A GENERALIZED STOCHASTIC COLLOCATION APPROACH
FOR RANDOM CONTROL AND
PARAMETER IDENTIFICATION PROBLEMS

Max Gunzburger  Catalin Trenchea  Clayton Webster
Florida State University  University of Pittsburgh  Oak Ridge National Laboratory

SIAM-UQ14 – April, 2014

MG: US Department of Energy Office of Science (ASCR) and US Air Force Office of Scientific Research

CT: US Air Force Office of Scientific Research

CW: US Department of Energy Office of Science (ASCR) and US Air Force Office of Scientific Research
If you are suffering from the curse of dimensionality when solving stochastic PDE problems,
you are going to suffer a lot more if you want to solve inverse problems for stochastic PDEs.
Suppose
\[ J = \text{degrees of freedom for spatial discretization} \]
\[ M = \text{degrees of freedom for stochastic discretization} \]

- Intrusive methods
  e.g. polynomial chaos
  for solving stochastic PDEs involve the solution of a single \( JM \times JM \) system

- Non-intrusive methods
  e.g., Monte Carlo or stochastic collocation
  involve the solution of \( MJ \times J \) systems

- Curse of dimensionality \( \iff \) in practice, both \( J \) and \( M \) are usually large
In the deterministic setting, inverse problems also have “intrusive” and “non-intrusive” solution approaches.

- **Intrusive**
  - have to solve a single system with at least 2 or 3 times the number of spatial degrees of freedom used for the forward model.

- **Non-intrusive**
  - have to solve 2 or 3 systems of the same size as the forward model, but have to do it several, say $K$, times.

- Those who work on deterministic inverse problems constrained by PDEs also think the suffer from a curse of dimensionality!!!
• In the stochastic inverse problem setting
  
  – an intrusive/intrusive method requires the solution of a single $2JM \times 2JM$ or $3JM \times 3JM$ discrete system

  – a non-intrusive/non-intrusive method requires $K$ solutions of $2M \times 2M$ or $3M \times 3M$ discrete systems

• Of course, you could mix intrusive and non-intrusive

---

**Moral of the story**

if you want to work on stochastic inverse problems, you will twice cursed!!!
STOCHASTIC OPTIMAL CONTROL AND PARAMETER IDENTIFICATION PROBLEMS
The forward problem, i.e., the state equations

- We focus on the stochastic boundary value problem for the state $u(x, y)$
  \[
  \begin{align*}
  -\nabla \cdot (\kappa(x, y) \nabla u(x, y)) &= f(x, y) \quad \text{in } D \times \Gamma \\
  u(x, y) &= 0 \quad \text{on } \partial D \times \Gamma,
  \end{align*}
  \]
  where
  - $D$ = bounded spatial domain in $\mathbb{R}^d$, $d = 1, 2, \text{ or } 3$, having boundary $\partial D$
  - $\Gamma = [-1, 1]^N$ is a stochastic domain
  - $y = (y_1, y_2, \ldots, y_n) \in \Gamma$ a vector of independent random parameters having a joint PDF $\rho(y)$

- We have assumed that the noise is finite dimensional so that
  - the problem is posed in terms of a finite number of parameters
  or
  - we have that $\kappa$ and/or $f$ are random fields having infinite expansions which have been truncated to $N$ terms
In general, $f$ and $\kappa$ are independent so that it is understood that we have

- $f(x, y_1, \ldots, y_{N_f})$ with a PDF $\rho_f(y_f) : [-1, 1]^{N_f} \rightarrow \mathbb{R}$

- $\kappa(x, y_{N_f+1}, \ldots, y_N)$ with a PDF $\rho_\kappa(y_\kappa) : [-1, 1]^{N_\kappa} \rightarrow \mathbb{R}$

and then

- $N = N_f + N_\kappa$

- $y_f$ and $y_\kappa$ independent

- $\rho(y) = \rho_f(y_f)\rho_\kappa(y_\kappa)$
• Concerning the random forcing function $f$, we assume that\(^1\)

$$f(x, y) \in H^{-1}(D) \otimes L^2_\rho(\Gamma)$$

• Concerning the random coefficient $\kappa$, we assume that

$$\kappa(x, y) \in L^\infty(D) \otimes L^\infty(\Gamma)$$

and there exist $\kappa_{\text{min}} > 0$ and $\kappa_{\text{max}} < \infty$ such that

$$P\left[y \in \Gamma : \kappa_{\text{min}} \leq \kappa(x, y) \leq \kappa_{\text{max}} \; \forall x \in \overline{D}\right] = 1$$

• As a result, the state equations are almost surely well posed

\[ L^\rho_\rho(\Gamma) \implies \nu(y) \text{ such that } \int_\Gamma \nu(y) \rho(y) \, dy < \infty \]
• Control problem

given $\kappa$ and a functional $J(u, f)$, find $u$ and $f$ such that

$J(u, f)$ is minimized

and

$u, \kappa, \text{ and } f$ almost surely satisfy the state equations

• Parameter identification problem

given $f$ and a functional $J(u, \kappa)$, find $u$ and $\kappa$ such that

$J(u, \kappa)$ is minimized

and

$u, \kappa, \text{ and } f$ almost surely satisfy the state equations

• Of course, we could try to find both $f$ and $\kappa$
  - here we focus on the parameter identification problem
Matching functionals

- For given
  \[ \hat{u}(x, y) \in L^2(D) \otimes L^2_\rho(\Gamma) \]
  \[ \beta > 0 \]
  we have the functional
  \[ J_1(u, \kappa) = \mathbb{E}\left[ \frac{1}{2} \| u(\cdot, y) - \hat{u}(\cdot, y) \|_{L^2(D)}^2 + \frac{\beta}{2} \| \kappa(\cdot, y) \|_{L^2(D)}^2 \right] \]
  this functional is the expected value of a functional that has the form commonly used in deterministic parameter identification

- Note that
  \[ \| \mathbb{E}[u - \hat{u}] \|_{L^2(D)} \leq \mathbb{E}[\| u - \hat{u} \|_{L^2(D)}] \]
  this motivates the definition of the next functional
For given
\[ \hat{u}(x, y) \in L^2(D) \otimes L^2_{\rho}(\Gamma) \]
\[ \beta > 0 \]
we have the functional
\[ J_2(u, \kappa) = \frac{1}{2} \left\| \mathbb{E}[u](x) - \mathbb{E}[\hat{u}](x) \right\|_{L^2(D)}^2 + \frac{\beta}{2} \left\| \mathbb{E}[\kappa](x) \right\|_{L^2(D)}^2 \]

- this functional is a functional that is commonly used in deterministic parameter identification problems applied to the expected values of \( u, \kappa, \) and \( \hat{u} \)

- This functional is easily generalized to do moment matching

- this motivates the definition of the next functional
For given
\[ \hat{u}(x, y) \in L^2(D : L^{2Q}(\Gamma)) \]
\[ \beta > 0 \]
we have the functional
\[
J_3(u, \kappa) = \sum_{q=1}^{Q} \frac{1}{2q} \| \mathbb{E}[u^q] - \mathbb{E}[\hat{u}^q] \|_{L^2(D)}^2 + \frac{\beta}{2} \| \mathbb{E}[\kappa] \|_{L^2(D)}^2
\]
Stochastic parameter identification problems

- We have three parameter identification problems: for $i = 1, 2, 3$

  Problem $P_i$: \[
  \min_{(u, \kappa) \in A_i} J_i(u, \kappa)
  \]

- For all three cases, the existence (but not necessarily the uniqueness) of optimal pairs $(u_i, \kappa_i)$, $i = 1, 2, 3$, is proved.

- In all three cases, optimal pairs $(u_i, \kappa_i)$ are shown to be necessarily characterized as solutions of an optimality system:
  - the optimality system consists of
    - the state equations
    - adjoint equations
    - optimality condition
  - the optimality condition embodies the necessary condition that the gradient of the functional $J$ with respect to $\kappa$ must vanish for optimal $\kappa$. 
Optimality systems, optimal states $u_i(x, y)$, optimal coefficients $\kappa_i(x, y)$, and adjoint variables $\xi_i(x, y)$

- For Problem $P_1$
  - the state equations
    \[
    \begin{cases}
    -\nabla \cdot (\kappa_1 \nabla u_1) = f & \text{in } D \times \Gamma \\
    u_1 = 0 & \text{on } \partial D \times \Gamma
    \end{cases}
    \]
  - the adjoint equations
    \[
    \begin{cases}
    -\nabla \cdot (\kappa_1 \nabla \xi_1) = u_1 - \hat{u} & \text{in } D \times \Gamma \\
    \xi_1 = 0 & \text{on } \partial D \times \Gamma
    \end{cases}
    \]
  - the optimality condition
    \[
    \kappa_1 = \max \left\{ \kappa_{\min}, \min \left\{ \kappa_{\max}, \frac{1}{\beta} \nabla u_1 \cdot \nabla \xi_1 \right\} \right\} \quad \text{in } D \times \Gamma
    \]
For Problem $P_2$

- the state equations

\[
\begin{align*}
-\nabla \cdot (\kappa_2 \nabla u_2) &= f \quad \text{in } D \times \Gamma \\
\kappa_2 &= 0 \quad \text{on } \partial D \times \Gamma
\end{align*}
\]

- the adjoint equations

\[
\begin{align*}
-\nabla \cdot (\kappa_2 \nabla \xi_2) &= \mathbb{E}[u_2 - \hat{u}] \quad \text{in } D \times \Gamma \\
\xi_2 &= 0 \quad \text{on } \partial D \times \Gamma
\end{align*}
\]

- the optimality condition

\[
\kappa_2 = \max \left\{ \kappa_{\min}, \min \left\{ \kappa_{\max}, \frac{1}{\beta} \nabla u_2 \cdot \nabla \xi_2 \right\} \right\} \quad \text{in } D \times \Gamma
\]

Note that the only difference in the optimality systems for problems $P_1$ and $P_2$ is in the right-hand side of the adjoint equation.
Remark

- with $u(\cdot, y) \in H^1_0(D)$, the functional $J_3(u, \kappa)$ is well defined for all integers $Q \in [1, \infty)$ for spatial dimension $d = 1$ or $2$

- however, for $d = 3$, it is well defined only for $2Q \in [2, 6]$, i.e., for $Q \in [1, 3]$
  - in this case, one can only match up to the 3rd moments

- however, if $u(\cdot, y) \in H^2(D) \cap H^1_0(D)$, then, even in three dimensions, one can match all moments

- fortunately, with further assumptions on $D, f$, and $\kappa$, one can show that $u(\cdot, y) \in H^2(D) \cap H^1_0(D)$
DISCRETIZATION OF THE OPTIMALITY SYSTEM
- Finite element method for spatial discretization

- Stochastic collocation method for stochastic discretization
  - specifically, Clenshaw-Curtis (nested) sparse grids
NUMERICAL ILLUSTRATIONS
The optimal state and coefficient functions are solutions of the optimality systems.

As such, one can solve for approximations of those solutions by

1. discretizing the optimality system with respect to spatial dependence $\Rightarrow$ finite element method, finite difference method, etc.

2. discretizing the optimality system with respect to stochastic dependence $\Rightarrow$ polynomial chaos, Monte Carlo, sparse grid, etc.

3. solving the discrete optimality system as a single coupled system or by interatively cycling through the components of the system.

- Polynomial chaos + the one-shot approach $\Rightarrow$ intrusive/intrusive method

- Stochastic collocation + iterative method $\Rightarrow$ non-intrusive/non-intrusive method
  - we use this type of approach
    stochastic collocation + gradient method optimization
- $D = (0, 1)$

- $y_n$ is uniformly distributed on $[0, 1]$ for $n = 1, \ldots, N$

- Target function $\hat{u}(x, y) = x(1 - x^2) + \sum_{n=1}^{N} \sin(2n\pi x) y_n$

- Let $\hat{\kappa}(x, y) = (1 + x^3) + \sum_{n=1}^{N} \cos \left( \frac{n\pi x}{2N} \right) y_n$

- From $\hat{u}$ and $\hat{\kappa}$, we manufacture the right-hand side function
  \[ f(x, y) = -\nabla \cdot (\hat{\kappa} \nabla \hat{u}) \]

- As a result, the target function is attainable
  - i.e., the exact solutions of the optimization problems are
    \[ u_i = \hat{u} \quad \text{and} \quad \kappa_i = \hat{\kappa} \]
Realizations and their expectations for
the target  \( \hat{u} \)
the true coefficient  \( \hat{\kappa} \)
the corresponding forcing function  \( f \)

\( N = 5 \) and 241 Clenshaw-Curtis sparse grid points
Realizations of $\hat{u}$ and their mean

Realizations of $\hat{v}$ and their mean

Corresponding realizations of $f$ and their mean
Comparison of using the functionals $J_1$ and $J_2$

$N = 1$ and 5 collocation points

$E[\hat{u}]$ (red)  $E[u_1]$ (dashed)  $E[u_2]$ (solid)

$E[\kappa], E[\kappa_1], E[\kappa_2]$

The exact solution $E(\hat{u}(x))$ versus $E(u_4(x))$ using SC

The exact solution $E(\kappa(x))$ versus $E(k^4(x))$ using SC
Comparison of using the functionals $J_1$ and $J_2$

$N = 5$ and 61 collocation points

$E[\hat{u}]$ (red)  
$E[u_1]$ (dashed)  
$E[u_2]$ (solid)  

$E[\hat{\kappa}]$, $E[\kappa_1]$, $E[\kappa_2]$

The exact target $E[M(x)]$ versus $E[\psi(x)]$ using SC

The exact coefficient $E[\sigma(x)]$ versus $E[\kappa(x)]$ using SC
Comparison of using the functionals $J_1$ and $J_2$

Comparison between stochastic collocation and Monte Carlo sampling

$N = 11$  
265 collocation points  
$1.2 \times 10^6$ Monte Carlo samples

---

The exact targets $\mathbb{E}[u(x)]$ versus $\mathbb{E}[w(x)]$

- SC (black): $\mathbb{E}[u_1]$ (dashed) $\mathbb{E}[u_2]$ (solid)
- MC (blue): $\mathbb{E}[u_1]$ (dashed) $\mathbb{E}[u_2]$ (solid)

The exact coefficients $\mathbb{E}[\hat{u}(x)]$ versus $\mathbb{E}[\hat{\kappa}(x)]$

- SC: $\mathbb{E}[\kappa_1], \mathbb{E}[\kappa_2]$
- MC: $\mathbb{E}[\kappa_1], \mathbb{E}[\kappa_2]$
Convergence of the functionals vs. number of gradient method iterations

$N = 11$

![Convergence of the cost functionals](image)

- Black: $J_2$ with 265 collocation points
- Dashed red: $J_1$ with 265 collocation points
- Dot-dashed red: $J_1$ with 2069 collocation points
- Dot red: $J_1$ with 12,497 collocation points
Cost comparison between SC and MC

<table>
<thead>
<tr>
<th>N</th>
<th>SG</th>
<th>MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>61</td>
<td>7e+03</td>
</tr>
<tr>
<td>11</td>
<td>1.581</td>
<td>9e+06</td>
</tr>
<tr>
<td>21</td>
<td>13,329</td>
<td>8e+09</td>
</tr>
</tbody>
</table>

Number of deterministic solutions required to reduce the value of the functional $J_2$ by a factor of $10^4$ using Clenshaw-Curtis sparse grid sampling and Monte Carlo sampling.
NUMERICAL ANALYSIS
• In addition to analyzing the model stochastic inverse problems to answer questions about the existence of optimal solutions and to characterize those solutions as satisfying optimality systems, we have analyzed fully-discrete discretization of the problems.

• In particular, we have derived error estimates for finite element spatial discretizations and sparse grid stochastic discretization.

• Results will be included in a forthcoming paper.