work of Marzouk and Najm (2007) for the gPC formulation of (43)).

Figure 2 shows the accuracy and computational efficiency of the DBFE and the gPC method for different numbers of the eigenfunctions used, N, and the order of the polynomial chaos basis, p. The accuracy is compared using the Monte Carlo method with 10000 samples, which are collected at the computational cost of 6616.17 seconds. The computational cost for solution of the forward problem increases with an increase in N and p for both the DBFE and gPC methods. Note that the stochastic dimension for the present problem is $N+2$ (N dimensions representing the truncated KL expansion, with

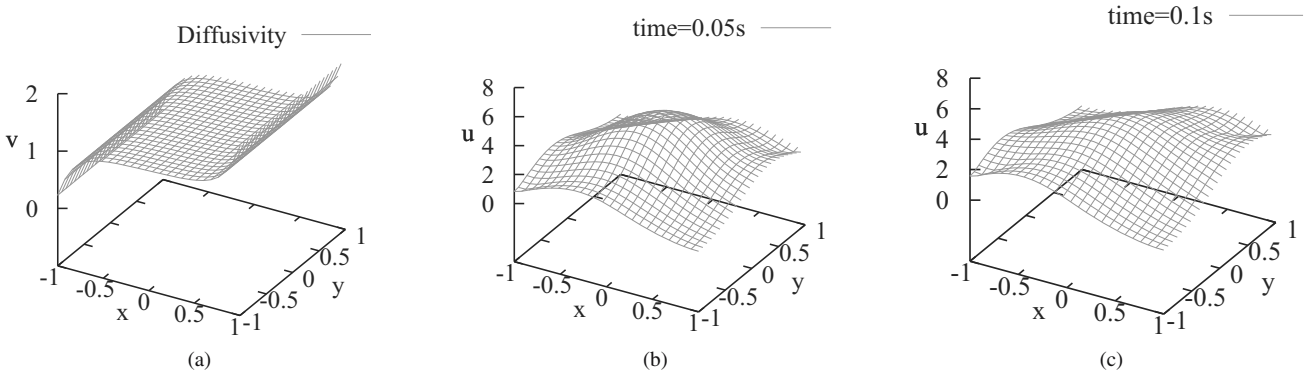


Fig. 1. Solution of a two-dimensional transient diffusion equation: spatial variation in diffusivity (a), u-field at time $t = 0.05$ s (b), u-field at $t = 0.1$ s (c).

two dimensions representing uncertainty in the source location), for which the number of polynomial chaos terms is given by

$$P = \frac{(N+2+p)!}{(N+2)!p!} + 1. \tag{53}$$

Since implementation of the gPC method requires numerical solution of P PDEs (see (Marzouk and Najm, 2009) for details), whereas the DBFE method involves numerical solution of $(N+1)$ PDEs and $N \times P$ ODEs, the increase in computational cost with p is significantly higher for the gPC method compared with the DBFE method. The computational cost of the gPC method is comparable to the Monte Carlo method for $N = 6$ and the second-order polynomial chaos basis, while the computational cost is higher than for the Monte Carlo method for the third-order polynomial chaos basis with $N \geq 3$, rendering the gPC method computationally intractable.

The DBFE method is numerically implemented at a computational cost comparable to the gPC method for $p = 1$, while the computational cost for the DBFE method for $p \geq 2$ is considerably lower than for the gPC method. Figure 2(b) shows the L_1-error in the variance, which is defined as

$$e = \frac{\int_{\mathcal{X}} \mid \hat{v}(\boldsymbol{x}) - v(\boldsymbol{x}) \mid \mathrm{d}\boldsymbol{x}}{\int_{\mathcal{X}} \mid \hat{v}(\boldsymbol{x}) \mid \mathrm{d}\boldsymbol{x}}, \tag{54}$$

where $\hat{v}(\boldsymbol{x})$ is the variance obtained using the Monte Carlo method, while $v(\boldsymbol{x})$ is the variance obtained using the gPC or the DBFE method. The error is comparable for both the DBFE and the gPC methods, which decreases with N and reaches a limiting value for $N \geq 4$, though the limiting value is higher for $p = 1$. Note that the transient diffusion equation involves multiplication of the diffusivity ν with ∇u. Thus, the appropriate spectral representation requires the use of the second-order polynomial chaos basis. From the results, it may be concluded that (43) can be numerically solved using the DBFE method at significantly lower computational cost than the gPC method with a comparable accuracy.

4.3. Solution of the source inversion problem. The proposed method is used for inference of the source location and the diffusivity. The prior uncertainty in the source location is given by independent Gaussian processes in the x and y directions, with $\mathcal{N}(0.0, 0.3)$ prior. The prior uncertainty in the diffusivity is specified using the Gaussian process with the mean (51) and the covariance function (52) with $\sigma^2 = 0.3$ and $\lambda = 1.5$. The prior uncertainty is propagated to the system response using the DBFE method with $N = 5$ and $p = 2$. Note that the accuracy and computational cost of the DBFE method depend on the choices of N and p, necessitating the appropriate compromise. The choice of p depends on the non-linearity associated with the governing equations. As can be seen from Fig. 2(b), $p = 2$ suffices for the present test case. The accuracy of the truncated Karhunen–Loève expansion depends on the number of the eigenfunctions used, N, through the variance of the associated random expansion coefficients.

Figure 3(a) shows the eigenvalues of the covariance function of the diffusivity. From the figure, it may be seen that the prior uncertainty in the diffusivity can be satisfactorily approximated with $N = 3$. However, to allow the possibility of higher order modes becoming active, $N = 5$ is used in the present paper. Figure 3(b) shows the variance of the random coefficient of the respective eigenfunctions. From the figure, it can be seen that variance for $N > 3$ is significantly lower than the variance of the first three eigenmodes, resulting in an acceptable approximation of the solution field over time evolution. Note that N also depends on the uncertainty in the source, with a higher uncertainty necessitating the

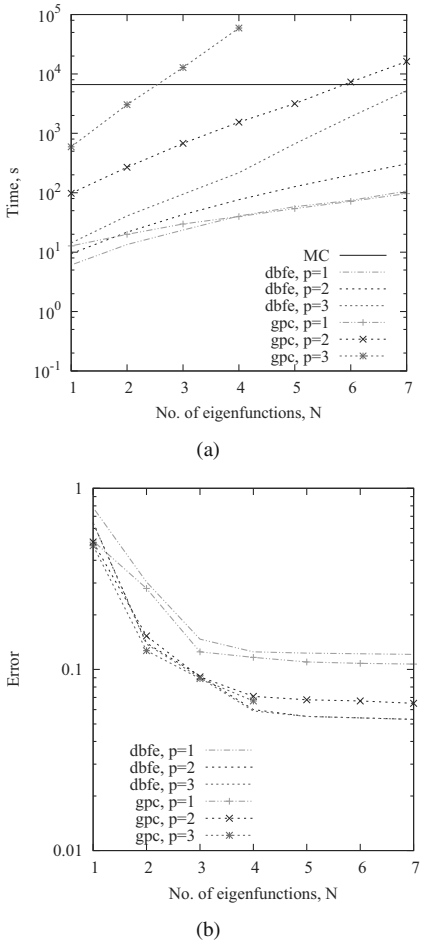


Fig. 2. Comparison of the accuracy and computational efficiency of the DBFE with gPC and Monte Carlo methods. Panel (a) shows a comparison of CPU time and (b) shows the L_1-error.

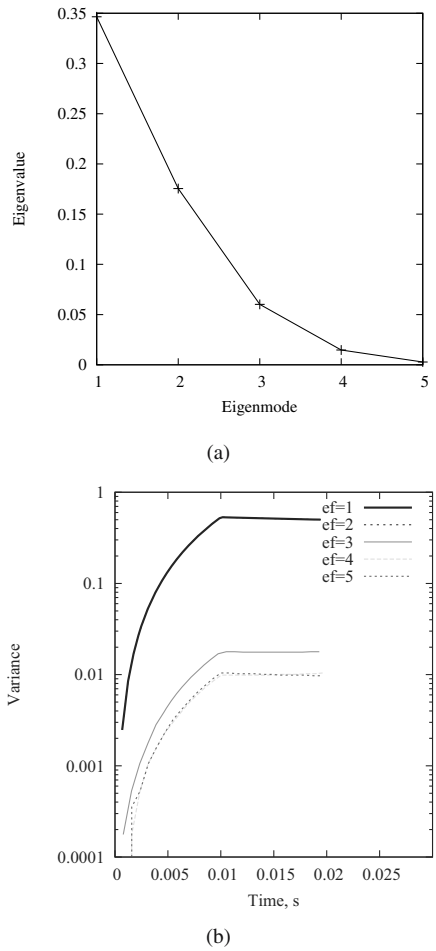


Fig. 3. Panel (a) shows the eigenvalues of the covariance function of diffusivity, and panel (b) displays time evolution of the variance of random coefficients.

use of a higher value of N. Further note that in the present implementation, eigenmodes are adaptively activated only after the variance crosses a given minimum value (1^{-5} in the present case).

The deterministic solution of the 2D transient diffusion equation at time $t = 0.02$ seconds with the source removed at $t = 0.01$ seconds is used as experimental observations. The source is assumed to be located in $[0.2, -0.2]$. A total of 25 uniformly spaced data points are used for the Bayesian inference. A 1% experimental uncertainty is assumed for each data point. The model structure uncertainty is defined by specifying the prior probability distribution for σ_δ^2 and λ_δ. The inverse Gamma distribution $IG(6.0, 2.0)$ is used for σ_δ^2, while the prior for λ_δ is given by the Gamma distribution $G(6.0, 2.0)$. The authors compared different numbers of MCMC samples to explore the posterior distribution. A total of 10,000 MCMC samples, after the burnout period of 1000 samples, provided an acceptable accuracy.

Figure 4 shows the prior and posterior probability density contours for the source location. The location of the experimental observations is also shown in Fig. 4(a). The contour is shown for the posterior density obtained using the proposed method (dashed line) and direct MCMC sampling from the posterior distribution (solid lines). The source location is predicted accurately, while the posterior density obtained using the proposed method agrees closely with direct MCMC sampling, demonstrating the efficacy of the DBFE based Bayesian inference. Figure 5(a) shows prior and posterior probability distributions of the gPC basis $\boldsymbol{\xi}$. The posterior distribution of ξ_1–ξ_5 represents an update in the uncertainty of the diffusivity, while the posterior distribution of ξ_6 and ξ_7 represents the inference of x and y co-ordinates of the source location, respectively. The probability distribution of ξ_1 is updated with variance

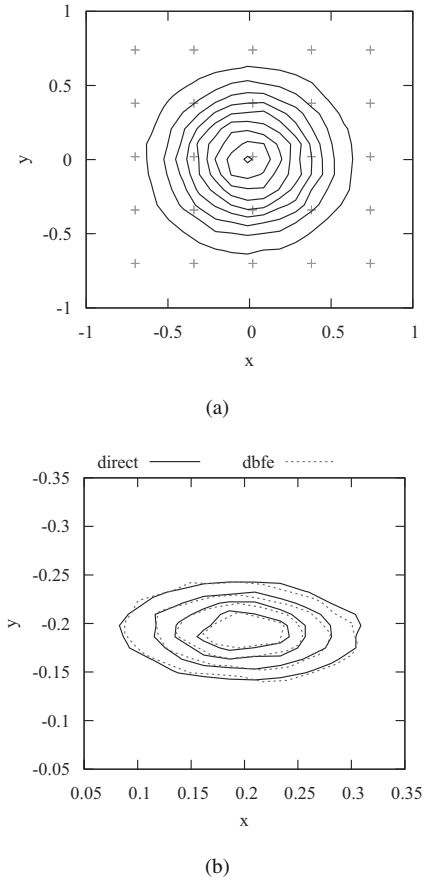


Fig. 4. Probability contours for the source location. The location of experimental observations is also shown in (a).

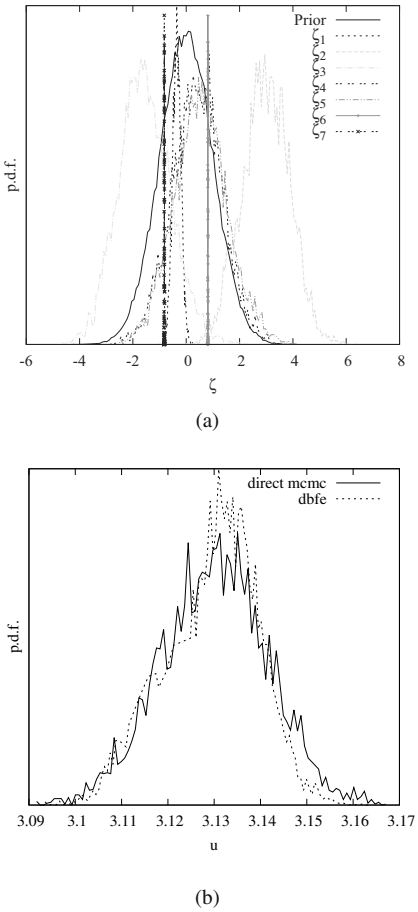
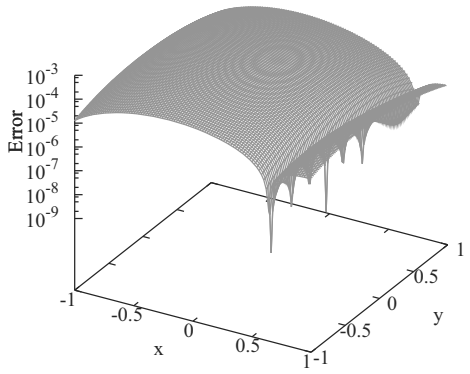


Fig. 5. Panel (a) shows the posterior probability distribution of $\boldsymbol{\xi}$, and panel (b) displays a comparison of the posterior probability distribution of u at $x = 0.2$ and $y = -0.2$ obtained using direct MCMC and the proposed method.

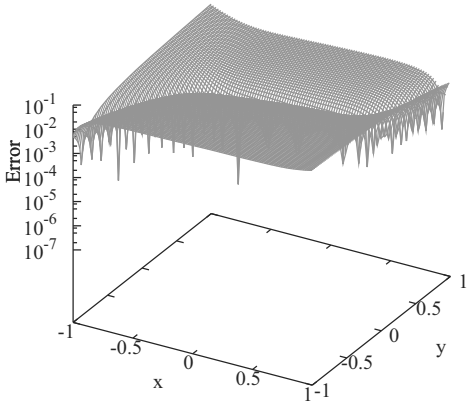
significantly lower than the prior, whereas the update of the probability distribution for ξ_2–ξ_5 is comparatively low. Note that the variance of the random coefficients associated with the eigenmodes $N > 3$ is negligible, resulting in the negligible update of the probability distribution of ξ_4 and ξ_5. The probability distribution of ξ_6 and ξ_7 is updated significantly, indicating high information contained in the experimental observations about the source location. The low variances of ξ_6 and ξ_7 indicate very high confidence on the inferred source location. Figure 5(b) shows a comparison of the probability distribution of $u(\boldsymbol{x}, t; \omega)$ at the source location ($x = 0.2$ and $y = -0.2$) obtained by running the forward model at posterior samples of direct MCMC and the proposed method. The probability distribution for the proposed method matches closely with the direct MCMC method. Similar results are obtained at other locations. However, they are not shown in the paper for brevity.

Figure 6(a) shows the L_1-error in the posterior variance of the diffusivity between the DBFE direct MCMC sampling. The maximum error is of the order of 10^{-3}, indicating the close agreement in variance for the posterior probability of the diffusivity obtained using the DBFE and direct MCMC sampling methods. Figure 6(b) shows the L_1-error for posterior mean of the diffusivity for the DBFE method obtained against the 'true' diffusivity. Note that the error reduces non-uniformly, indicating the effect of the location of the experimental observations on the proposed Bayesian inference.

Figure 7 draws a comparison of the posterior probability distributions of σ_δ^2 and λ_δ obtained using direct MCMC sampling and the proposed method. The posterior distribution for λ_δ obtained using the proposed method matches closely direct sampling. However, the match is comparatively poor for the posterior distribution of σ_δ^2. Note that the bi-orthogonal expansion obtained using the DBFE method acts as an emulator of the 2D transient diffusion equation, which is used in the Bayesian inference against the simulator in direct MCMC sampling. Thus, any remnant error in the bi-orthogonal expansion is regarded as the uncertainty in the model structure, resulting in the difference in the posterior probability distribution for σ_δ^2. The posterior probability

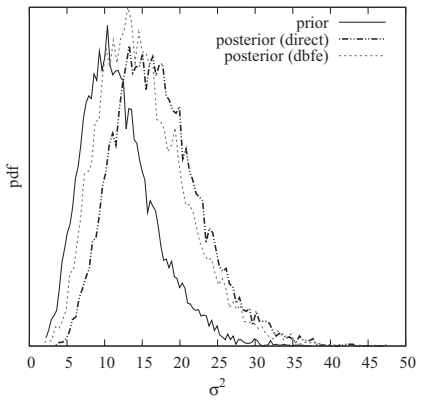


(a)

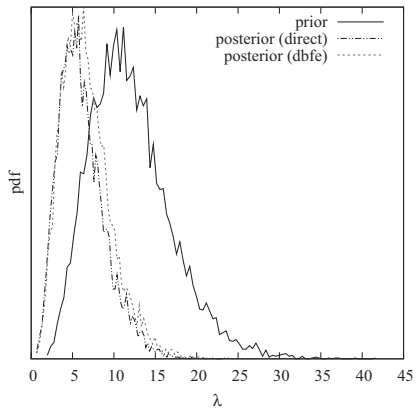


(b)

Fig. 6. L_1-errors in the posterior variance for the DBFE and direct MCMC sampling (a) and the posterior mean for diffusivity.



(a)



(b)

Fig. 7. Comparison of probability distributions of σ_δ^2 (a) and λ_δ (b).

for λ_δ has moved towards the left for both the cases, indicating an increased correlation for the model structure uncertainty, while the posterior distribution for σ_δ^2 moves towards the right, indicating higher posterior confidence on the simulator. Detailed guidelines for inference on the veracity and validity of the model based on the posterior distribution of hyper-parameters of the discrepancy function are provided by Tagade and Choi (2013).

Figure 8 shows the L_1-error in the posterior mean for the system response u, which is defined against the true spatial distribution of u. The posterior mean of the system response is calculated by substituting the mean of $\boldsymbol{\xi}$ in the bi-orthogonal expansion. The maximum L_1-error is of the order of 10^{-1}, which is located in the boundary region where experimental data are not provided for Bayesian inference. In the region where experimental observations are available, the error is significantly low with the minimum value of the order of 10^{-7}.

5. Concluding remarks

The paper has presented a dynamic bi-orthogonality based approach to computationally efficient implementation of Bayesian inference. The proposed method can be applied for calibration of a simulator represented using a partial differential equation with high dimensional uncertainty. Though the method requires reformulation of the governing equations, the existing schemes can be extended in a straightforward manner for the numerical solution of the DBFE.

A key innovation of the proposed approach is a gPC expansion of the stochastic coefficients of the DO expansion. The resultant DBFEs are solved using numerical integration techniques like the central difference method and the fourth-order order Runge–Kutta method. As the proposed approach does not require Monte Carlo sampling for a solution, numerical implementation of the DBFE method is computationally more efficient compared with the state-of-the-art DOFE methods. Numerical examples presented in this paper have demonstrated the computational efficiency of the proposed approach.

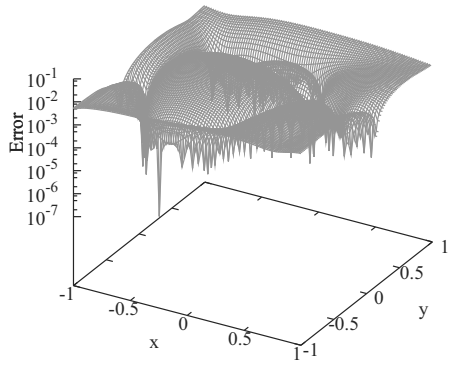


Fig. 8. L_1-error in the posterior mean for u.

Efficacy of the proposed method is demonstrated for calibration of a two-dimensional transient diffusion equation with uncertain source location and diffusivity. The computational cost of the proposed method for uncertainty propagation is compared with the gPC and Monte Carlo methods. It has been observed that as the dimensionality of the uncertainty increases, the DBFE method provides the solution of the SPDE at a significantly less computational cost than the gPC method with a comparable accuracy. It has also been demonstrated that the method provides accurate inference of the source location with the marginal posterior distribution matching closely MCMC sampling. The method is found to accurately infer the posterior distribution of the spatially/temporally varying parameters.

Acknowledgment

This work was supported in part by the AFOSR grant FA9550-12-1-0313 and in part by the Climate Change Research Hub Project of the KAIST EEWS Research Center (grant no. #EEWS-2016-N11160018).

References

Bayarri, M., Berger, J., Paulo, R., Sacks, J., Cafeo, J., Cavendish, J., Lin, C. and Tu, J. (2007). A framework for validation of computer models, *Technometrics* **49**(2): 138–153.

Besag, J., Green, P., Higdon, D. and Mengersen, K. (1995). Bayesian computation and stochastic systems, *Statistical Science* **10**(1): 3–41.

Bieri, M. and Schwab, C. (2009). Sparse high order FEM for elliptic sPDEs, *Computer Methods in Applied Mechanics and Engineering* **198**(13–14): 1149–1170.

Bortz, A., Kalos, M. and Lebowitz, J. (1975). A new algorithm for Monte Carlo simulation in Ising spin systems, *Journal of Computational Physics* **17**(1): 10–18.

Cai, B., Meyer, R. and Perron, F. (2008). Metropolis–Hastings algorithms with adaptive proposals, *Statistics and Computing* **18**(4): 421–433.

Cameron, R. and Martin, W. (1947). The orthogonal development of non-linear functionals in series of Fourier–Hermite functionals, *The Annals of Mathematics* **48**(2): 385–392.

Cheng, M., Hou, T. and Zhang, Z. (2013a). A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations. I: Derivation and algorithms, *Journal of Computational Physics* **242**(1): 843–868.

Cheng, M., Hou, T. and Zhang, Z. (2013b). A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations. II: Adaptivity and generalizations, *Journal of Computational Physics* **242**(1): 753–776.

Cotter, S., Dashti, M., Robinson, J. and Stuart, A. (2012). Variational data assimilation using targeted random walks, *International Journal for Numerical Methods in Fluids* **68**(4).

Cotter, S., Roberts, G., Stuart, A. and White, D. (2013). MCMC methods for functions: Modifying old algorithms to make them faster, *Statistical Science* **28**(3): 424–446.

Dumbser, M. and Munz, C.-D. (2007). On source terms and boundary conditions using arbitrary high order discontinuous Galerkin schemes, *International Journal of Applied Mathematics and Computer Science* **17**(3): 297–310, DOI: 10.2478/v10006-007-0024-1.

Gamerman, D. and Lopes, H. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, Chapman and Hall/CRC, Boca Raton, FL.

Ghanem, R. and Red-Horse, J. (1999). Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element approach, *Physica D* **133**(1–4): 137–144.

Ghanem, R. and Spanos, P. (1991). Spectral stochastic finite-element formulation for reliability analysis, *Journal of Engineering Mechanics* **117**(10): 2351–2372.

Ghanem, R. and Spanos, P. (2003). *Stochastic Finite Elements: A Spectral Approach*, Dover Publications, New York, NY.

Gilks, W., Roberts, G. and Sahu, S. (1998). Adaptive Markov chain Monte Carlo through regeneration, *Journal of American Statistical Association* **93**(443): 1045–1054.

Goldstein, M. and Rougier, J. (2005). Probabilistic formulations for transferring inferences from mathematical models to physical systems, *SIAM Journal of Scientific Computing* **26**(2): 467–487.

Hastings, W. (1970). Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* **57**(1): 97–109.

Higdon, D., Kennedy, M., Cavendish, J., Cafeo, J. and Ryne, R. (2005). Combining field data and computer simulations for calibration and prediction, *SIAM Journal of Scientific Computing* **26**(2): 448–446.

Hoang, V., Schwab, C. and Stuart, A. (2013). Complexity analysis of accelerated MCMC methods for Bayesian inversion, *Inverse Problems* **29**(8): 085010

Janiszowski, K. and Wnuk, P. (2016). Identification of parameteric models with *a priori* knowledge of process properties, *International Journal of Applied Mathematics and Computer Science* **26**(4): 767–776, DOI: 10.1515/amcs-2016-0054.

Kamiński, M. (2015). Symbolic computing in probabilistic and stochastic analysis, *International Journal of Applied Mathematics and Computer Science* **25**(4): 961–973, DOI: 10.1515/amcs-2015-0069.

Karczewska, A., Pozmej, P., Szczeciński, M. and Boguniewicz, B. (2016). A finite element method for extended KdV equations, *International Journal of Applied Mathematics and Computer Science* **26**(3): 555–567, DOI: 10.1515/amcs-2016-0039.

Kelly, D. and Smith, C. (2009). Bayesian inference in probabilistic risk assessment—the current state of the art, *Reliability Engineering and System Safety* **94**(2): 628–643.

Kennedy, M. and O'Hagan, A. (2001). Bayesian calibration of computer models, *Journal of the Royal Statistical Society B: Statistical Methodology* **63**(3): 425–464.

Knio, O. and Maitre, O. (2006). Uncertainty propagation in CFD using polynomial chaos decomposition, *Fluid Dynamics Research* **38**(9): 616–640.

Lucor, D., Xiu, D., Su, C. and Karniadakis, G. (2003). Predictability and uncertainty in CFD, *International Journal for Numerical Methods in Fluids* **43**(5): 483–505.

Marzouk, Y. and Najm, H. (2007). Stochastic spectral methods for efficient Bayesian solution of inverse problems, *Journal of Computational Physics* **224**(2): 560–586.

Marzouk, Y. and Najm, H. (2009). Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems, *Journal of Computational Physics* **228**(6): 18621902.

Mathelin, L., Hussaini, M., Zang, T. and Bataille, F. (2004). Uncertainty propagation for a turbulent, compressible nozzle flow using stochastic methods, *AIAA Journal* **42**(8): 1669–1676.

Mehta, U. (1991). Some aspects of uncertainty in computational fluid dynamics results, *Journal of Fluid Engineering* **113**(4): 538–543.

Mehta, U. (1996). Guide to credible computer simulations of fluid flows, *Journal of Propulsion and Power* **12**(5): 940–948.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. and Teller, E. (1953). Equation of state calculations by fast computing machines, *The Journal of Chemical Physics* **21**(6): 1087–1092.

Narayanan, V. and Zabaras, N. (2004). Stochastic inverse heat conduction using spectral approach, *International Journal for Numerical Methods in Engineering* **60**(9): 1569–1593.

Oberkampf, W., DeLand, S., Rutherford, B., Diegert, K. and Alvin, K. (2002). Error and uncertainty in modeling and simulation, *Reliability Engineering and System Safety* **75**(3): 335–357.

O'Hagan, A. (2006). Bayesian analysis of computer code outputs: A tutorial, *Reliability Engineering and System Safety* **91**(10–11): 1290–1300.

Oreskes, N., Shrader-Frechett, K. and Belitz, K. (1994). Verification, validation and confirmation of numerical models in earth sciences, *Science* **263**(5147): 641–647.

Paulo, R. (2005). Default priors for Gaussian processes, *The Annals of Statistics* **33**(2): 556–582.

Poette, G., Despres, B. and Lucor, D. (2009). Uncertainty quantification for systems of conservation laws, *Journal of Computational Physics* **228**(7): 2443–2467.

Sapsis, T. and Lermusiaux, P. (2009). Dynamically orthogonal field equations for continuous stochastic dynamical systems, *Physica D* **238**: 2347–2360.

Sapsis, T. and Lermusiaux, P. (2012). Dynamical criteria for the evolution of the stochastic dimensionality in flows with uncertainty, *Physica D* **241**(1): 60–76.

Schwab, C. and Gittelson, C. (2011). Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs, *Acta Numerica* **20**: 291–467.

Schwab, C. and Stuart, A. (2012). Sparse deterministic approximation of Bayesian inverse problems, *Inverse Problems* **28**(4): 1–32.

Tagade, P. and Choi, H.-L. (2012). An efficient Bayesian calibration approach using dynamically biorthogonal field equations, *ASME International Design Engineering Technical Conference and Computers and Information in Engineering Conference, Chicago, IL, USA*, pp. 873–882.

Tagade, P. and Choi, H.-L. (2013). A generalized polynomial chaos-based method for efficient Bayesian calibration of uncertain computational models, *Inverse Problems in Science and Engineering* **22**(4): 602–624.

Tagade, P. and Sudhakar, K. (2011). Inferencing component maps of gas turbine engine using Bayesian framework, *Journal of Propulsion and Power* **27**(1): 94–104.

Tagade, P., Sudhakar, K. and Sane, S. (2009). Bayesian framework for calibration of gas turbine simulator, *Journal of Propulsion and Power* **25**(4): 987–992.

Trucano, T., Swiler, L., Igusa, T., Oberkampf, W. and M., P. (2006). Calibration, validation, and sensitivity analysis: What's what, *Reliability Engineering and System Safety* **91**(10–11): 1331–1357.

Venturi, D. (2011). A fully symmetric nonlinear biorthogonal decomposition theory for random fields, *Physica D* **240**(4–5): 415–425.

Wiener, N. (1938). The homogeneous chaos, *American Journal of Mathematics* **60**(4): 897–936.

Wiener, N. (1958). *Nonlinear Problems in Random Theory*, John Wiley & Sons, New York, NY.

Xiu, D. and Karniadakis, E. (2002). The Weiner–Askey polynomial chaos for stochastic differential equations, *SIAM Journal of Scientific Computing* **24**(2): 619–644.

Xiu, D. and Karniadakis, G. (2003). Modeling uncertainty in flow simulations via generalized polynomial chaos, *Journal of Computational Physics* **187**(1): 137–167.

Zaidi, A., Ould Bouamama, B. and Tagina, M. (2012). Bayesian reliability models of Weibull systems: State of the art, *International Journal of Applied Mathematics and Computer Science* **22**(3): 585–600, DOI: 10.2478/v10006-012-0045-2.

Piyush M. Tagade is a chief engineer at the Samsung Advanced Institute of Technology, Samsung R&D Institute, Bangalore, India. He holds a PhD degree in aerospace engineering from the Indian Institute of Technology, Bombay, India. Before joining Samsung, he was a postdoctoral research associate at the Massachusetts Institute of Technology, USA, and the Korea Advanced Institute of Science and Technology, Republic of Korea. In his scientific research work he is mostly concerned with developing an efficient Bayesian framework for large scale system simulators. His areas of interest include Bayesian inference, uncertainty propagation, data assimilation, optimization and machine learning.

Han-Lim Choi is an associate professor of aerospace engineering at the KAIST (Korea Advanced Institute of Science and Technology), and a visiting scholar in the School of Aeronautics and Astronautics at Purdue University. He received his BSc and MSc degrees in aerospace engineering from the KAIST, Daejeon, Korea, in 2000 and 2002, respectively, and his PhD degree in aeronautics and astronautics from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2009. He then worked at the MIT as a postdoctoral associate until he joined the KAIST in 2010. His current research interests include planning and control of multi-agent systems, planning and control under uncertainty, resource management in radars, and Bayesian inference for large-scale systems. He (together with Dr. Jonathan P. How) is the recipient of the 2011 Automatica Applications Prize.

Received: 14 September 2016
Revised: 30 January 2017
Re-revised: 5 February 2017
Accepted: 8 March 2017