

MIT Open Access Articles

Model Order Reduction for Stochastic Optimal Control

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Nguyen, Ngoc-Hien, et al. "Model Order Reduction for Stochastic Optimal Control." ASME 2012 11th Biennial Conference on Engineering Systems Design and Analysis, 2-4 July, 2012, Nantes, France, ASME, 2012, p. 599.

As Published: <http://dx.doi.org/10.1115/ESDA2012-82061>

Publisher: ASME International

Persistent URL: <http://hdl.handle.net/1721.1/115152>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



ESDA2012-82061

MODEL ORDER REDUCTION FOR STOCHASTIC OPTIMAL CONTROL

Ngoc-Hien Nguyen*

Singapore-MIT Alliance
National University of Singapore
Singapore 117576

Karen Willcox

Massachusetts Institute of Technology
Cambridge, Massachusetts 02139-4307, USA

Boo Cheong Khoo

Department of Mechanical Engineering
National University of Singapore
Singapore 117576

ABSTRACT

This work presents an approach to solve stochastic optimal control problems in the application of flow quality management in reservoir systems. These applications are challenging because they require real-time decision-making in the presence of uncertainties such as wind velocity. These uncertainties must be accounted for as stochastic variables in the mathematical model. In addition, computational costs and storage requirements increase rapidly due to the stochastic nature of the simulations and optimisation formulations. To overcome these challenges, an approach is developed that uses the combination of a reduced-order model and an adjoint-based method to compute the optimal solution rapidly. The system is modelled by a system of stochastic partial differential equations. The finite element method together with collocation in the stochastic space provide an approximate numerical solution—the “full model”, which cannot be solved in real-time. The proper orthogonal decomposition and Galerkin projection technique are applied to obtain a reduced-order model that approximates the full model. The conjugate-gradient method with Armijo line-search is then employed to find the solution of the optimal control problem under the uncertainty of input parameters. Numerical results show that the stochastic control yields solutions that are above the bound of the set solutions of the deterministic control. Applying the reduced model to the stochastic optimal control problem yields a speed-up in computational time by a factor of about 80 with acceptable accuracy

in comparison with the full model. Application of the optimal control strategy shows the potential effectiveness of this computational modeling approach for managing flow quality.

INTRODUCTION

Finding the solution of optimal flow control problems can be a computationally expensive undertaking. For simulations to support real-time decision-making in applications governed by partial differential equations (PDEs), the discretized models may have many thousands or even millions degrees of freedom. The situation is even more challenging for stochastic control problems because of the presence of uncertainties such as wind velocity. These uncertainties must be accounted for as stochastic variables in the mathematical model. The computational costs and storage requirements increase rapidly due to the stochastic nature of the simulations and optimization formulation. In such situations, the use of traditional discretization methods, such as finite element or finite volume methods, to achieve real-time simulations may be infeasible.

The goal of this work is to present an efficient computational approach to solve stochastic optimal control problems. The approach uses the combination of a reduced-order model (ROM) and an adjoint-based method to compute the optimal solution rapidly. The system is modelled by a system of stochastic partial differential equations (SPDEs). The adjoint-based method, also known as the one-shot or Lagrange multiplier method, has

*Correspondence to: g0801993@nus.edu.sg

been widely used in optimization and optimal control problems [1]. For stochastic optimal control problems, simulated solutions need to be evaluated repeatedly over different realizations of the uncertain input parameters. Our “full model” utilizes the finite element method [2] together with sparse grid stochastic collocation to approximate solution of the SPDEs.

In the collocation framework, candidate solutions are computed at sample points in the multi-dimensional stochastic space. The global solution of the SPDEs is then represented using interpolation functions [3–5]. The Smolyak algorithm provides a minimal number of collocation points to construct the interpolation functions, which for many problems leads to efficient and accurate representation of the stochastic solutions [3,6]. The sparse grid collocation method has been widely applied for stochastic applications, such as natural convection problems [7], source inversion and flow through porous media [8].

Model order reduction techniques aim to reduce the dimension of a state-space system, while retaining the characteristic dynamics of the system and preserving the input-output relationship [9]. Many large-scale model reduction frameworks are based on a projection approach. The idea is to approximate any solution of the PDEs of interest as a linear combination of solutions that have been pre-computed and to project the large-scale governing equations onto the subspace spanned by a reduced-space basis, yielding a low-order dynamical system. The most popular technique to find the basis is the proper orthogonal decomposition (POD). POD provides an orthogonal basis for a set of data, which may be theoretical, experimental or computational data. Sirovich introduced the method of snapshots, where each snapshot contains spatial data obtained from numerical simulation at a fixed time, as a efficient way for determining the POD basis vectors for large-scale problems [10].

This paper is outlined as follows. In the next two sections, we briefly introduce the mathematical model, the numerical method for stochastic optimal control of the problems of interest, and the model reduction techniques. In the subsequent section, we use the numerical example to demonstrate the solution of stochastic control problems and the reduced-order model performance. We provide some concluding remarks in the final section.

STOCHASTIC OPTIMAL CONTROL PROBLEM SETUP

Suppose that $\mathcal{D} \subset \mathbb{R}^d$, $d = 1, 2, 3$ is a physical domain with boundary Γ . Let the diffusivity $\kappa(\mathbf{x}, \mathbf{t}; \omega)$ be function mapping the product space $\mathcal{D} \times [t_0, t_f] \times \Omega \rightarrow \mathbb{R}$, where $\mathbf{x} \in \mathcal{D}$ denotes the spatial coordinates and $\mathbf{t} \in [t_0, t_f]$ denotes time. The randomness of the diffusivity is contained in $\omega \in \Omega$, where Ω is the sample space. A contaminant concentration which is represented by a function $\mathbf{c} : \equiv \mathbf{c}(\mathbf{x}, \mathbf{t}; \omega)$ satisfies the stochastic parabolic differential equation (SPDE), boundary conditions and initial conditions

as follows:

$$\frac{\partial \mathbf{c}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{c} - \nabla \cdot (\kappa(\mathbf{x}, \mathbf{t}; \omega) \nabla \mathbf{c}) = \mathbf{f} \quad \text{in } \mathcal{D} \times [t_0, t_f], \quad (1)$$

$$\mathbf{c} = \mathbf{g} \quad \text{on } \Gamma_D \times [t_0, t_f], \quad (2)$$

$$\frac{\partial \mathbf{c}}{\partial \mathbf{n}} = 0 \quad \text{on } \Gamma_N \times [t_0, t_f], \quad (3)$$

$$\mathbf{c}(\mathbf{x}, t_0; \omega) = \mathbf{c}_0(\mathbf{x}; \omega) \quad \text{in } \mathcal{D}, \quad (4)$$

where \mathbf{f} is the external source and \mathbf{c}_0 the given initial condition. The inlet boundary Γ_D is subjected to a Dirichlet condition, while the remainder of the boundary $\Gamma_N = \Gamma \setminus \Gamma_D$ satisfies Neumann conditions. The velocity field $\mathbf{u} \in \mathbb{R}^d$ in the convective term is used as a control parameter. In general, \mathbf{u} can be a function of \mathbf{x} and \mathbf{t} , although in this work we take it to be constant. For any given \mathbf{u} , one can solve for the solution $\mathbf{c}(\mathbf{x}, \mathbf{t}; \omega)$.

The goal of our control problem is to flush the contaminant out of the domain by controlling the velocity of the fluid pump. The objective functional is to seek a velocity over an admissible control set $\mathbf{u} \in \mathbf{U}_{ad}$ that minimizes a weighted combination of the L_2 -norm of the expected contaminant field and the velocity field:

$$\min_{\mathbf{u} \in \mathbf{U}_{ad}} \mathcal{J} = \frac{1}{2} \int_{t_0}^{t_f} \mathbb{E} [\|\mathbf{c}(\mathbf{x}, \mathbf{t}; \omega)\|_{L_2}^2] d\mathbf{t} + \frac{\beta}{2} \int_{t_0}^{t_f} \|\mathbf{u}\|_{L_2}^2 d\mathbf{t}, \quad (5)$$

subject to the constraints Eqns. (1)–(4). Here, β is a constant controlling the relative weighting of the components of the objective function and $\mathbb{E}[\cdot]$ denotes the expectation operator.

Stochastic Collocation Method

In the collocation framework, the stochastic problem is transformed into a parameterized family of deterministic PDEs using an assumption of finite-dimensional noise [4, 11]. The approximation of the SPDE solution is then computed based on a weighted combination of the solutions at each sample in the collocation space. Under the finite-dimensional noise assumption, the uncertain diffusivity field κ can be re-written as $\kappa(\mathbf{x}, \mathbf{t}; \omega) = \kappa(\mathbf{x}, \mathbf{t}; \mathbf{Y}(\omega))$. Here $\mathbf{Y}(\omega) = \{\mathbf{Y}_i(\omega)\}_{i=1}^{N_Y}$ are independent random variables. The multi-dimensional stochastic space Θ^{N_Y} is then defined based on vector $\mathbf{Y}(\omega)$. For more details, refer to [5]. We then represent κ as

$$\kappa(\mathbf{x}, \mathbf{t}; \mathbf{Y}) = \mathbb{E}[\kappa](\mathbf{x}, \mathbf{t}) + \sum_{i=1}^{N_Y} \kappa_i(\theta) \mathbf{Y}_i(\omega), \quad (6)$$

where the κ_i are now deterministic functions and θ represents the coordinates in the stochastic space. This expansion in Eqn. (6)

could be computed for example using the Karhunen-Loève decomposition [12].

We now define a collocation space $\mathbb{P}^{P-1}(\Theta) \subset \mathbf{L}^2(\Theta^{N_Y})$ as the span of tensor product polynomials with degree at most $P-1$. The collocation space has two attributes: the collocation points $\{\theta^k = (\xi, \eta)\}_{k=1}^P \subset \Theta$ and the collocation weights $\{\mathbf{w}^k\}_{k=1}^P$. The uncertain diffusivity field κ in Eqn. (6) can now be considered as a function of variable θ^k if the random vector $\mathbf{Y}(\omega)$ is fixed. As a result, the stochastic collocation requires evaluation of the solution $\mathbf{c}(\mathbf{x}, \mathbf{t}; \mathbf{Y})$ at each collocation point $\{\theta^k = (\xi, \eta)\}_{k=1}^P$. Hence, the SPDE problem with an uncertain input parameter is now written as a deterministic parameterized PDE where θ is the input parameter. Let $\mathbf{c}_k(\mathbf{x}, \mathbf{t}; \mathbf{Y})$ be the solution of the deterministic PDE at each θ^k . The solution of the SPDE is a global approximation constructed by linear combination of the solution at collocation points.

$$\mathbf{c}_F(\mathbf{x}, \mathbf{t}; \mathbf{Y}) = \sum_{k=1}^P \mathbf{c}_k(\mathbf{x}, \mathbf{t}; \mathbf{Y}) \mathbf{L}_k(\theta), \quad (7)$$

where $\mathbf{L}_k(\theta)$ is the Lagrange interpolation function corresponding to the k^{th} collocation point.

Finite Element Approximations

The finite element method (FEM) [2] is employed to obtain a semi-discrete set of equations with the following form

$$\mathbf{M}\dot{\mathbf{c}} + (\mathbf{C}(\mathbf{u}) + \mathbf{K}(\mathbf{t}; \theta^k))\mathbf{c} = \mathbf{F}, \quad (8)$$

$$\mathbf{c}(\mathbf{t}_0; \mathbf{Y}) = \mathbf{c}_0(\mathbf{Y}). \quad (9)$$

Here, $\mathbf{c}(\mathbf{t}; \mathbf{Y}) \in \mathbb{R}^N$ is the discretized approximation of $\mathbf{c}(\mathbf{x}, \mathbf{t}; \mathbf{Y})$ and contains N state unknowns. $\dot{\mathbf{c}}$ is the derivative of \mathbf{c} with respect to time. $\mathbf{M} \in \mathbb{R}^{N \times N}$ is the mass matrix, $\mathbf{C}(\mathbf{u}) \in \mathbb{R}^{N \times N}$ is the convective matrix, $\mathbf{K}(\mathbf{t}; \theta^k) \in \mathbb{R}^{N \times N}$ is the stiffness matrix, and $\mathbf{F} \in \mathbb{R}^N$ is the external source. Here, N is the number of grid points and θ^k the k^{th} collocation point.

We now consider optimal control with the cost functional as given in Eqn. (5). In the collocation framework, the expected value is approximated via a quadrature rule (such as Clenshaw-Curtis quadrature [13]) built on the collocation points. Define $\{\mathbf{w}^k\}_{k=1}^P$ to be the quadrature weights associated with the collocation points,

$$\mathbf{w}^k = \int_{\Theta} \rho(\mathbf{Y}) \mathbf{L}_k^2(\theta) d\theta, \quad \text{for } k = 1, \dots, P, \quad (10)$$

where $\rho(\mathbf{Y})$ is the probability density of the random vector \mathbf{Y} . The cost functional is replaced by the discretized problem as follows

low

$$\min_{\mathbf{u} \in \mathbf{U}_{\text{ad}}} \hat{\mathcal{J}}(\mathbf{u}) = \frac{1}{2} \int_{t_0}^{t_f} \sum_{k=1}^P \mathbf{w}^k \mathbf{c}^T(\mathbf{t}; \mathbf{Y}) \mathbf{M} \mathbf{c}(\mathbf{t}; \mathbf{Y}) d\mathbf{t} + \frac{\beta}{2} \int_{t_0}^{t_f} \|\mathbf{u}\|_{L_2}^2 d\mathbf{t}. \quad (11)$$

Here, the solution $\mathbf{c}(\mathbf{t}; \mathbf{Y}), k = 1, \dots, P$, solves the ordinary differential equations (ODEs) (8)-(9).

The Optimality System

We introduce the Lagrange multiplier functional with the adjoint state $\mathbf{p}(\mathbf{t}; \mathbf{Y})$ and adjoint initial condition $\chi \in \mathbb{R}^N$ as follows

$$\begin{aligned} \mathcal{L}(\mathbf{c}, \mathbf{u}, \mathbf{p}, \chi) = & \hat{\mathcal{J}}(\mathbf{u}) - \chi^T (\mathbf{c}(\mathbf{t}_0; \mathbf{Y}) - \mathbf{c}_0(\mathbf{Y})) \\ & - \mathbf{p}^T (\mathbf{M}\dot{\mathbf{c}} + (\mathbf{C}(\mathbf{u}) + \mathbf{K}(\mathbf{t}; \theta^k))\mathbf{c} - \mathbf{F}). \end{aligned} \quad (12)$$

The first-order necessary conditions, known as the Karush-Kuhn-Tucker (KKT) optimality conditions, are determined by taking variations with respect to $\mathbf{c}, \mathbf{p}, \mathbf{u}$ and χ of Eqn. (12).

Taking variations with respect to \mathbf{p} and χ recovers the state equation and initial condition:

$$\mathbf{M}\dot{\mathbf{c}} + (\mathbf{C}(\mathbf{u}) + \mathbf{K}(\mathbf{t}; \theta^k))\mathbf{c} = \mathbf{F}, \quad (13)$$

$$\mathbf{c}(\mathbf{t}_0; \mathbf{Y}) = \mathbf{c}_0(\mathbf{Y}). \quad (14)$$

Taking variations with respect to \mathbf{c} gives the adjoint equation with a final time condition:

$$-\mathbf{M}^T \dot{\mathbf{p}} + (\mathbf{C}^T(\mathbf{u}) + \mathbf{K}^T(\mathbf{t}; \theta^k))\mathbf{p} = \mathbf{M}\mathbf{c}, \quad (15)$$

$$\mathbf{p}(\mathbf{t}_f; \mathbf{Y}) = 0. \quad (16)$$

Taking variations with respect to \mathbf{u} yields the optimality condition,

$$\frac{\delta \mathcal{L}}{\delta \mathbf{u}} = \beta \int_{t_0}^{t_f} \mathbf{u} d\mathbf{t} - \int_{t_0}^{t_f} \sum_{k=1}^P \mathbf{w}^k \mathbf{c}^T(\mathbf{u}) \mathbf{p} d\mathbf{t} = 0. \quad (17)$$

In summary, the state equation, adjoint equation and optimality condition form the optimality system, solutions of which provide the optimal state \mathbf{c} , co-state \mathbf{p} and control variable \mathbf{u} .

To solve the KKT system, the Crank-Nicolson method [14] is used to discretize the state, adjoint and optimality condition

equations in time. The conjugate gradient method [15] is employed to solve the linearized system; the Armijo line-search [16] is used to ensure convergence. Algorithm 1 summarizes the procedure to solve the stochastic optimal control problem using the collocation method.

Algorithm 1

1. Initial work
 - 1a. Given \mathbf{P} , \mathcal{D} , Θ , initial velocity \mathbf{u}_0 , tolerance ε . Set $\mathbf{j} = 0$
 - 1b. Given the FEM basis ϕ_l for $l = 1, \dots, N$, where N is the number of grid points
 - 1c. Compute the matrices \mathbf{M} , $\mathbf{C}(\mathbf{u})$, and vector \mathbf{F}
 - 1d. Compute collocation point $\{\theta^k = (\xi, \eta)\}_{k=1}^P$ and collocation weights $\{\mathbf{w}^k\}_{k=1}^P$.
2. Solve the KKT system

For $k = 1 : P$

 - 2a. Compute the inputs $\kappa(\mathbf{x}, \mathbf{t}; \mathbf{Y})$ at each θ^k
 - 2b. Compute stiffness matrix $\mathbf{K}(\mathbf{t}; \theta^k)$ where $\mathbf{K}(\mathbf{t}; \theta^k) = \int_{\mathcal{D}} \kappa(\mathbf{x}, \mathbf{t}; \mathbf{Y}) \nabla \phi_l(\mathbf{x}) \cdot \nabla \phi_l(\mathbf{x}) d\mathbf{x}$
 - 2c. Solve the state equation with input \mathbf{u}_j
 - 2d. Solve the adjoint equation
 - 2e. Store results

end
3. Compute the optimal control
 - 3.a Compute the cost-functional $\hat{\mathcal{J}}(\mathbf{u}_j)$ and the gradient $\text{grad}(\mathbf{u}_j)$
 - 3.b If $\|\text{grad}(\mathbf{u}_j)\| < \varepsilon \rightarrow \text{stop}$.
 - 3.c Perform Armijo line search

Set $\mathbf{s}_j = -\text{grad}(\mathbf{u}_j)$

Set $\alpha_j = 1$ then evaluate $\hat{\mathcal{J}}(\mathbf{u}_j + \alpha_j \mathbf{s}_j)$, and $\text{gtol} = 10^{-4} \alpha_j \mathbf{s}_j^T \text{grad}(\mathbf{u}_j)$

While $\hat{\mathcal{J}}(\mathbf{u}_j + \alpha_j \mathbf{s}_j) > \hat{\mathcal{J}}(\mathbf{u}_j) + \text{gtol}$

Set $\alpha_j = \alpha_j / 2$ and evaluate $\hat{\mathcal{J}}(\mathbf{u}_j + \alpha_j \mathbf{s}_j)$.
 - 3.d Set $\mathbf{u}_{j+1} = \mathbf{u}_j + \alpha_j \mathbf{s}_j$, and $\mathbf{j} = \mathbf{j} + 1$. Go to step 2.

Discretization of the KKT system in space yields a high-dimensional discrete state-space system in the form of ODEs (Eqns. (13)–(16)). In addition, the collocation method and optimal control work require evaluating repeatedly the solutions of both the state and adjoint equations. Thus, these simulations in real-time are computationally expensive and may not be feasible. Model order reduction is applied to obtain a reduced-order approximation of the large model, which allows efficient simulations.

MODEL ORDER REDUCTION

Reduced-order modeling has been widely used in computational fluid dynamics for the simulation of large-scale systems.

Applications involving repeated evaluations of the model (such as inverse problems and control problems) become computationally expensive in the large-scale setting. To reduce the computational costs and storage requirements, model order reduction can be used to replace the large-scale models with approximate models of lower dimensions that capture the essential characteristics of the full models.

Reduction via Projection

We consider the system of ODEs as they appeared in Eqns. (13)–(14). We also interest in the output of contaminant solution at sensor locations in the domain, which is given by

$$\mathbf{y} = \mathbf{B}\mathbf{c}, \quad (18)$$

where matrix $\mathbf{B} \in \mathbb{R}^{N_0 \times N}$ and vector $\mathbf{y}(\mathbf{t}; \mathbf{Y}) \in \mathbb{R}^{N_0}$ contains the N_0 outputs of the system. A reduced order model of this system can be derived by approximating the full state vector \mathbf{c} as a linear combination of \mathbf{m} basis vectors as follows,

$$\mathbf{c} \approx \mathbf{V}\mathbf{c}_r, \quad (19)$$

where $\mathbf{c}_r \in \mathbb{R}^m$ is the reduced order state and $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_m] \in \mathbb{R}^{N \times m}$ is an orthonormal basis, i.e., $\mathbf{V}^T \mathbf{V} = \mathbf{I}$. Projecting the governing system (13)–(14) onto the reduced space formed by the column span of basis \mathbf{V} yields the lower order model in (20)–(22)

$$\mathbf{M}_r \dot{\mathbf{c}}_r + (\mathbf{C}_r(\mathbf{u}) + \mathbf{K}_r(\mathbf{t}; \theta^k)) \mathbf{c}_r = \mathbf{F}_r, \quad (20)$$

$$\mathbf{c}_r(\mathbf{t}_0, \mathbf{Y}) = \mathbf{c}_{0r}(\mathbf{Y}), \quad (21)$$

$$\mathbf{y}_r = \mathbf{B}_r \mathbf{c}_r, \quad (22)$$

where

$$\mathbf{M}_r = \mathbf{V}^T \mathbf{M} \mathbf{V}, \quad (23)$$

$$\mathbf{K}_r(\mathbf{t}; \theta^k) = \mathbf{V}^T \mathbf{K}(\mathbf{t}; \theta^k) \mathbf{V}, \quad (24)$$

$$\mathbf{C}_r(\mathbf{u}) = \mathbf{V}^T \mathbf{C}(\mathbf{u}) \mathbf{V}, \quad (25)$$

$$\mathbf{F}_r = \mathbf{V}^T \mathbf{F}, \quad (26)$$

$$\mathbf{B}_r = \mathbf{B} \mathbf{V}, \quad (27)$$

$$\mathbf{c}_{0r}(\mathbf{Y}) = \mathbf{V}^T \mathbf{c}_0(\mathbf{Y}). \quad (28)$$

The model reduction task is then to find a suitable basis \mathbf{V} so that $m \ll N$. In the literature there exist various methods for the computation of proper basis in the case of large-scale system, such as balanced truncation, Krylov-subspace and POD methods. This study will consider POD as the method to compute the basis.

Proper Orthogonal Decomposition

Proper orthogonal decomposition (POD) provides a method to compute the reduced-order basis \mathbf{V} and construct the low-order system by projection as described above. Here we briefly describe the general POD method (more details may be found in [10]).

Let $\mathbf{X} = [\mathbf{c}^1(\mathbf{t}_1; \cdot) \mathbf{c}^1(\mathbf{t}_2; \cdot) \cdots \mathbf{c}^1(\mathbf{t}_T; \cdot) \mathbf{c}^2(\mathbf{t}_1; \cdot) \cdots \mathbf{c}^S(\mathbf{t}_T; \cdot)] \in \mathbb{R}^{N \times Q}$ be a collection of a total Q snapshot state solutions $\mathbf{c}^s(\mathbf{t}_j; \cdot)$, $j = 1, \dots, T$, where T is the number of time steps, of the system in (13) for $s = 1, \dots, S$ input parameters. The POD basis is optimal in the sense that vectors \mathbf{V} are chosen to maximize the averaged projection of $\mathbf{c}(\mathbf{t}; \cdot)$ onto \mathbf{V} , suitably normalized

$$\max_{\mathbf{V}} \frac{\langle \mathbf{c}, \mathbf{V} |^2 \rangle}{\|\mathbf{V}\|^2}, \quad (29)$$

where $|\cdot|$ is the inner product of basis vector \mathbf{V} with the field \mathbf{c} , $\langle \cdot \rangle$ the time averaged operator and $\|\cdot\|$ the L_2 norm.

The POD basis vectors are the m left singular vectors of \mathbf{X} corresponding to the largest singular values ($m \leq Q$). Let σ_i , $i = 1, 2, \dots, Q$ be the singular values of \mathbf{X} in non-increasing order. We determine the number of POD vectors to retain in the reduced-order model by choosing $m \leq Q$ vectors so that

$$\sum_{i=1}^m \sigma_i^2 / \sum_{j=1}^Q \sigma_j^2 \geq \varepsilon_E, \quad (30)$$

where $\varepsilon_E(\%)$ is the required amount of energy, typically taken to be 99% or higher.

NUMERICAL EXAMPLE

We present the 2D mathematical model to which we apply stochastic optimal control with the full model using Algorithm 1. Then we apply model reduction to obtain the reduced-order model. We compare the stochastic optimal control result using the reduced model and the full model. Finally, we compare the behavior of the stochastic control with a deterministic control strategy.

Model Setup

In order to implement the contaminant transport problem, we consider the computational domain as in Figure 1. The domain is rectangular with $\mathcal{D} = [0, 1] \times [0, 0.5]$. The inflow boundary, which is defined on $\mathbf{x} = 0, 0 \leq y \leq 0.5$, satisfies a homogeneous Dirichlet condition, Γ_D ; the remaining boundaries satisfy homogeneous Neumann conditions, Γ_N . The velocity vector with x and y -component is chosen as uniform and constant in time, given by $\mathbf{u} = [\mathbf{u} \ \mathbf{v}]^T$. A velocity of $\mathbf{u} = [1 \ 0]^T$ is used

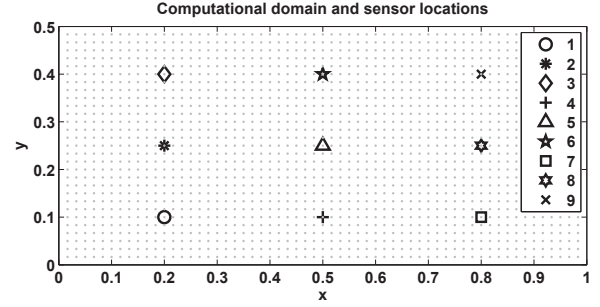


FIGURE 1. THE COMPUTATIONAL DOMAIN WITH $N_0 = 9$ SENSORS.

as an initial guess for finding an optimal velocity. In this example, we discretize the KKT system on a $\mathbf{n}_x \times \mathbf{n}_y = 61 \times 31$ grid, where \mathbf{n}_x and \mathbf{n}_y are the number of grid points in x and y -direction, respectively. This results in $N = 1891$ spatially discrete unknowns using the standard finite element method. The Crank-Nicolson method is employed to discretize the system in time, where $\mathbf{t} \in [\mathbf{t}_0, \mathbf{t}_f]$ with $\mathbf{t}_0 = 0, \mathbf{t}_f = 1.4$ and the time-step size $\Delta \mathbf{t} = 0.02$ or $T = 70$ time steps.

The input is a random diffusivity field κ . To generate the diffusivity coefficients under the finite dimensional noise assumption, we use the formulation similar to that in [5]. The random diffusivity coefficient is a nonlinear function of the random vector \mathbf{Y} , namely

$$\kappa(\mathbf{x}, \mathbf{t}; \mathbf{Y}) = \kappa_0 + \exp \left\{ \left[\mathbf{Y}_1(\omega) \cos(\pi \eta) + \mathbf{Y}_3(\omega) \sin(\pi \eta) \right] e^{-\frac{1}{8}} + \left[\mathbf{Y}_2(\omega) \cos(\pi \xi) + \mathbf{Y}_4(\omega) \sin(\pi \xi) \right] e^{-\frac{1}{8}} \right\} / \sigma_Y. \quad (31)$$

Here, $\theta = (\xi, \eta) \in \mathbb{P}$ are the coordinates of the collocation points. We choose $\kappa_0 = 1/125, \sigma_Y = 200$. The initial Péclet number $\text{Pe}_0 = \frac{\|\mathbf{u}\|L}{\kappa_0} = 125$, where the length of the domain is used as the characteristic length $L = 1$. The real random variables $\mathbf{Y}_n, n = 1, \dots, 4$ are independent and identically distributed with zero mean value and unit variance.

The source function $\mathbf{f}(\mathbf{x}, \mathbf{t})$ is described as a Gaussian distribution as follows

$$\mathbf{f}(\mathbf{x}, \mathbf{t}) = \sum_{k=1}^{n_s} \frac{\mathbf{h}_k}{2\pi\sigma_{sk}^2} \exp \left(-\frac{|\bar{\mathbf{x}}_k - \mathbf{x}|^2}{2\sigma_{sk}^2} \right) \delta(\mathbf{t} - \mathbf{t}_{0k}). \quad (32)$$

Here, we choose the number of sources to be $n_s = 1$, located at $\bar{\mathbf{x}}_1 = (\mathbf{x}_c, \mathbf{y}_c) = (0.3, 0.25)$, with the strength $\mathbf{h}_1 = 1$ and width $\sigma_{s1} = 0.05$. The active time of the source is $\mathbf{t}_{01} \in [\mathbf{t}_0, \mathbf{t}_{off}]$ with $\mathbf{t}_{off} = 0.4$.

Figure 2 shows the contaminant solution $\mathbf{c}(\mathbf{x}, \mathbf{t}; \kappa_0)$ of the full model at specific times. The contaminant field increases in

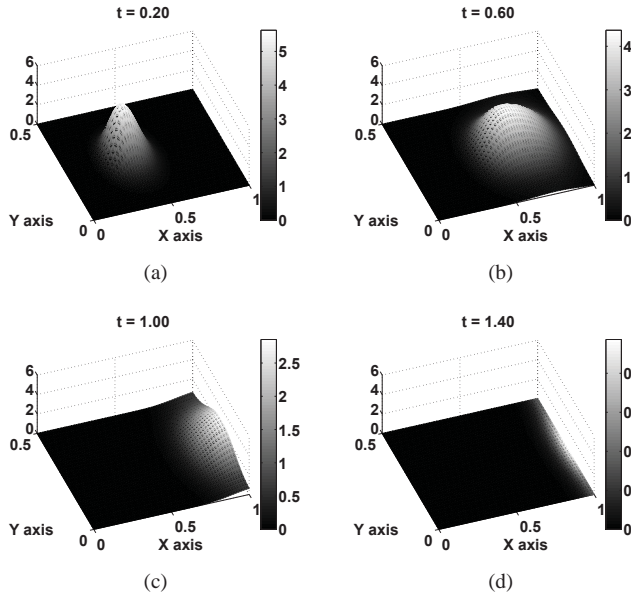


FIGURE 2. CONTAMINANT FIELD OF FULL MODEL AT SPECIFIC TIMES.

magnitude while the source is active. After the shutoff time of the source, the contaminant moves away and spreads out due to convection and diffusion until it flows out of the domain.

Full Stochastic Control Model

The stochastic optimal control now can be solved by following Algorithm 1 as described above. To illustrate the behavior of the collocation, we simulate the unbounded random variables Y_n via the Gaussian density distribution function. We employ the Smolyak algorithm [4–6] to determine the collocation points and collocation weights. We evaluate the optimal solution with Smolyak nodes which represent exactly polynomials of total degree 5 ($P = 29$), degree 7 ($P = 65$), degree 9 ($P = 145$) and degree 11 ($P = 321$) as shown in Figure 3. To estimate the relative error of the solution, we choose the solution corresponding to the finest collocation scheme ($P = 321$) as a “truth” solution. The relative error of the estimated optimal velocity is given as:

$$\epsilon_{\text{error}} = \frac{\|\mathbf{u}_{\text{truth}} - \mathbf{u}\|_2^2}{\|\mathbf{u}_{\text{truth}}\|_2^2}. \quad (33)$$

We then set the control parameter $\beta = 0.1$. Table 1 shows the results of the optimal control with different numbers of collocation points. Figure 4 shows the relative error of the stochastic optimal control solutions based on the finest solution. When the number of collocation points increases, the relative error in the estimated optimal solution decreases. However the computational time also increases when the number of collocation points

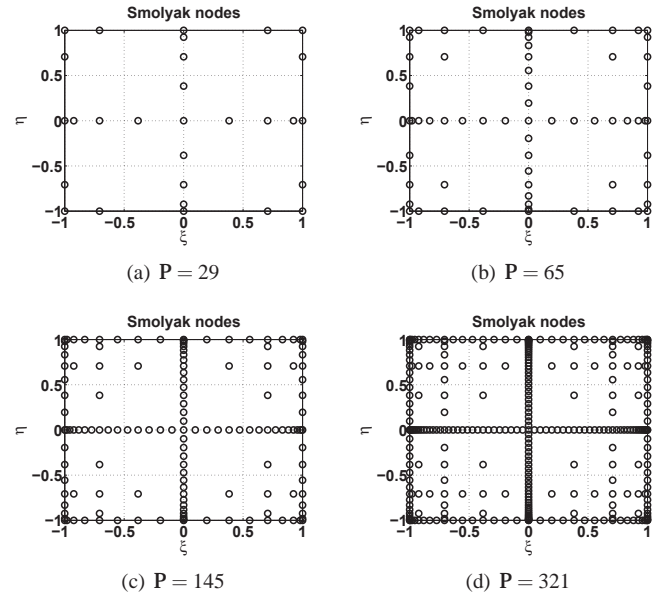


FIGURE 3. THE SMOLYAK QUADRATURE NODES.

TABLE 1. ESTIMATED OPTIMAL CONTROL FOR DIFFERENT NUMBERS OF COLLOCATION POINTS.

P	u	v	$\hat{\mathcal{J}}$	time (s)
29	1.35012	0.00434	0.31215	3.2638e+3
65	1.35089	0.00483	0.31188	8.1935e+3
145	1.35121	0.00489	0.31197	1.8046e+4
321	1.35124	0.00489	0.31198	3.5432e+4

increases. We observe that the computational time is approximately 6 hours when $P = 321$ Smolyak nodes.

Reduced Stochastic Control Model

To generate the snapshots needed for the POD basis, we choose N_k evenly-spaced samples, κ_t , on the interval $[\kappa_{\min}, \kappa_{\max}]$. In this example, $N_k = 10$. To determine an appropriate number of POD modes we use the energy capture as in Eqn. (30). Table 2 shows the relative error of the approximation (for a randomly chosen value of κ not in the snapshot set) for different sizes of the reduced-order model. In practice, we need both the dimensions of the reduced-order model and the relative error to be small. Here, we choose the case with $\epsilon_E = 99.99\%$ yielding a POD basis of size $m = 46$. The outputs of interest are the values of contaminant solution c at selected sensor locations. The outputs of the full model, y , and reduced model of order $m = 46$, y_r , are shown in Figure 5 at four different sensor locations. These lo-

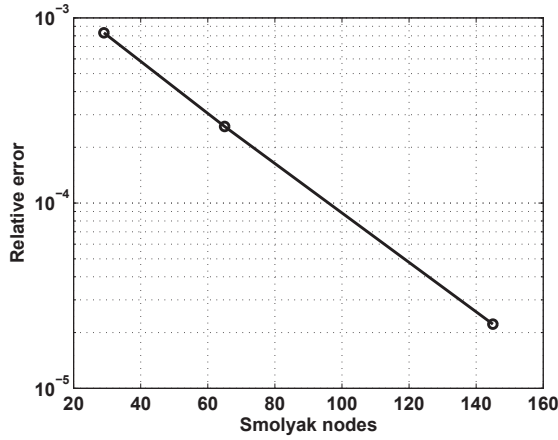


FIGURE 4. RELATIVE ERROR OF THE ESTIMATED STOCHASTIC CONTROL SOLUTION WITH NUMBER OF COLLOCATION POINTS.

TABLE 2. PROPERTIES OF VARIOUS MOR MODELS.

$\varepsilon_E(\%)$	POD	$\varepsilon_{\text{state}}$	$\varepsilon_{\text{adjoint}}$
99.0	18	5.48143e-3	1.10702e-2
99.5	21	3.45869e-3	5.28466e-3
99.9	30	6.28582e-4	8.11021e-4
99.99	46	1.13589e-4	1.80933e-4
99.999	65	2.05810e-5	6.68017e-5
99.9999	86	6.85014e-6	2.39363e-5

cations correspond to sensors 4, 5, 6, and 8 in Figure 1. It can be seen that the magnitude of the sensor reading varies depending on the location of the sensor relative to the source. In all cases the reduced-order model is able to capture well the behavior of the full model at the sensor locations.

Applying Algorithm 1 for the reduced-order model, we obtain the optimal result as in Table 3. The comparison of accuracy and computational time between the full model and reduced model are given in Table 4. The reduced model of order $m = 46$ has a relative error around 10^{-5} . The computational time is reduced by approximately 80 times in comparison with full control model.

Stochastic Control vs. Deterministic Control

To make the comparison between the stochastic control and deterministic control, we choose the solution of the stochastic control at the degree of polynomial 9 or $P = 145$ Smolyak nodes. We then choose a subset of Smolyak nodes in the collocation

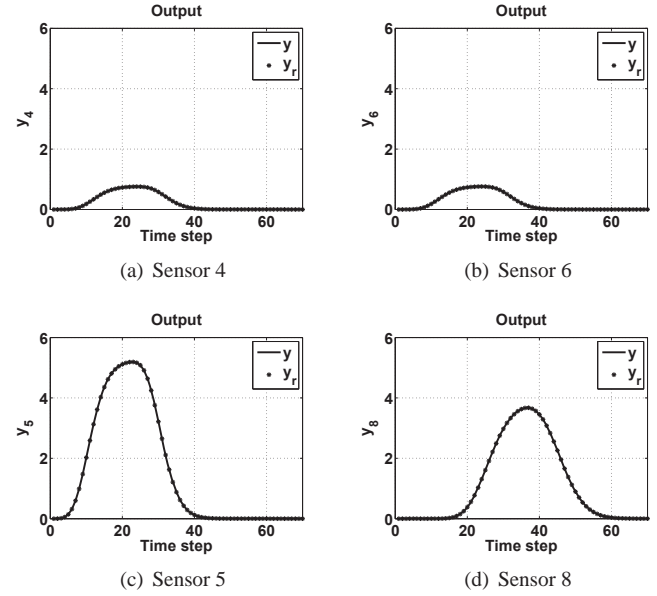


FIGURE 5. A COMPARISON OF THE FULL MODEL ($N = 1891$) AND REDUCED MODEL ($m = 46$) OUTPUT OF INTEREST AT SENSOR LOCATIONS.

TABLE 3. OPTIMAL CONTROL OF REDUCED MODEL.

P	u	v	\mathcal{J}	time (s)
29	1.35019	0.00513	0.31212	40
65	1.35087	0.00512	0.31185	98
145	1.35124	0.00510	0.31196	214
321	1.35122	0.00510	0.31196	460

TABLE 4. FULL MODEL VS. REDUCED ORDER MODEL.

P	ε_u	$\frac{\text{time}_{\text{Full}}}{\text{time}_{\text{MOR}}}$
29	5.3899e-5	81
65	1.4014e-5	83
145	2.2777e-5	84
321	1.4227e-5	76

space \mathbb{P} , for example we choose $\mathbb{P}_S \in \mathbb{P}$ such that $-1 \leq \xi \leq 1$ and $\eta = -1$. For each pair $\theta^k = (\xi, \eta)$ we compute the diffusivity coefficient $\kappa(\mathbf{x}, t; \mathbf{Y})$. We then compute the deterministic optimal control for the mean value of κ to find the optimal velocity and estimate its cost functional. Figure 6 shows that the stochastic optimal control always has the value above the average of the set of deterministic control.

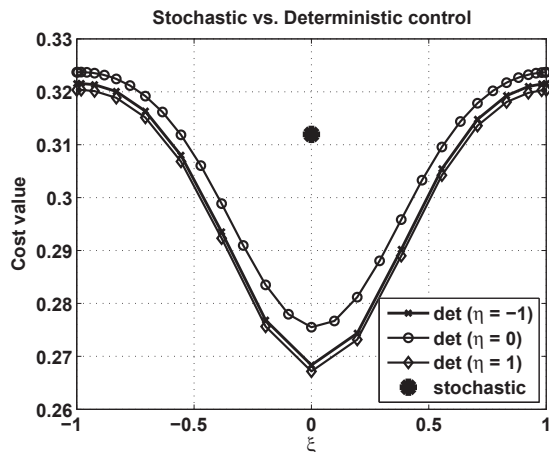


FIGURE 6. STOCHASTIC VS. DETERMINISTIC CONTROL.

CONCLUSION

This study has applied the combination of model order reduction techniques based on POD and an adjoint-based method to solve a stochastic optimal control problem. The reduced model with order $m = 46$ decreases the computational time of solution by a factor of about 80 while retaining acceptable accuracy with a relative error around 10^{-5} as compared to the full model with size $N = 1891$. This speed up is important in real-time decision-making applications because it provides a rapid solution and reduces time cost and storage requirements. Application of the optimal control strategy shows the potential effectiveness of this computational modeling approach for managing flow quality.

ACKNOWLEDGMENT

This work was supported by the Singapore-MIT Alliance.

REFERENCES

- [1] Gunzburger, M., 2003. **Perspectives in Flow Control and Optimization** (Advances in Design and Control). SIAM.
- [2] Zienkiewicz, O. C., and Morgan, K., 1983. **Finite Elements and Approximation**. Wiley, New York.
- [3] Ma, X., and Zabarar, N., 2009. "An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations". *Journal of Computational Physics*, **228**(8), pp. 3084–3113.
- [4] Nobile, F., Tempore, R., and Webster, C., 2008. "A sparse grid stochastic collocation method for partial differential equations with random input data". *SIAM Journal on Numerical Analysis*, **46**(5), May, pp. 2309–2345.
- [5] Babuska, I., Nobile, F., and Tempore, R., 2007. "A stochastic collocation method for elliptic partial differential equations with random input data". *SIAM Journal on Numerical Analysis*, **45**(3), pp. 1005–1034.

- [6] Bungartz, H. J., and Griebel, M., 2004. "Sparse grids". *Acta Numerica*, **13**, pp. 1–123.
- [7] Ganapathysubramanian, B., and Zabarar, N., 2007. "Sparse grid collocation schemes for stochastic natural convection problems". *Journal of Computational Physics*, **225**(1), July, pp. 652–685.
- [8] Ma, X., and Zabarar, N., 2009. "An efficient bayesian inference approach to inverse problems based on an adaptive sparse grid collocation method". *Inverse Problems*, **25**(3), February, pp. 013–035.
- [9] Antoulas, A. C., 2005. **Approximation of Large-Scale Dynamical Systems**. SIAM Advances in Design and Control, Philadelphia.
- [10] Sirovich, L., 1987. "Turbulence and the dynamics of coherent structures. part 1: Coherent structures". *Quarterly of Applied Mathematics*, **45**(3), October, pp. 561–571.
- [11] Kouri, D. P., 2010. "Optimization Governed by Stochastic Partial Differential Equations". MS Thesis, Rice University, Houston, Texas, May.
- [12] Ghanem, R. G., and Spanos, P. D., 1991. **Stochastic Finite Elements: A Spectral Approach**. Springer - Verlag, New York.
- [13] Clenshaw, C. W., and Curtis, A. R., 1960. "A method for numerical integration on an automatic computer". *Numerische Mathematik*, **2**, October, pp. 197–205.
- [14] Crank, J., and Nicolson, P., 1996. "A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type". *Advances in Computational Mathematics*, **6**(1), pp. 207–226.
- [15] Shewchuk, J. R., 1994. An introduction to the conjugate gradient method without the agonizing pain. Technical report CMU-CS-94-125, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, August.
- [16] Armijo, L., 1966. "Minimization of functions having lipschitz continuous first partial derivatives". *Pacific Journal of Mathematics*, **16**(1), pp. 1–3.