A Stochastic Collocation Approach to Bayesian Inference in Inverse Problems

Youssef Marzouk[1,*] and Dongbin Xiu[2]

[1] *Department of Aeronautics & Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.*
[2] *Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA.*

Abstract. We present an efficient numerical strategy for the Bayesian solution of inverse problems. Stochastic collocation methods, based on generalized polynomial chaos (gPC), are used to construct a polynomial approximation of the forward solution over the support of the prior distribution. This approximation then defines a surrogate posterior probability density that can be evaluated repeatedly at minimal computational cost. The ability to simulate a large number of samples from the posterior distribution results in very accurate estimates of the inverse solution and its associated uncertainty. Combined with high accuracy of the gPC-based forward solver, the new algorithm can provide great efficiency in practical applications. A rigorous error analysis of the algorithm is conducted, where we establish convergence of the approximate posterior to the true posterior and obtain an estimate of the convergence rate. It is proved that fast (exponential) convergence of the gPC forward solution yields similarly fast (exponential) convergence of the posterior. The numerical strategy and the predicted convergence rates are then demonstrated on nonlinear inverse problems of varying smoothness and dimension.

AMS subject classifications: 41A10, 60H35, 65C30, 65C50

Key words: Inverse problems, Bayesian inference, stochastic collocation, generalized polynomial chaos, uncertainty quantification.

1   Introduction

The indirect estimation of model parameters or inputs from observations constitutes an *inverse problem*. Such problems arise frequently in science and engineering, with applica-

---

*Corresponding author. *Email addresses:* `ymarz@mit.edu` (Y. Marzouk), `dxiu@purdue.edu` (D. Xiu)

tions ranging from subsurface and atmospheric transport to chemical kinetics. In practical settings, observations are inevitably noisy and may be limited in number or resolution. Quantifying the resulting *uncertainty* in inputs or parameters is then essential for predictive modeling and simulation-based decision-making.

The Bayesian approach to inverse problems [6,13,18,22,23] provides a foundation for inference from noisy and incomplete data, a natural mechanism for incorporating physical constraints and heterogeneous sources of information, and a quantitative assessment of uncertainty in the inverse solution. Indeed, the Bayesian setting casts the inverse solution as a *posterior* probability distribution over the model parameters or inputs. Though conceptually straightforward, this setting presents challenges in practice; the posterior probability distribution is typically not of analytical form and, especially in high dimensions, cannot be easily interrogated. Many numerical approaches have been developed in response, mostly seeking to approximate the posterior distribution or posterior expectations via samples [9]. These approaches require repeated solutions of the forward model; when the model is computationally intensive, e.g., specified by partial differential equations (PDEs), the Bayesian approach then becomes prohibitive.

Several efforts at accelerating Bayesian inference in inverse problems have appeared in recent literature; these have relied largely on reductions or surrogates for the forward model [3, 14, 17, 24], or instead have sought more efficient sampling from the posterior [4,5,11]. Recent work [17] used (generalized) polynomial chaos (gPC)-based stochastic Galerkin methods [8, 29] to propagate prior uncertainty through the forward model, thus yielding a polynomial approximation of the forward solution over the support of the prior. This approximation then entered the likelihood function, resulting in a posterior density that was inexpensive to evaluate. This scheme was used to infer parameters appearing nonlinearly in a transient diffusion equation, demonstrating exponential convergence to the true posterior and multiple order-of-magnitude speedup in posterior exploration via Markov chain Monte Carlo (MCMC). The gPC stochastic Galerkin approach has also been extended to Bayesian inference of spatially-distributed quantities, such as inhomogeneous material properties appearing as coefficients in a PDE [16].

An alternative to the stochastic Galerkin approach to uncertainty propagation is stochastic collocation [25,27]. A key advantage of stochastic collocation is that it requires only a finite number of uncoupled deterministic simulations, with no reformulation of the governing equations of the forward model. Also, stochastic collocation can deal with highly nonlinear problems that are challenging, if not impossible, to handle with stochastic Galerkin methods. A spectral representation may also be applied to arbitrary functionals of the forward solution; moreover, many methods exist for addressing high input dimensionality via efficient low-degree integration formulae or sparse grids. For an extensive discussion of gPC-based algorithms, see [26].

This paper extends the work of [17] by using gPC stochastic collocation to construct posterior surrogates for efficient Bayesian inference in inverse problems. We also conduct a rigorous error analysis of the gPC Bayesian inverse scheme. Convergence of the approximate posterior distribution to the true posterior distribution is established and

its asymptotic convergence rate obtained. Numerical examples are provided for a variety of nonlinear inverse problems to verify the theoretical findings and demonstrate the efficiency of the new algorithms.

2 Formulation

Let $D \subset \mathbb{R}^\ell$, $\ell = 1, 2, 3$, be a physical domain with coordinates $x = (x_1, \cdots, x_\ell)$ and let $T > 0$ be a real number. We consider the following general stochastic partial differential equation

$$\begin{cases} u_t(x,t,Z) = \mathcal{L}(u), & D \times (0,T] \times \mathbb{R}^{n_z}, \\ \mathcal{B}(u) = 0, & \partial D \times [0,T] \times \mathbb{R}^{n_z}, \\ u = u_0, & D \times \{t = 0\} \times \mathbb{R}^{n_z}, \end{cases} \tag{2.1}$$

where $\mathcal{L}$ is a (nonlinear) differential operator, $\mathcal{B}$ is the boundary condition operator, u_0 is the initial condition, and $Z = (Z_1, \cdots, Z_{n_z}) \in \mathbb{R}^{n_z}, n_z \geq 1$, are a set of independent random variables characterizing the random inputs to the governing equation. The solution is therefore a stochastic quantity,

$$u(x,t,Z) : \bar{D} \times [0,T] \times \mathbb{R}^{n_z} \to \mathbb{R}^{n_u}. \tag{2.2}$$

We assume that each random variable Z_i has a *prior distribution*

$$F_i(z_i) = P(Z_i \leq z_i) \in [0,1],$$

where P denotes probability. In this paper we will focus on continuous random variables. Subsequently each Z_i has a probability density function $\pi_i(z_i) = dF_i(z_i)/dz_i$. The joint prior density function for Z is

$$\pi_Z(z) = \prod_{i=1}^{n_z} \pi_i(z_i). \tag{2.3}$$

Throughout this paper, we will neglect the subscript of each probability density and use $\pi(z)$ to denote the probability density function of the random variable Z, $\pi_Z(z)$, unless confusion arises otherwise. Note that it is possible to loosen the independence assumption on the input random variables Z by assuming some dependence structure; see, for example, discussions in [1, 21]. As the focus of this paper is not on methods for the stochastic problem (2.1), we follow the usual approach by assuming prior independence on Z.

Let

$$d^t = g(u) \in \mathbb{R}^{n_d} \tag{2.4}$$

be a set of variables that one observes, where $g : \mathbb{R}^{n_u} \to \mathbb{R}^{n_d}$ is a function relating the solution u to the true observable d^t. We then define a "forward model" $G : \mathbb{R}^{n_z} \to \mathbb{R}^{n_d}$ to describe the relation between the random parameters Z and observable d^t:

$$d^t = G(Z) \triangleq g \circ u(Z). \tag{2.5}$$

In practice, measurement error is inevitable and the observed data d may not match the true value of d^t. Assuming additive observational errors, we have

$$d = d^t + e = G(Z) + e, \tag{2.6}$$

where $e \in \mathbb{R}^{n_d}$ are mutually independent random variables with probability density functions $\pi(e) = \prod_{i=1}^{n_d} \pi(e_i)$. We make the usual assumption that e are also independent of Z.

The present Bayesian inference problem is concerned with estimating the parameters Z given a set of observations d. To this end, Bayes' rule takes the form

$$\pi(z|d) = \frac{\pi(d|z)\pi(z)}{\int \pi(d|z)\pi(z)dz} , \tag{2.7}$$

where $\pi(z)$ is the prior probability density of Z; $\pi(d|z)$ is the likelihood function; and $\pi(z|d)$, the density of Z conditioned on the data d, is the *posterior probability density* of Z. For notational convenience, we will use $\pi^d(z)$ to denote the posterior density $\pi(z|d)$ and $L(z)$ to denote the likelihood function $\pi(d|z)$. That is, (2.7) can be written as

$$\pi^d(z) = \frac{L(z)\pi(z)}{\int L(z)\pi(z)dz} . \tag{2.8}$$

Following the independence assumption on the measurement noise e, the likelihood function is

$$L(z) \triangleq \pi(d|z) = \prod_{i=1}^{n_d} \pi_{e_i}(d_i - G_i(z)). \tag{2.9}$$

3 Algorithm

In this section we describe a stochastic collocation scheme, based on generalized polynomial chaos (gPC) expansions, for the Bayesian solution of the inverse problem (2.7).

3.1 Generalized polynomial chaos

The generalized polynomial chaos (gPC) is an orthogonal polynomial approximation to random functions. Without loss of generality, in this subsection we describe the gPC approximation to the forward problem (2.5) for $n_d = 1$. When $n_d > 1$ the procedure will be applied to each component of G and is straightforward.

Let $\mathbf{i} = (i_1, \cdots, i_{n_z}) \in \mathbb{N}_0^{n_z}$ be a multi-index with $|\mathbf{i}| = i_1 + \cdots + i_{n_z}$, and let $N \geq 0$ be an integer. The Nth-degree gPC expansion of $G(Z)$ is defined as

$$G_N(Z) = \sum_{|\mathbf{i}|=0}^{N} a_{\mathbf{i}} \Phi_{\mathbf{i}}(Z), \tag{3.1}$$

where

$$a_{\mathbf{i}} = \mathbb{E}[G(Z)\Phi_{\mathbf{i}}(Z)] = \int G(z)\Phi_{\mathbf{i}}(z)\pi(z)dz, \tag{3.2}$$

are the expansion coefficients, $\mathbb{E}$ is the expectation operator, and $\Phi_{\mathbf{i}}(Z)$ are the basis functions defined as

$$\Phi_{\mathbf{i}}(Z) = \phi_{i_1}(Z_1)\cdots\phi_{i_{n_z}}(Z_{n_z}), \quad 0 \le |\mathbf{i}| \le N. \tag{3.3}$$

Here $\phi_m(Z_k)$ is the mth-degree one-dimensional orthogonal polynomial in the Z_k direction satisfying, for all $k = 1, \cdots, n_z$,

$$\mathbb{E}_k[\phi_m(Z_k)\phi_n(Z_k)] = \int \phi_m(z_k)\phi_n(z_k)\pi(z_k)dz_k = \delta_{m,n}, \qquad 0 \le m, n \le N, \tag{3.4}$$

where the expectation $\mathbb{E}_k$ is taken in terms of Z_k only and the basis polynomials have been normalized. Consequently $\{\Phi_{\mathbf{i}}(Z)\}$ are n_z-variate orthonormal polynomials of degree up to N satisfying

$$\mathbb{E}\left[\Phi_{\mathbf{i}}(Z)\Phi_{\mathbf{j}}(Z)\right] = \int \Phi_{\mathbf{i}}(z)\Phi_{\mathbf{j}}(z)\pi(z)dz = \delta_{\mathbf{i},\mathbf{j}}, \quad 0 \le |\mathbf{i}|, |\mathbf{j}| \le N, \tag{3.5}$$

where $\delta_{\mathbf{i},\mathbf{j}} = \prod_{k=1}^{n_z} \delta_{i_k j_k}$. From (3.4), the distribution of Z_k will determine the polynomial type. For example, Hermite polynomials are associated with the Gaussian distribution, Jacobi polynomials with the beta distribution, Laguerre polynomials with the gamma distribution, etc. For a detailed discussion of these correspondences and their resulting computational efficiency, see [28].

Following classical approximation theory, the gPC expansion (3.1) converges when $G(Z)$ is square integrable with respect to $\pi(z)$, that is,

$$\|G(Z) - G_N(Z)\|_{L^2_{\pi_Z}}^2 \triangleq \int (G(z) - G_N(z))^2 \pi(z)dz \to 0, \qquad N \to \infty. \tag{3.6}$$

Furthermore, the rate of convergence depends on the regularity of G such that

$$\|G(Z) - G_N(Z)\|_{L^2_{\pi_Z}} \le CN^{-\alpha}, \tag{3.7}$$

where C is a constant independent of N, and $\alpha > 0$ depends on the smoothness of G. When G is relatively smooth, the convergence rate can be large. This implies that a relatively low-degree expansion can achieve high accuracy and is advantageous in practical stochastic simulations. Many studies have been devoted to the convergence properties of gPC, numerically or analytically, and the computational efficiency of gPC methods. See, for example, [2, 8, 15, 28, 29].

3.2 Stochastic collocation

In the pseudo-spectral stochastic collocation method [25], an approximate gPC expansion is sought, similar to (3.1), in the following form,

$$\widetilde{G}_N(Z) = \sum_{|\mathbf{i}|=0}^{N} \widetilde{a}_{\mathbf{i}}\Phi_{\mathbf{i}}(Z), \tag{3.8}$$

where the expansion coefficients are obtained by

$$\widetilde{a}_{\mathbf{i}} = \sum_{m=1}^{Q} G(Z^{(m)})\Phi_{\mathbf{i}}(Z^{(m)})w^{(m)}, \tag{3.9}$$

where $Z^{(m)} = (Z_1^{(m)}, \cdots, Z_{n_z}^{(m)})$ are a set of nodes and $w^{(m)}$ are the corresponding weights for $m = 1, \cdots, Q$, of an integration rule (cubature) on $\mathbb{R}^{n_z}$ such that

$$\widetilde{a}_{\mathbf{i}} \approx \int G(z)\Phi_{\mathbf{i}}(z)\pi(z)dz = a_{\mathbf{i}}. \tag{3.10}$$

The expansion of (3.8) thus becomes an approximation to the exact expansion (3.1); that is,

$$\widetilde{G}_N(Z) \approx G_N(Z).$$

The difference between the two expansions is the so-called "aliasing error" [25] and is induced by the error of using the integration rule in (3.9). If a convergent integration rule is employed such that

$$\lim_{Q\to\infty} \widetilde{a}_{\mathbf{i}} = a_{\mathbf{i}}, \qquad \forall \mathbf{i},$$

then

$$\lim_{Q\to\infty} \widetilde{G}_N(Z) = G_N(Z), \qquad \forall Z, \tag{3.11}$$

and convergence of $\widetilde{G}_N$ to the exact forward model G follows naturally,

$$\left\| G(Z) - \widetilde{G}_N(Z) \right\|_{L^2_{\pi_Z}}^2 \to 0, \qquad N\to\infty, \ Q\to\infty. \tag{3.12}$$

A prominent feature of the pseudo-spectral collocation method is that it only requires simulations of the forward model $G(Z)$ at fixed nodes $Z^{(m)}, m = 1, \cdots, Q$, which are uncoupled deterministic problems with different parameter settings. This significantly facilitates its application in practical simulations, as long as the aliasing error is under control. For detailed presentation and analysis, see [25].

3.3   gPC-based Bayesian algorithm

In the gPC-based Bayesian method, we use the approximate gPC solution (3.8) to replace the exact (but unknown) forward problem solution (3.1) in Bayes' rule (2.8) and define the following approximate posterior probability,

$$\widetilde{\pi}_N^d(z) = \frac{\widetilde{L}_N(z)\pi(z)}{\int \widetilde{L}_N(z)\pi(z)dz}, \tag{3.13}$$

where $\pi(z)$ is again the prior density of Z and $\widetilde{L}_N$ is the approximate likelihood function defined as

$$\widetilde{L}_N(z) \triangleq \widetilde{\pi}_N(d|z) = \prod_{i=1}^{n_d} \pi_{e_i}(d_i - \widetilde{G}_{N,i}(z)), \tag{3.14}$$

where $\widetilde{G}_{N,i}$ is the i-th component of $\widetilde{G}_N$.

The advantage of this algorithm is that, upon obtaining an accurate gPC solution $\widetilde{G}_N$, dependence on the random parameters Z is known analytically (in polynomial form). Subsequently, the approximate posterior density $\widetilde{\pi}_N^d$ of (3.13) can be evaluated at arbitrary values of z and for an arbitrarily large number of samples, without resorting to additional simulations of the forward problem. Very high accuracy in sampling the posterior distribution can thus be achieved at negligible computational cost. Combined with an efficient forward problem solver employing gPC collocation, this scheme provides a fast and accurate method for Bayesian inference.

4 Convergence study

To establish convergence of the gPC-based Bayesian algorithm, we quantify the difference between the approximate posterior $\widetilde{\pi}_N^d$ and the exact posterior π^d via Kullback-Leibler divergence. The Kullback-Leibler divergence (KLD) measures the difference between probability distributions and is defined, for probability density functions $\pi_1(z)$ and $\pi_2(z)$, as

$$D(\pi_1 \| \pi_2) \triangleq \int \pi_1(z) \log \frac{\pi_1(z)}{\pi_2(z)} dz. \tag{4.1}$$

It is always non-negative, and $D(\pi_1 \| \pi_2) = 0$ when $\pi_1 = \pi_2$.

Similar to (3.13), we define π_N^d as a posterior density obtained in terms of the exact Nth-degree gPC expansion (3.1). That is,

$$\pi_N^d(z) = \frac{L_N(z)\pi(z)}{\int L_N(z)\pi(z)dz}, \tag{4.2}$$

where L_N is the likelihood function obtained by using the exact Nth-degree gPC expansion (3.1),

$$L_N(z) \triangleq \pi_N(d|Z) = \prod_{i=1}^{n_d} \pi_{e_i}(d_i - G_{N,i}(z)), \tag{4.3}$$

where $G_{N,i}$ is the i-th component of G_N. By the definitions of $\widetilde{\pi}_N^d$ and π_N^d, we immediately have the following lemma by following the pointwise convergence of $\widetilde{G}_N$ to G_N in (3.11).

Lemma 4.1. *If $\widetilde{G}_N$ converges to G_N in the form of (3.11), i.e.,*

$$\lim_{Q \to \infty} \widetilde{G}_{N,i}(Z) = G_{N,i}(Z), \qquad 1 \le i \le n_d, \qquad \forall Z, \tag{4.4}$$

then

$$\lim_{Q\to\infty}\widetilde{\pi}_N^d(z)=\pi_N^d(z), \qquad \forall z, \tag{4.5}$$

and

$$\lim_{Q\to\infty}D\left(\widetilde{\pi}_N^d\|\pi_N^d\right)=0. \tag{4.6}$$

Hereafter we employ the common assumption that the observational error in (2.6) is i.i.d. Gaussian, and without loss of generality, assume

$$e\sim N(0,\sigma^2\mathbb{I}), \tag{4.7}$$

where $\sigma>0$ is the standard deviation and $\mathbb{I}$ is the identity matrix of size $n_d\times n_d$.

Lemma 4.2. *Assume that the observational error in (2.6) has an i.i.d. Gaussian distribution (4.7). If the gPC expansion G_N (3.1) of the forward model converges to G in the form of (3.6), i.e.,*

$$\|G_i(Z)-G_{N,i}(Z)\|_{L^2_{\pi_Z}}\to 0, \qquad 1\le i\le n_d, \qquad N\to\infty, \tag{4.8}$$

then the posterior probability π_N^d (4.2) converges to the true posterior probability (2.8) in the sense that the Kullback-Leibler divergence (4.1) converges to zero, i.e,

$$D(\pi_N^d\|\pi^d)\to 0, \qquad N\to\infty. \tag{4.9}$$

Proof. Let

$$\gamma=\int L(z)\pi(z)dz, \qquad \gamma_N=\int L_N(z)\pi(z)dz. \tag{4.10}$$

Obviously, $\gamma>0$ and $\gamma_N>0$. By following the definitions of the likelihood functions $L(z)$ (2.9) and $L_N(z)$ (4.3) and utilizing the fact that function e^{-x} is (uniformly) Lipschitz continuous for $x\ge 0$, i.e., $|e^{-x}-e^{-y}|\le\Lambda|x-y|$ for all $x,y\ge 0$, where Λ is a positive constant, we have

$$\begin{aligned}
|\gamma_N-\gamma|&=\left|\int(L_N(z)-L(z))\pi(z)dz\right|\\
&\le\prod_{i=1}^{n_d}\int\frac{1}{\sqrt{2\pi\sigma^2}}\left|e^{-\frac{(d_i-G_{N,i}(z))^2}{2\sigma^2}}-e^{-\frac{(d_i-G_i(z))^2}{2\sigma^2}}\right|\pi(z)dz\\
&\le\prod_{i=1}^{n_d}\int\frac{\Lambda}{2\sigma^2\sqrt{2\pi\sigma^2}}\left|(d_i-G_{N,i}(z))^2-(d_i-G_i(z))^2\right|\pi(z)dz\\
&\le\prod_{i=1}^{n_d}\frac{\Lambda}{2\sigma^2\sqrt{2\pi\sigma^2}}\|G_{N,i}-G_i\|_{L^2_{\pi_Z}}\|2d_i-G_{N,i}-G_i\|_{L^2_{\pi_Z}}\\
&\le C_1\prod_{i=1}^{n_d}\|G_i-G_{N,i}\|_{L^2_{\pi_Z}},
\end{aligned} \tag{4.11}$$

where Hölder's inequality has been used. Note the positive constant C_1 is independent of N. Therefore, by the $L^2_{\pi_Z}$ convergence of (4.8), we have

$$\gamma_N \to \gamma, \qquad N \to \infty. \tag{4.12}$$

Also,

$$\frac{\pi^d_N}{\pi^d} = \frac{L_N}{L} \frac{\gamma}{\gamma_N} = \frac{\gamma}{\gamma_N} \prod_{i=1}^{n_d} \frac{\pi_{e_i}(d_i - G_{N,i})}{\pi_{e_i}(d_i - G_i)} = \frac{\gamma}{\gamma_N} \prod_{i=1}^{n_d} \exp\left(-\frac{(d_i - G_{N,i})^2 - (d_i - G_i)^2)}{2\sigma^2} \right).$$

Therefore,

$$\log \frac{\pi^d_N}{\pi^d} = -\frac{1}{2\sigma^2} \sum_{i=1}^{n_d} \left[(d_i - G_{N,i})^2 - (d_i - G_i)^2 \right] + \log\left(\frac{\gamma}{\gamma_N} \right),$$

and

$$\begin{aligned}
D(\pi^d_N \| \pi^d) &= \frac{1}{2\sigma^2 \gamma_N} \sum_{i=1}^{n_d} \int L_N(z) \left[(d_i - G_i)^2 - (d_i - G_{N,i})^2 \right] \pi(z) dz \\
&\quad + \frac{1}{\gamma_N} \int L_N(z) \log\left(\frac{\gamma}{\gamma_N} \right) \pi(z) dz \\
&= \frac{1}{2\sigma^2 \gamma_N} \sum_{i=1}^{n_d} \int L_N(z) \left[(d_i - G_i)^2 - (d_i - G_{N,i})^2 \right] \pi(z) dz + \log \frac{\gamma}{\gamma_N}. \tag{4.13}
\end{aligned}$$

Since both $\gamma > 0$ and $\gamma_N > 0$ are constants and $L_N(z)$ is bounded, i.e., $0 < L_N(z) \le C_2$, we obtain immediately

$$\begin{aligned}
D(\pi^d_N \| \pi^d) &\le \frac{C_2}{2\sigma^2 \gamma_N} \sum_{i=1}^{n_d} \int \left| (d_i - G_{N,i})^2 - (d_i - G_i)^2 \right| \pi(z) dz + \left| \log \frac{\gamma}{\gamma_N} \right| \\
&\le \frac{C_3}{2\sigma^2 \gamma_N} \sum_{i=1}^{n_d} \| G_i - G_{N,i} \|_{L^2_{\pi_Z}} + \left| \log \frac{\gamma}{\gamma_N} \right|. \tag{4.14}
\end{aligned}$$

Again, the Hölder inequality has been used. The first term converges by following (4.8). Along with (4.12), the convergence (4.9) is established. □

Lemma 4.3. *Assume the convergence of $G_{N,i}$ takes the form of (3.7), i.e.,*

$$\| G_i(Z) - G_{N,i}(Z) \|_{L^2_{\pi_Z}} \le C N^{-\alpha}, \quad 1 \le i \le n_d, \quad \alpha > 0, \tag{4.15}$$

where the constant C is independent of N, and let γ and γ_N be defined as in (4.10). Then, for sufficiently large N,

$$\left| \frac{\gamma}{\gamma_N} - 1 \right| \le \frac{C_\gamma N^{-\alpha \cdot n_d}}{1 - C_\gamma N^{-\alpha \cdot n_d}} \sim N^{-\alpha \cdot n_d}, \quad \left| \log \frac{\gamma}{\gamma_N} \right| \le \left| \log\left(1 - C_\gamma N^{-\alpha \cdot n_d} \right) \right| \sim N^{-\alpha \cdot n_d}. \tag{4.16}$$

Proof. By using (4.11) and (4.15), we immediately have

$$|\gamma_N - \gamma| \le C_4 N^{-\alpha \cdot n_d}, \tag{4.17}$$

where the constant $C_4 > 0$ is independent of N. To prove the inequality (4.16), we divide the above inequality by γ and require N to be sufficiently large such that

$$0 < 1 - C_\gamma N^{-\alpha \cdot n_d} \le \frac{\gamma_N}{\gamma} \le 1 + C_\gamma N^{-\alpha \cdot n_d},$$

where $C_\gamma = C_4 / \gamma$ is independent of N. The inequality (4.16) is then straightforward. $\qquad\square$

Theorem 4.1. *Assume that the observational error in* (2.6) *has an i.i.d. Gaussian distribution* (4.7) *and the gPC expansion G_N* (3.1) *of the forward model converges to G in the form of* (4.15), *then for sufficiently large N*

$$D(\pi_N^d \| \pi^d) \lesssim N^{-\alpha}. \tag{4.18}$$

Proof. The proof of (4.18) starts from (4.14). By slightly rewriting (4.14), we have

$$\begin{aligned} D(\pi_N^d \| \pi^d) &\le \frac{C_3}{2\sigma^2 \gamma} \frac{\gamma}{\gamma_N} \sum_{i=1}^{n_d} \|G_{N,i} - G_i\|_{L_{\pi_Z}^2} + \left| \log \frac{\gamma}{\gamma_N} \right| \\ &\le \frac{C_3}{2\sigma^2 \gamma} \left(1 + \left| \frac{\gamma}{\gamma_N} - 1 \right| \right) \sum_{i=1}^{n_d} \|G_{N,i} - G_i\|_{L_{\pi_Z}^2} + \left| \log \frac{\gamma}{\gamma_N} \right|. \end{aligned} \tag{4.19}$$

We then establish (4.18) by following (4.15) and (4.16). $\qquad\square$

The result indicates that the asymptotic convergence rate of the posterior distribution π_N^d to π^d, measured by Kullback-Leibler divergence, is at least the same as the $L_{\pi_Z}^2$ convergence rate of the forward model $G_N(Z)$ to $G(Z)$. This result is based on the assumptions of Gaussian measurement noise and mean-square integrability of G and G_N but not continuity. Therefore while the convergence rate of the forward model can be a sharp estimate, the rate (4.18) may not be, and in practice one may see convergence faster than (4.18). Nevertheless, we immediately have

Corollary 4.1. *If $\|G_N - G\|_{L_{\pi_Z}^2}$ converges to zero exponentially fast for sufficiently large* N, *then $D(\pi_N^d \| \pi^d)$ converges to zero exponentially fast for sufficiently large N.*

Finally we have the convergence of $\widetilde{\pi}_N^d$ to π^d.

Theorem 4.2. *If the convergence of G_N is in the form of* (4.8) *and that of $\widetilde{G}_N$ is in the form of* (4.4), *then the posterior density $\widetilde{\pi}_N^d$ converges to the true posterior density π^d in the sense that the KLD converges, i.e.,*

$$D\left(\widetilde{\pi}_N^d \| \pi^d \right) \to 0, \quad N \to \infty, \quad Q \to \infty. \tag{4.20}$$

Proof.

$$
\begin{aligned}
D\left(\widetilde{\pi}_N^d \| \pi^d\right) &= \int \widetilde{\pi}_N^d(z) \log \frac{\widetilde{\pi}_N^d(z)}{\pi^d(z)} dz \\
&= \int \widetilde{\pi}_N^d(z) \log \frac{\widetilde{\pi}_N^d(z)}{\pi_N^d(z)} dz + \int \widetilde{\pi}_N^d(z) \log \frac{\pi_N^d(z)}{\pi^d(z)} dz \\
&= D\left(\widetilde{\pi}_N^d \| \pi_N^d\right) + \int \pi_N^d(z) \log \frac{\pi_N^d(z)}{\pi^d(z)} dz + \int (\widetilde{\pi}_N^d(z) - \pi_N^d(z)) \log \frac{\pi_N^d(z)}{\pi^d(z)} dz \\
&= D\left(\widetilde{\pi}_N^d \| \pi_N^d\right) + D\left(\pi_N^d \| \pi^d\right) + \int (\widetilde{\pi}_N^d(z) - \pi_N^d(z)) \log \frac{\pi_N^d(z)}{\pi^d(z)} dz. \quad (4.21)
\end{aligned}
$$

All three terms converge to zero when $N \to \infty$ and $Q \to \infty$, following Lemma 4.1 and Lemma 4.2. $\square$

5 Numerical examples

In this section we provide numerical examples to verify our theoretical findings and demonstrate the efficacy of the stochastic collocation approach to Bayesian inference.

5.1 Burgers' equation

We consider the viscous Burgers' equation which, under proper conditions, exhibits supersensitivity to a random boundary condition [30]:

$$
\begin{aligned}
u_t + u u_x &= \nu u_{xx}, \qquad x \in [-1,1] \\
u(-1) &= 1 + \delta(Z), \qquad u(1) = -1.
\end{aligned}
$$

Here $\delta(Z) > 0$ is a small perturbation to the left boundary condition and $\nu > 0$ is the viscosity. At steady state, this system has an exact solution,

$$
u(x,Z) = \tanh\left[\frac{A}{2\nu}(x - z_{ex})\right], \tag{5.1}
$$

where z_{ex} is the location of the "transition layer," defined as the zero of the solution profile $u(x = z_{ex}) = 0$, and A is given by the slope at z_{ex}:

$$
-A = \left.\frac{\partial u}{\partial x}\right|_{x=z_{ex}}.
$$

With boundary conditions specified in (5.1), A and z_{ex} may be obtained by solving a nonlinear system of equations. Details are provided in [30].

We now pose a simple one-parameter inverse problem: given noisy observation(s) of the steady-state transition layer location z_{ex}, what is the initial perturbation δ? In the

Bayesian setting, we seek the posterior density of δ conditioned on observations $d_i = z_{ex} + e_i, i = 1 \cdots n_d$. The measurement noise is assumed to be Gaussian, $e_i \sim N(0, \sigma^2)$ with $\sigma = 0.05$. The prior distribution on δ is chosen to be uniform between 0 and an upper bound $\delta_{max} = 0.1$. For convenience, we transform the problem to

$$Z = 2\delta / \delta_{max} - 1,$$

such that the prior distribution on Z is $U(-1, 1)$. The forward model G then maps Z to z_{ex}. (Note that because the definition of z_{ex} does not have an explicit closed formula, a gPC Galerkin approximation is impossible to obtain.)

Using a gPC expansion consisting of Legendre polynomials in Z, the pseudo-spectral stochastic collocation procedure yields an approximation $\widetilde{G}_N(Z)$ to the forward model. This approximation then defines a posterior probability density $\widetilde{\pi}_N^d(z)$, as described in Section 3.3. For comparison, we also compute the exact posterior density $\pi^d(z)$ using the exact forward model G. Fig. 1 shows the resulting densities at two values of the viscosity ν. In both cases, $n_d = 5$ observations were used to estimate δ; these observations are independent random perturbations to the z_{ex} resulting from a "true" value of $\delta = 0.5$.

Posterior probability densities in Fig. 1(a)-(b) are non-Gaussian, reflecting the non-linearity of the forward model. A lower value of the viscosity, $\nu = 0.05$, results in a broader posterior density than the larger value $\nu = 0.10$. This phenomenon is a result of the steady-state solution profile steepening as viscosity decreases. Given a fixed range on δ, the resulting distribution of transition layer locations tightens with smaller ν [30]; conversely, given a fixed observational error in z_{ex}, a wider range of δ values correspond to transitions that fall near the center of the observational distribution—thus spreading the posterior probability density over a wider range of δ. In both cases, however, the approximate posterior densities $\widetilde{\pi}_N^d(z)$ approach the exact density with increasing gPC order N.

Convergence of the posterior with respect to polynomial order is analyzed more quantitatively in Fig. 2. Again, results are shown for $\nu = 0.05$ and $\nu = 0.10$. We plot the L^2 error in the forward model,

$$\|G - \widetilde{G}_N\|_{L_{\pi_Z}^2}^2,$$

and the Kullback-Leibler divergence of the exact posterior from the approximate posterior, $D(\widetilde{\pi}_N^d \| \pi^d)$. A large number of collocation points ($Q = 800$) are employed so that aliasing errors are well-controlled, particularly since results are computed at high order. Since the forward model is smooth, we find the expected exponential convergence of $\widetilde{G}_N$ to G at large N. We also observe exponential convergence of the Kullback-Leibler divergence at large N. (Dashed lines show log-linear fits at large N.) Moreover, we find that the posterior Kullback-Leibler divergence converges somewhat faster than the L^2 error in the forward model, thus exceeding the (minimal) convergence rate guaranteed by Theorem 4.1.
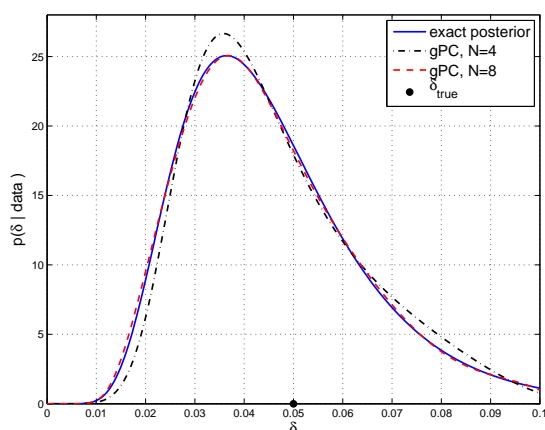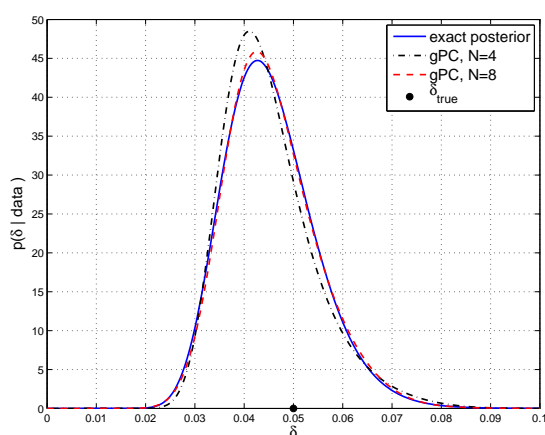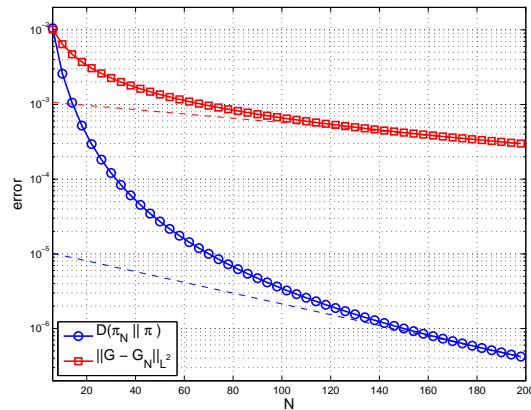
(a) $\nu = 0.05$.

(b) $\nu = 0.10$.

Figure 1: Posterior probability density of δ, the boundary condition of Burgers' equation, conditioned on observations of the transition layer location. Dash-dotted and dashed lines correspond to approximate posteriors $\widetilde{\pi}_N^d$ with gPC order $N=4$ and $N=8$, respectively; solid line is the exact posterior π^d.

5.2 Discontinuous forward model

The Burgers' example above yields a transition layer location z that, while nonlinear, is an infinitely smooth function of the input random parameter Z. For contrast, we consider a forward model whose output depends discontinuously on the input parameter. A simple but illustrative case is the step function $H(z)$ defined on $[-1,1]$,

$$H(Z) = \begin{cases} 0, & Z \in [-1,0], \\ 1, & Z \in (0,1]. \end{cases} \tag{5.2}$$

We take H to be our forward model (i.e., we put $G(Z) = H(Z)$) and use a single observation $d = G(z_{true}) + e$ to define a posterior density $\pi^d(z)$. As before, e is Gaussian;

(a) $\nu = 0.05$.
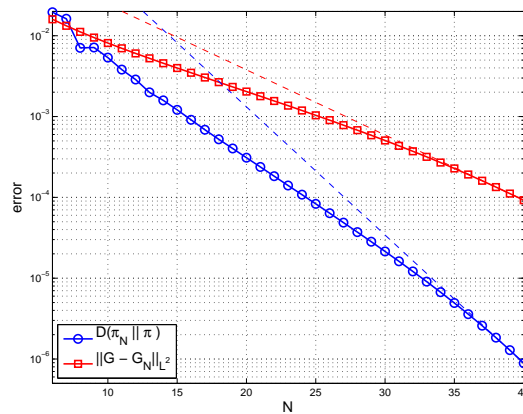
(b) $\nu = 0.10$.

Figure 2: Convergence of the forward model and the posterior density for boundary-condition inversion of Burgers' equation, with respect to gPC order N. Lines marked with squares show L^2 error in the gPC forward model, $\|G - \widetilde{G}_N\|^2_{L^2_{\pi_Z}}$; lines marked with circles show Kullback-Leibler divergence of the exact posterior from the approximate posterior, $D(\widetilde{\pi}^d_N \| \pi^d)$.

$e \sim N(0, \sigma^2)$ with $\sigma = 0.1$. The prior distribution on Z is uniform on the entire domain $[-1, 1]$. The original input $z_{true} = 0.2$, and thus we expect most posterior probability to lie on the right half of the domain.

Pseudo-spectral stochastic collocation is used to construct a Legendre polynomial chaos approximation to the forward model $G_N(Z)$. Fig. 3(a) shows the forward model approximation at $N = 9$. Oscillations characteristic of Gibbs' phenomena are observed; these are expected, given the discontinuity in the exact forward model. Fig. 3(b) shows the corresponding posterior density $\widetilde{\pi}^d_N$; here, the oscillations of the approximate forward model $\widetilde{G}_N$ are inherited and indeed amplified by the nonlinearity of the Gaussian density of e.
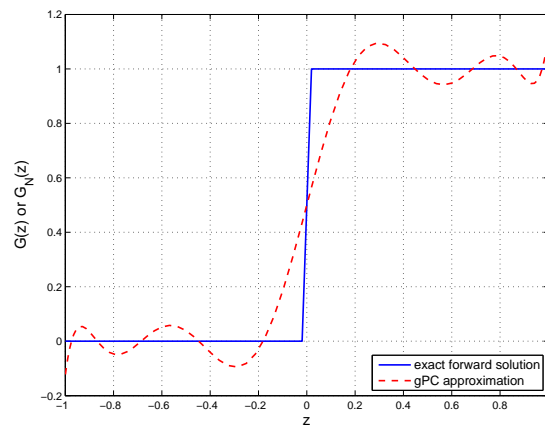
(a) Forward model approximation.
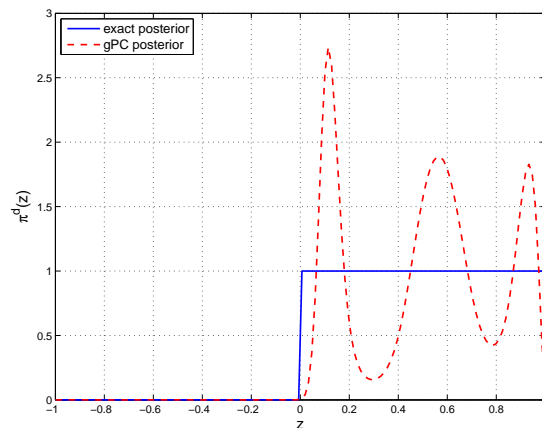
(b) Posterior density.

Figure 3: Exact and approximated forward model and posterior density for a step-function $G(Z)$, using gPC stochastic collocation with $N=9$.

Fig. 4 shows convergence of the forward model and posterior with respect to gPC order: we plot the L^2 error in the forward model, $\|G - \widetilde{G}_N\|^2_{L^2_{\pi_Z}}$, and the Kullback-Leibler divergence of the exact posterior from the approximate posterior, $D(\widetilde{\pi}^d_N \| \pi^d)$. $Q = 600$ collocation points are employed. Convergence is algebraic, since the forward model is not smooth. But the convergence rate of the Kullback-Leibler divergence still exceeds that of the L^2 error in the forward model, consistent with Theorem 4.1.

5.3 Genetic toggle switch

A larger-scale example is given by the kinetics of a genetic toggle switch, first synthesized in [7] and studied numerically in [25]. The toggle switch consists of two repressible
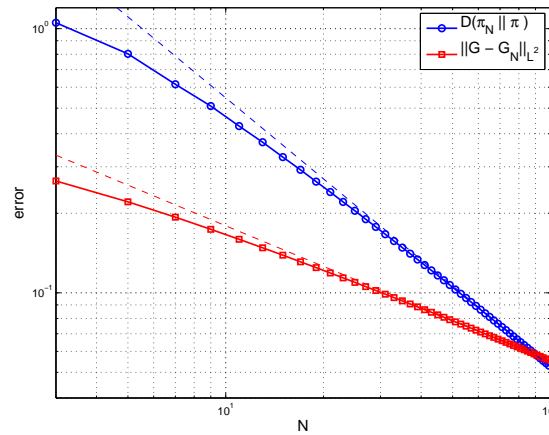
Figure 4: Errors in the forward model and posterior density approximations for a step-function $G(Z)$, as a function of gPC order N.

promotors arranged in a mutually inhibitory network: promoter 1 transcribes a repressor for promoter 2, while promoter 2 transcribes a repressor for promoter 1. Either repressor may be induced by an external chemical or thermal signal. Genetic circuits of this form have been implemented on *E. coli* plasmids, and the following differential-algebraic (DAE) model has been proposed [7]:

$$\frac{du}{dt} = \frac{\alpha_1}{1+v^\beta} - u,$$
$$\frac{dv}{dt} = \frac{\alpha_2}{1+w^\gamma} - v,$$
$$w = \frac{u}{(1+[\text{IPTG}]/K)^\eta}. \tag{5.3}$$

Here u is the concentration of the first repressor and v is the concentration of the second repressor; α_1 and α_2 are the effective rates of synthesis of the repressors; γ and β represent cooperativity of repression of promotors 1 and 2, respectively; and [IPTG] is the concentration of IPTG, the chemical compound that induces the switch. Parameters K and η describe binding of IPTG with the first repressor. At low values of [IPTG], the switch is in the 'low' state, reflected in low values of v; conversely, high values of [IPTG] lead to strong expression of v. Experimental measurements [7] of steady-state expression levels $v(t=\infty)$ are reproduced in Fig. 5. Observations over a wide range of IPTG concentrations clearly reveal the two-state behavior of the switch.

Values of the six parameters $Z_\theta = (\alpha_1, \alpha_2, \beta, \gamma, \eta, K) \in \mathbb{R}^6$ are highly uncertain. Nominal values were estimated in [7], but here we compute the joint posterior probability density of these parameters from experimental data. This density will reflect not only nominal values (e.g., mean or maximum *a posteriori* estimates), but variances, correlations, and any other desired measure of uncertainty in the parameter vector Z. Our data consist
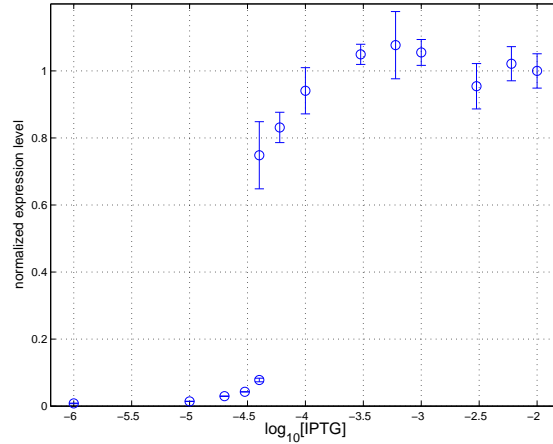
Figure 5: Experimental measurements of steady-state gene expression levels in the toggle switch, reproduced from [7]. Values are normalized by the mean expression level at the largest concentration of IPTG, and correspond to $v(t=\infty)$ in (5.3).

of normalized steady-state values of v at selected IPTG concentrations.[†] We choose one IPTG concentration on the 'low' side of the switch and five IPTG concentrations on the 'high' side. For the purpose of the present demonstration, we avoid IPTG values immediately around the transition region, as the DAE state there exhibits bimodality over the range of the prior, which is not efficiently captured by a global gPC expansion.

The experimental error is assumed Gaussian and zero-mean, but with a standard deviation that depends on whether the expression level is low or high: $\sigma_{\text{low}} = 4 \times 10^{-5}$, $\sigma_{\text{high}} = 5 \times 10^{-3}$. This simplified error model is consistent with experimental observations. Priors are uniform and independent for each parameter, centered at the nominal values $\bar{\theta}$ suggested by Gardner *et al.* in [7]. In other words, we have

$$Z_{\theta,i} = \bar{\theta}_i (1 + \varsigma_i Z_i),$$

where Z is a vector of six uniform random variables, $Z_i \sim U(-1,1)$. The entries of ς are $(0.20, 0.15, 0.15, 0.15, 0.30, 0.20)$.

Since the dimension of the problem ($n_z = 6$) renders a tensor-product formula impractical, we use sparse grid stochastic collocation to construct a gPC approximation $G_N(Z)$ of the forward model over the support of the prior. In particular, we employ a Smolyak algorithm based on a one-dimensional Clenshaw-Curtis quadrature rule [20,25]. The resulting 6-dimensional posterior distributions clearly cannot be visualized directly; rather, we must simulate samples from the posterior using Markov chain Monte Carlo (MCMC) [9]. Because the posterior distribution contains strong correlations among components of Z, along with differing scales of variability and sharp bounds, it is advantageous to use an adaptive MCMC algorithm. We thus employ the delayed-rejection adaptive Metropolis (DRAM) scheme of Haario *et al.* [10].

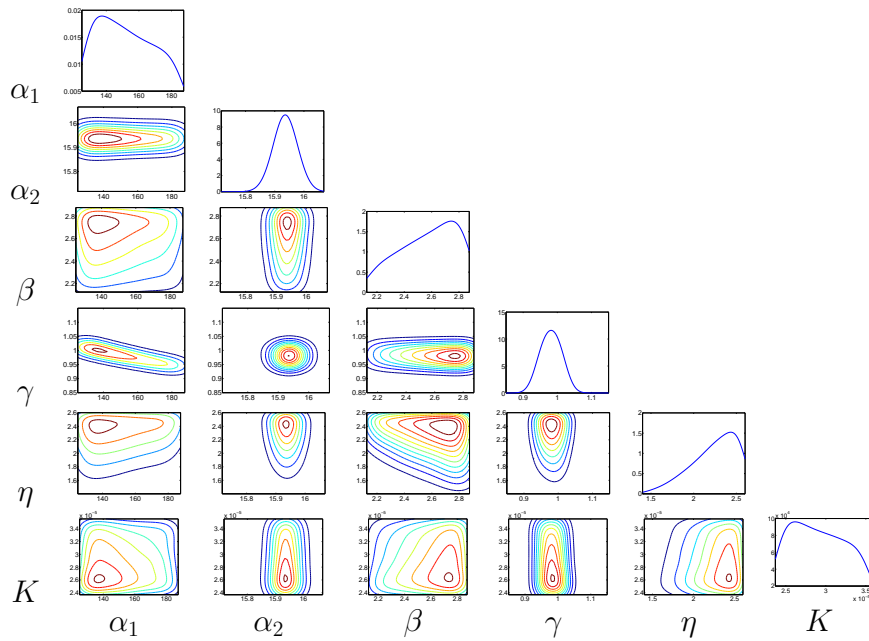[†]Experimental data are courtesy of T. S. Gardner.

Figure 6: 1-D and 2-D posterior marginals of parameters in the differential-algebraic model of a genetic toggle switch, conditioned on experimental data using the full forward model (i.e., with no gPC approximation).
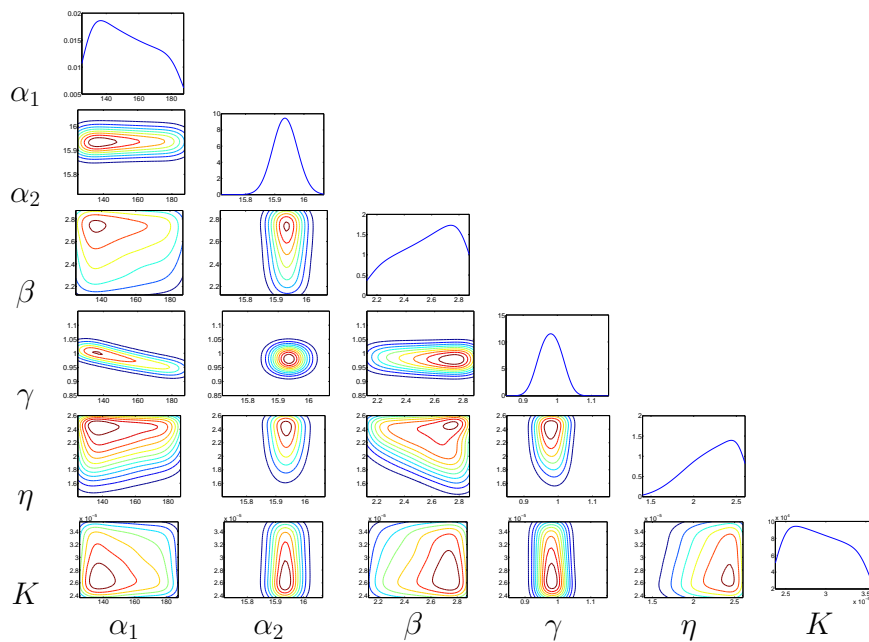

Figure 7: 1-D and 2-D posterior marginals of parameters in the differential-algebraic model of a genetic toggle switch, conditioned on experimental data using the stochastic collocation Bayesian approach with $N=3$.
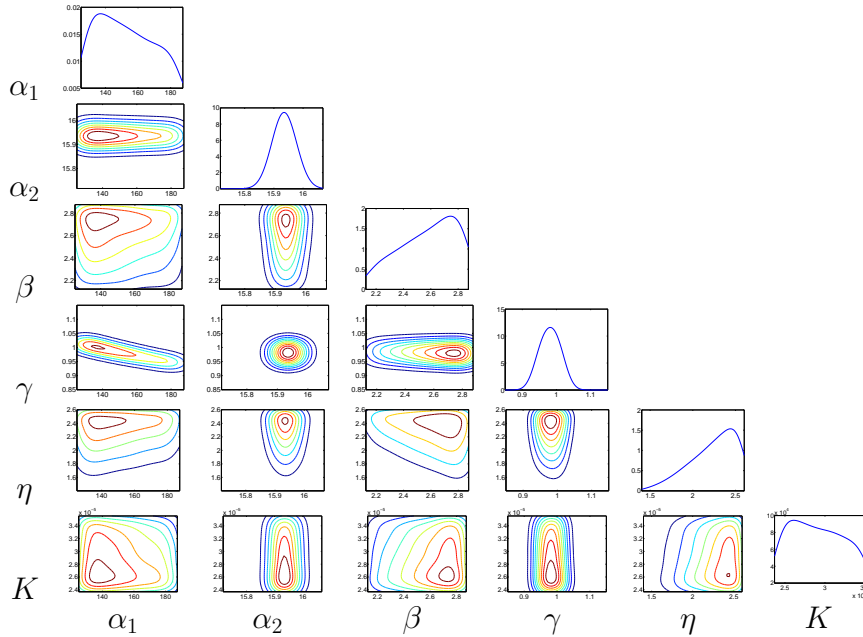
Figure 8: 1-D and 2-D posterior marginals of parameters in the kinetic model of a genetic toggle switch, conditioned on experimental data using the stochastic collocation Bayesian approach with $N=4$.

Posterior distributions for the toggle-switch parameters are shown in Figs. 6-8. Each figure contains all of the one- and two-parameter marginal densities of $\pi^d(z)$ or $\widetilde{\pi}_N^d(z)$. We obtain 3×10^5 MCMC samples in each case, and use kernel density estimation [12,19] to construct the marginal densities after discarding a "burn-in" interval of 10^4 samples. Fig. 6 shows results of simulation from the exact posterior π^d (2.8); these are obtained by applying MCMC with the exact forward model $G(Z)$. Fig. 7 depicts the approximate posterior $\widetilde{\pi}_N^d(Z)$ with gPC order $N=3$, using quadrature on a 5-level sparse grid to obtain the gPC coefficients. Fig. 8 shows $\widetilde{\pi}_N^d(z)$ with gPC order $N=4$, using quadrature on a 6-level sparse grid to obtain the coefficients. Excellent agreement with the true posterior is observed; indeed, the $N=3$ approximation appears sufficient to capture the essential features of the posterior distribution. We note that some of the parameters are not strictly identifiable within the prior range, while other parameters (e.g., α_2, γ) are endowed with much sharper posterior bounds on their possible values. Very strong correlation between α_1 and γ is also evident.

A simple check on the efficiency of MCMC sampling is provided by the empirical autocorrelation at lag s for each component of the Markov chain. Rapid decay of the autocorrelation is indicative of good "mixing": MCMC iterates are less correlated, and the variance of any MCMC estimate at a given number of iterations is reduced [9]. Autocorrelations for the present DRAM-based MCMC simulations are shown in Fig. 9; while these plots are specifically for simulation from $\widetilde{\pi}_N^d(z)$ with $N=4$, autocorrelation plots at
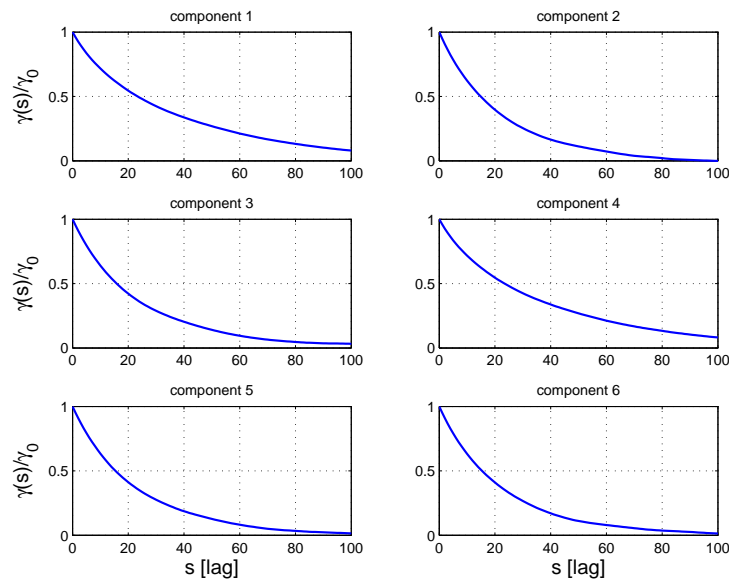
Figure 9: Autocorrelation at lag s of each component of the MCMC chain, for Bayesian estimation of parameters in the genetic toggle switch.

lower N or for simulation from the exact posterior π^d are comparable. The decay of $\gamma(s)$ suggests that 3×10^5 samples are sufficient for the present purposes; indeed, essentially no change in the posterior marginals of Figs. 6-8 is observed with further iterations.

The stochastic collocation Bayesian scheme provides dramatic gains in computational efficiency. Using $N = 3$ and a 5-level sparse grid, computing the gPC coefficients takes 1130 sec; this calculation involves forward simulations at 4865 collocation points. Subsequently, 467 sec are required to complete 3×10^5 DRAM iterations. By contrast, the same number of MCMC iterations on the exact posterior consumes approximately 80000 sec. We also note that calculation of the gPC coefficients may be considered an "offline" cost, performed before seeing any data. Restricting comparison to MCMC times alone, the stochastic collocation Bayesian inference scheme thus provides approximately a factor of 170 in speedup.

6   Conclusions

This paper develops efficient and accurate algorithms for the Bayesian solution of inverse problems, using generalized polynomial chaos (gPC) expansions and stochastic collocation. Given a prior distribution on model parameters or inputs, we use stochastic collocation to construct a gPC approximation of the forward model; gPC expansions then replace the full forward model in the likelihood function and define an approximate posterior probability density. The approximate posterior can be evaluated at arbitrary values of the inputs/parameters and for an arbitrarily large number of samples, at mini-

mal computational cost.

We prove the convergence of the approximate posterior to the true posterior, in terms of the Kullback-Leibler divergence (KLD), with increasing gPC order, and obtain an estimate of the rate of convergence. In particular, we show that the asymptotic convergence rate of the posterior density is at least the same as the L^2 convergence rate of the gPC expansion for the forward solution, and therefore if the gPC representation of the forward solution converges exponentially fast, so does the posterior density.

Convergence properties of our algorithm are then demonstrated numerically: first on an infinitely smooth problem involving parameter estimation in the viscous Burgers' equation, and second with a forward model exhibiting discontinuous dependence on its input. In both cases, consistency with the predicted convergence rates is obtained. We then present an example of kinetic parameter estimation from real experimental data, using stochastic collocation on sparse grids. The latter example shows the utility of sparse grid constructions for the solution of inverse problems in higher dimensions, and demonstrates that large computational speedups can be obtained with the present stochastic collocation Bayesian inference scheme.

Acknowledgments

References

[1] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. SIAM J. Numer. Anal., 45(3):1005–1034, 2007.
[2] I. Babuška, R. Tempone, and G.E. Zouraris. Galerkin finite element approximations of stochastic elliptic differential equations. SIAM J. Numer. Anal., 42:800–825, 2004.
[3] S. Balakrishnan, A. Roy, M. G. Ierapetritou, G. P. Flach, and P. G. Georgopoulos. Uncertainty reduction and characterization for complex environmental fate and transport models: an empirical Bayesian framework incorporating the stochastic response surface method. Water Resources Res., 39(12):1350, 2003.
[4] J. A. Christen and C. Fox. MCMC using an approximation. J. Comput. Graph. Stat., 14(4):795–810, 2005.
[5] Y. Efendiev, T. Y. Hou, and W. Luo. Preconditioning Markov chain Monte Carlo simulations using coarse-scale models. SIAM J. Sci. Comput., 28:776–803, 2006.
[6] S. N. Evans and P. B. Stark. Inverse problems as statistics. Inverse Problems, 18:R55–R97, 2002.

  [7] T.S. Gardner, C.R. Cantor, and J.J. Collins. Construction of a genetic toggle switch in Escherichia coli. Nature, 403:339–342, 2000.
  [8] R.G. Ghanem and P. Spanos. Stochastic Finite Elements: a Spectral Approach. Springer-Verlag, 1991.
  [9] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. Markov Chain Monte Carlo in Practice. Chapman and Hall, 1996.
 [10] H. Haario, M. Laine, A. Mira, and E. Saksman. DRAM: efficient adaptive MCMC. Statistics and Computing, 16:339–354, 2006.
 [11] D. Higdon, H. Lee, and C. Holloman. Markov chain Monte Carlo-based approaches for inference in computationally intensive inverse problems. Bayesian Statistics, 7:181–197, 2003.
 [12] A. T. Ihler. Kernel density estimation toolbox for MATLAB. http://www.ics.uci.edu/~ihler/code/kde.html.
 [13] J. Kaipio and E. Somersalo. Statistical and Computational Inverse Problems. Springer, 2005.
 [14] M. C. Kennedy and A. O'Hagan. Bayesian calibration of computer models. J. Royal Statist. Soc. Series B, 63(3):425–464, 2001.
 [15] O. Le Maitre, O. Knio, H. Najm, and R. Ghanem. Uncertainty propagation using Wiener-Haar expansions. J. Comput. Phys., 197:28–57, 2004.
 [16] Y. M. Marzouk and H. N. Najm. Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems. J. Comput. Phys., 228(6):1862–1902, 2009.
 [17] Y. M. Marzouk, H. N. Najm, and L. A. Rahn. Stochastic spectral methods for efficient Bayesian solution of inverse problems. J. Comput. Phys., 224(2):560–586, 2007.
 [18] A. Mohammad-Djafari. Bayesian inference for inverse problems. In Bayesian inference and Maximum Entropy Methods in Science and Engineering, 21:477–496, 2002.
 [19] B. W. Silverman. Density Estimation for Statistics and Data Analysis. Chapman & Hall, 1986.
 [20] S. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. Soviet Math. Dokl., 4:240–243, 1963.
 [21] Ch. Soize and R. Ghanem. Physical systems with random uncertainties: chaos representations with arbitrary probability measure. SIAM. J. Sci. Comput., 26(2):395–410, 2004.
 [22] A. Tarantola. Inverse Problem Theory and Methods for Model Parameter Estimation. SIAM, Philadelphia, 2005.
 [23] J. Wang and N. Zabaras. Hierarchical Bayesian models for inverse problems in heat conduction. Inverse Problems, 21:183–206, 2005.
 [24] J. Wang and N. Zabaras. Using Bayesian statistics in the estimation of heat source in radiation. Int. J. Heat Mass Trans., 48:15–29, 2005.
 [25] D. Xiu. Efficient collocational approach for parametric uncertainty analysis. Comm. Comput. Phys., 2(2):293–309, 2007.
 [26] D. Xiu. Fast numerical methods for stochastic computations: a review. Comm. Comput. Phys., 5:242–272, 2009.
 [27] D. Xiu and J. S. Hesthaven. High-order collocation methods for differential equations with random inputs. SIAM J. Sci. Comput., 27(3):1118–1139, 2005.
 [28] D. Xiu and G.E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. SIAM J. Sci. Comput., 24(2):619–644, 2002.
 [29] D. Xiu and G.E. Karniadakis. Modeling uncertainty in flow simulations via generalized polynomial chaos. J. Comput. Phys., 187:137–167, 2003.
 [30] D. Xiu and G.E. Karniadakis. Supersensitivity due to uncertain boundary conditions. Int. J. Numer. Meth. Engng., 61(12):2114–2138, 2004.