Stochastic Collocation Methods for Polynomial Chaos: Analysis and Applications

Dongbin Xiu

Department of Mathematics, Purdue University

Support: AFOSR FA9550-08-1-0353 (Computational Math)
NSF CAREER DMS-0645035 (Computational Math)
DOE DE-FC52-08NA28617 (PSAAP)
Overview

• Generalized polynomial chaos

• Stochastic collocation
  • Lagrange interpolation
  • Pseudo spectral gPC

• Applications
  • Bayesian inverse problem
  • Data assimilation
Uncertainty Quantification via gPC

• Stochastic PDE:
\[
\begin{aligned}
\frac{\partial u}{\partial t}(t, x, Z) &= \mathcal{L}(u), & (0, T] \times D \times \mathbb{R}^{n_z} \\
\mathcal{B}(u) &= 0, & [0, T] \times \partial D \times \mathbb{R}^{n_z} \\
u &= u_0(x, Z), & \{t = 0\} \times D \times \mathbb{R}^{n_z}
\end{aligned}
\]

• \(N\)th-order gPC expansion:
\[
u_N(t, x, Z) \triangleq \sum_{|k| = 0}^N \hat{u}_k(t, x)\Phi_k(Z), \quad \# \text{ of basis} = \begin{pmatrix} n_z + N \\ n_z \end{pmatrix}
\]
\[
\hat{u}_k = \mathbb{E}[u(Z)\Phi_k(Z)] = \int u(Z)\Phi_k(Z)\rho(Z)dZ, \quad 0 \leq |k| \leq N,
\]

• Orthogonal basis:
\[
\mathbb{E}_Z[\Phi_i(Z)\Phi_j(Z)] \triangleq \int \Phi_i(Z)\Phi_j(Z)\rho(Z)dZ = \delta_{ij}
\]

• Optimality:
\[
\|u - u_N\|_{L^2_\rho(Z)} = \inf_{\Psi \in \Pi_N^Z} \|u - \Psi\|_{L^2_\rho(Z)}
\]
Generalized Polynomial Chaos (gPC)

• **Basis functions:**
  - Hermite polynomials: seminal work by *R. Ghanem*
  - Global orthogonal polynomials (*Xiu & Karniadakis, 02*)
  - Wavelet basis (*Le Maitre et al, 04*)
  - Piecewise basis (*Babuska et al 04, Wan & Karniadakis, 05*)

• **Implementations:**
  - Stochastic Galerkin
  - Stochastic collocation

• **Properties:**
  - Rigorous mathematics
  - High accuracy, fast convergence
  - Curse-of-dimensionality
Stochastic Collocation

- **Collocation**: To satisfy governing equations at nodes

- **Sampling**: (solution statistics only)
  - Random (Monte Carlo)
  - Deterministic (lattice rule, tensor grid, cubature)

- **Stochastic collocation**: To construct polynomial approximations

- **Lagrange interpolation**
  - Can not be constructed for any given nodes
  - Interpolation error hard to control

- **Pseudo spectral**
  - Utilize gPC polynomial basis
  - Becomes multivariate integration

- **Response surface method**
  - Multivariate interpolation
  - Many ad hoc approaches
Stochastic Collocation – Lagrange Interpolation

• Nodal set: \[ \Theta_Q = \{ Z^i \}^Q \subseteq \mathbb{R}^{n_z} \]

• Lagrange interpolation: \[ u^Q(Z) \triangleq \sum_{j=1}^{Q} u(Z^j)L_j(Z) \quad L_i(Z^j) = \delta_{ij}, \quad 1 \leq i, j \leq Q \]

• Solution: for \( j=1, \ldots, Q \), \[ \frac{\partial u}{\partial t}(t, x, Z^j) = \mathcal{L}(u), \quad \text{in} \ (0, T] \times D, \]
\[ \mathcal{B}(u) = 0, \quad [0, T] \times \partial D, \]
\[ u = u_0(x, Z^j), \quad \{t = 0\} \times D \]

• Tensor product: \[ \left( U^{i_1} \otimes \ldots \otimes U^{i_{n_z}} \right) \]

• Sparse grid (Smolyak): \[ \sum_{q-N+1 \leq |\mathbf{i}| \leq q} (-1)^{q-|\mathbf{i}|} \binom{N-1}{q-|\mathbf{i}|} (U^{i_1} \otimes \ldots \otimes U^{i_{n_z}}) \]

(Xiu & Hesthaven, SIAM J. Sci. Comput., 05)
Stochastic Collocation: Pseudo Spectral Approach

• $N^\text{th}$-order gPC projection

$$u_N(t, x, Z) = \sum_{|k|=0}^{N} \hat{u}_k(t, x) \Phi_k(Z), \quad \hat{u}_k = \int u(Z) \Phi_k(Z) \rho(Z) dZ.$$  

• gPC-collocation approximation

$$w_N(t, x, Z) = \sum_{|k|=0}^{N} \hat{w}_k(t, x) \Phi_k(Z),$$

$$\hat{w}_k = \sum_{j=1}^{Q} u(t, x, Z_j^j) \Phi_k(Z_j^j) \alpha^j \approx \int u(Z) \Phi_k(Z) \rho(Z) dZ$$

$$\hat{w}_k(t, x) \rightarrow \hat{u}_k(t, x), \quad Q \rightarrow \infty$$

• Aliasing Error: \[ \varepsilon_Q \triangleq \left\| u_N - w_N \right\|_{L^2(Z)} \]

(Xiu, Comm. Comput Phys, vol. 2, 07)
1. Choose a nodal set $\{Z^j, \alpha^j\}_{j=1}^Q$ in $\mathbb{R}^{n_z}$

2. Solve for each $j = 1, \ldots, Q$,
\[
\frac{\partial u}{\partial t}(t, x, Z^j) = L(u), \quad \text{in } (0, T] \times D,
\]
\[
B(u) = 0, \quad [0, T] \times \partial D,
\]
\[
u = u_0(x, Z^j), \quad \{t = 0\} \times D
\]

3. Evaluate the approximate gPC expansion coefficient
\[
\hat{w}_k = \sum_{j=1}^Q u(t, x, Z^j) \Phi_k(Z^j) \alpha^j, \quad 0 \leq |k| \leq N;
\]

4. Construct the $N^{th}$-order gPC approximation
\[
w_N(t, x, Z) = \sum_{|k|=1}^N \hat{w}_k \Phi_k(Z).
\]

• Error bound (Xiu, 07):
\[
\varepsilon = \frac{\Delta t}{2} \left( \frac{\Delta x}{2} \right)^2 \leq \left( \frac{\varepsilon_N^2}{Q^2} + \frac{\varepsilon_Q^2}{C_Q^2} \right)^{1/2}
\]

Error $\leq$ Finite-term projection error + aliasing error + Numerical error
Parameter Estimation: Bayesian Inverse Approach

- **Stochastic PDE:**
  \[
  \begin{aligned}
  \frac{\partial u}{\partial t}(t, x, Z) &= \mathcal{L}(u), \quad (0, T] \times D \times \mathbb{R}^{n_z} \\
  \mathcal{B}(u) &= 0, \quad [0, T] \times \partial D \times \mathbb{R}^{n_z} \\
  u &= u_0(x, Z), \quad \{t = 0\} \times D \times \mathbb{R}^{n_z}
  \end{aligned}
  \]

- **Solution:**
  \[u(t, x, Z) : [0, T] \times \bar{D} \times \mathbb{R}^{n_z} \mapsto \mathbb{R}^{n_u}\]

- **Prior distribution:**
  \[\pi_Z(z) = \prod_{i=1}^{n_z} \pi_i(z_i)\]

- Estimation of the prior distribution
  - Requires direct measurements of the parameters
  - No/not enough direct measurements? (Use experience/intuition …)
  - How to take advantage of measurements of other variables?
Bayesian Inference

• **Data:** \( d = G(Z) + e, \quad e \in \mathbb{R}^{n_d} \) is i.i.d.

• **Posterior distribution:** \( \pi^d(Z) \triangleq \pi(Z \mid d) = \frac{\pi(d \mid Z)\pi(Z)}{\int \pi(d \mid Z)\pi(Z) dZ} \)

• **Likelihood function:** \( L(Z) \triangleq \pi(d \mid Z) = \prod_{i=1}^{n_d} \pi_{e_i}(d_i - G_i(Z)) \)

• **Notes:**
  - Difficult to manipulate
  - Classical sampling approaches can be time consuming (MCMC, etc)
  - GPC (Galerkin) based approach: \((Marzouk, Najm, Rahn, JCP, 07)\)

---

• **gPC approximation:**

\[
\pi^d_N(Z) = \frac{L_N(Z)\pi(Z)}{\int L_N(Z)\pi(Z) dZ}
\]

\[
L_N(Z) \triangleq \pi_N(d \mid Z) = \prod_{i=1}^{n_d} \pi_{e_i}(d_i - G_{N,i}(Z))
\]

• **Properties:**
  - Allows direct sampling in term of \( Z \) with arbitrarily large samples
  - (Virtually) no additional computational cost – forward problem solver only
  - Convergence seems natural
Convergence of gPC Bayesian Inference

- **Kullback-Leibler divergence:** 
  \[ D(p_1 \| p_2) \equiv \int p_1(z) \log \frac{p_1(z)}{p_2(z)} \, dz \]

- **Observation error:** 
  \[ e \sim N(0, \sigma^2 I), \text{ i.i.d. Normal} \]

**Theorem.** If the gPC expansion \( G_N \) converges to \( G \) in \( L^2_{\pi_Z} \), then the posterior density \( \pi^d_N \) converges to \( \pi^d \) in the sense

\[ D\left( \pi^d_N \| \pi^d \right) \to 0, \quad N \to \infty. \]

Moreover, if

\[ \left\| G_i(Z) - G_{N,i}(Z) \right\|_{L^2_{\pi_Z}} \leq CN^{-\alpha}, \quad 1 \leq i \leq n_d, \alpha > 0, \ C \text{ independent of } N, \]

then for sufficiently large \( N \),

\[ D\left( \pi^d_N \| \pi^d \right) \sim N^{-2\alpha}. \]

**Notes:**
- Fast (exponential) convergence rate is retained
- Factor of 2 in the convergence rates

*(Marzouk & Xiu, Comm. Comput. Phys. 08)*
Parameter Estimation: Supersensitivity Example

- Burgers’ equation: \( \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2}, \quad x \in [-1,1] \)

- Boundary conditions: \( u(-1) = 1 + \delta(Z); \quad u(1) = -1; \quad 0 < \delta << 1 \)

Deterministic results with no uncertainty


1% uncertainty in left BC
Prior distribution is uniform

Measurement noise: $e \sim N(0,0.05^2)$
$N_{\text{error}} \quad D(\pi_N \parallel \pi) ||G - G_N||_2$

Factor = 2.10 (theory = 2)
Parameter Estimation: Step Function

- Assume the forward model is a step function
- Posterior distribution is discontinuous
- Gibb’s oscillations exist
- Slow convergence with global gPC basis functions

Forward model and its approximation

Posterior distribution and its approximation
\begin{align*}
\text{Factor} = 1.99 \quad (\text{theory} = 2)
\end{align*}
Kalman Filter for Data Assimilation

• **True state (unknown):** \( u' \in \mathbb{R}^m, \quad m \geq 1 \)

• **Forecast:**
  \[
  \begin{aligned}
  &\frac{du^f}{dt}(t, Z) = F(t, u^f), \quad t \in (0, T] \\
  &u^f(0, Z) = u_0
  \end{aligned}
  \]

• **Observation:**
  \( d = Hu^f + \epsilon \in \mathbb{R}^\ell, \quad H : \mathbb{R}^m \rightarrow \mathbb{R}^\ell \)

• **Analysis:**
  \[ u^a = u^f + K(d - Hu^f) \]

  \[ K = P^fH^T(HP^fH^T + R)^{-1} \quad \text{(Kalman gain matrix)} \]

  \[ P^f = \mathbb{E}\left[ (u^f - u')(u^f - u')^T \right] \quad R = \mathbb{E}\left[ \epsilon \epsilon^T \right] \]

• **Properties:**
  
  • Straightforward for linear dynamic equations
  
  • Extension to nonlinear equations: Extended KF (EKF)
  
  • Optimal for Gaussian
  
  • Explicit calculation of covariance can be costly
Ensemble Kalman Filter (EnKF)

- **Ensemble:**
  \[
  (u^f)_i \triangleq u^f(t, Z^i), \quad i = 1, \ldots, M \quad (d)_i = d + (\varepsilon)_i, \quad i = 1, \ldots, N
  \]

  \[
  (u^a)_i = (u^f)_i + K_e \left[ (d)_i - H(u^f)_i \right], \quad i = 1, \ldots, M
  \]

  \[
  K_e = P^f e H^T (H P^f e H^T + R_e)^{-1}
  \]

  \[
  P^f e = (u^f - \bar{u}^f)(u^f - \bar{u}^f)^T \approx P^f \quad R_e = \varepsilon \varepsilon^T \approx R
  \]

- **Properties:**
  - nonlinear dynamics
  - sampling errors
    - Measurement. Can be eliminated by square-root filter (EnSRF)
    - Solution states.
  - Computational cost is of great concern
Error Analysis of the EnKF

• Assimilation step size: \( \Delta T = t_{n+1} - t_n \)

• **Lemma** (local error):

\[
e_{n+1} \leq \|M\| \cdot \|\mathbf{e}_{\Delta t}\| + \|\Delta \mathbf{K}\| \cdot \|\mathbf{e}^f\| + \|\Delta \mathbf{K}\| \cdot \|\mathbf{H}\| \cdot \|\mathbf{e}_{\Delta u}\|
\]

\[
\sim O\left(\Delta t^p, \sigma N^{-\alpha}\right)
\]

\[
\mathbf{M} = \mathbf{I} - \mathbf{KH} \quad \Delta \mathbf{K} = \mathbf{K}_e - \mathbf{K}
\]

• **Theorem** (global error):

\[
E_n \leq \left( E_0 + \sum_{k=1}^{n} e_k \right) \exp \left( \Lambda \cdot t_n \right)
\]

\[
\Lambda \propto \Delta T^{-1}
\]

• Note the **inverse dependence on assimilation step size**

*(Li & Xiu, vol. 197, CMAME 08)*
EnKF Example: Linear Wave Equation

Model description:

- Linear advection equation;
- Periodic domain of length $L=1000$;
- Wave speed = 1; grid spacing=1; time step = 1;
- True states are sampled from a Gaussian process, with zero mean and unit variance, and a spatial de-correlation length of 20. The dimension of random space is $n_z=50$.
- Four measurements uniformly in space are made every 5 time units.
- Measurement variance is 0.01.
- No model error.

$(x, Z) \in \mathbb{R} \times \mathbb{R}^{50}$

Long-term Wave propagation
Error Behavior of EnKF

w.r.t. ensemble size

w.r.t. data noise level
EnKF Error Behavior

qEnSRF: EnSRF combined with deterministic sampling using optimal cubature
GPC Collocation based Kalman Filter

Errors of assimilated results

<table>
<thead>
<tr>
<th></th>
<th>$T=100$</th>
<th>$T=500$</th>
<th>$T=1,000$</th>
<th>$T=1,500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EnKF ($N=100$)</td>
<td>$5.3 \times 10^{-3}$</td>
<td>$3.4 \times 10^{-3}$</td>
<td>$2.3 \times 10^{-3}$</td>
<td>$1.7 \times 10^{-3}$</td>
</tr>
<tr>
<td>EnKF ($N=10^3$)</td>
<td>$1.5 \times 10^{-3}$</td>
<td>$1.0 \times 10^{-3}$</td>
<td>$7.4 \times 10^{-4}$</td>
<td>$6.2 \times 10^{-4}$</td>
</tr>
<tr>
<td>EnKF ($N=10^4$)</td>
<td>$4.6 \times 10^{-4}$</td>
<td>$2.8 \times 10^{-4}$</td>
<td>$2.4 \times 10^{-4}$</td>
<td>$1.9 \times 10^{-4}$</td>
</tr>
<tr>
<td>gPC-KF ($N=51$)</td>
<td>$3.5 \times 10^{-4}$</td>
<td>$1.6 \times 10^{-4}$</td>
<td>$1.1 \times 10^{-4}$</td>
<td>$9.0 \times 10^{-5}$</td>
</tr>
<tr>
<td>gPC-KF ($N=100$)</td>
<td>$1.8 \times 10^{-4}$</td>
<td>$7.9 \times 10^{-5}$</td>
<td>$5.6 \times 10^{-5}$</td>
<td>$4.6 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

• 50 dimensional random space

(Li & Xiu, vol. 197, CMAME 08)
Accuracy Improvement of EnKF via gPC

• Use cubature – equally weighted

• Use pseudo-spectral gPC

  ▪ Analytical expression in $Z$

  $$u_N^f(t, Z) \triangleq \sum_{|k|=0}^{N} \hat{u}_k^f(t)\Phi_k(Z),$$

  ▪ Statistics

  $$\bar{u}_N^f = \hat{u}_0, \quad P_N^f = \sum_{0<|k|\leq N} \left(\hat{u}_k^f (\hat{u}_k^f)^T\right)$$

*(Li & Xiu, J. Comput. Phys. 08)*
GPC Based Ensemble Kalman Filter

• Ensemble:

\[
\left( u_N^f \right)_i = \sum_{|k|=0}^{N} \hat{u}_k^f (t) \Phi_k (Z_i), \quad i = 1, \ldots, M, \ M \gg 1
\]

• Square-root update:

\[
\left( u_N^f \right)_i = \bar{u}_N^f + \left( u_N^f \right)_i, \quad \left( u_N^a \right)_i = \bar{u}_N^a + \left( u_N^a \right)_i, \quad i = 1, \ldots, M
\]

- Mean state update:

\[
\bar{u}_N^a = \bar{u}_N^f + K_N \left( d - H \bar{u}_N^f \right), \quad K_N = P_N^f H^T \left( H P_N^f H^T + R \right)^{-1}
\]

- Perturbation update:

\[
\left( u_N^a \right)_i = \left( u_N^f \right)_i + \tilde{K}_N H \left( u_N^f \right)_i, \quad i = 1, \ldots, M
\]

\[
\tilde{K}_N = P_N^f H^T \left( \sqrt{H P_N^f H^T + R} \right)^{-1} \left( \sqrt{H P_N^f H^T + R + \sqrt{R}} \right)^{-1}
\]
Example: Nonlinear Population Dynamics

\[
\frac{du^f}{dt} = -r \left( 1 - \frac{u^f}{A} \right) u^f, \quad u^f(0) = u_0
\]
Error Convergence

\[ N=8, \ Q=10 \] is sufficient
Comparison: gPC-KF vs EnKF

![Graph showing the comparison between gPC-KF and EnKF]
Nonlinear System Example: Lorenz Equations

\[
\begin{align*}
\frac{dx}{dt} &= \sigma(y - x) \\
\frac{dy}{dt} &= \rho x - y - xz \\
\frac{dz}{dt} &= xy - \beta z
\end{align*}
\]

\[\sigma = 10, \ \rho = 28, \ \beta = 8/3\]

\[(x_0, y_0, z_0) = (1.508870, -1.531271, 25.46091)\]

Small deviation in initial condition (0.001 in \(x_0\)) causes large deviation
Qualitative Comparison: gPC EnKF vs EnKF

\[ ||\text{Estimate} - \text{True state}|| \]

- GPC EnKF: \( Q = 5^3 = 125 \)
- EnKF: ensemble size = 104
Summary

- Point selection is crucial for the efficacy of stochastic collocation

- GPC expansion is much more than a forward UQ method
  - Bayesian inverse (*Marzouk & Xiu, Comm. Comput. Phys, 08*)
  - Kalman filter for data assimilation (*Li & Xiu, CMAME 08; JCP 08*)