Comput Visual Sci (2010) 13:153–160
DOI 10.1007/s00791-010-0134-4

REGULAR ARTICLE

Multigrid and sparse-grid schemes for elliptic control problems with random coefficients

A. Borzì

Received: 20 January 2009 / Accepted: 26 March 2010 / Published online: 23 April 2010

Abstract A multigrid and sparse-grid computational approach to solving nonlinear elliptic optimal control problems with random coefficients is presented. The proposed scheme combines multigrid methods with sparse-grids collocation techniques. Within this framework the influence of randomness of problem's coefficients on the control provided by the optimal control theory is investigated. Numerical results of computation of stochastic optimal control solutions and formulation of mean control functions are presented.

Keywords Multigrid method · Sparse grids · Reaction-diffusion problems · Random fields · Optimal control theory

Mathematics Subject Classification (2000) 35J55 · 60H25 · 49K20 · 65N55 · 65C20

1 Introduction

An established tool for the construction of control strategies for real systems is provided by optimal control theory [17,19] where an optimal control problem is formulated as the minimization of an objective, that models the purpose of the control and describes the cost of its action, under the constraint given by the modeling equations. We focus on the control of a semilinear elliptic PDE with random coefficients where the nonlinearity is nonmonotonic. Therefore the resulting optimal control problem is singular [20] in the sense that without control the model equation my not admit solutions.

With this benchmark, we investigate an elliptic optimal control problem with a random coefficient that describes the reaction nonlinearity. The present investigation, is based on previous work on deterministic models [6,7] and on recent research [22,31] on modeling and optimization problems where the coefficients of the problem are described by random fields. In this paper, we present an efficient multigrid and sparse-grids methodology to solve these problems. Our setting is based on the work in [1,12,23,29] on elliptic problems with random inputs. We assume that the reaction coefficient is modeled by random fields that can be approximated by a truncated Karhunen–Loève expansion on the probability space. With this representation, we can use the Smolyak sparse-grid algorithm [10,18] to model a high-dimensional stochastic coefficient space.

We use a stochastic collocation method, where the solution of the stochastic optimal control problem is obtained solving, in the physical space, a deterministic optimality system for each point of the sparse-grids coefficient space. For the solution of the deterministic optimality system, we use a collective-smoothing multigrid (CSMG) scheme presented in [3–5] that provides optimal computational performance independently of the values of the optimization parameters and of the problem's coefficient [6]. The combination of sparse-grids and multigrid techniques results in a solution process with optimal computational complexity with respect to the sizes of the physical and probability grids. In a companion

Supported in part by the Austrian Science Fund FWF project P18136-N13 "Quantum optimal control of semiconductor nanostructures" and F3205-N18 "Fast Multigrid Methods for Inverse Problems".

A. Borzì (✉)
Università degli Studi del Sannio, Dipartimento e Facoltà di Ingegneria, Palazzo Dell'Aquila Bosco Lucarelli, Corso Garibaldi 107, 82100 Benevento, Italy
e-mail: alfio.borzi@unisannio.it

A. Borzì
Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstr. 36, 8010 Graz, Austria
e-mail: alfio.borzi@uni-graz.at

paper [8], the methodology presented here is extended to the case of parabolic control problems.

In the next section, we discuss the modeling of random fields and their representation by the Karhunen–Loève expansion and define the concept of solution of a stochastic PDE problem. In Sect. 3, we formulate a representative nonlinear elliptic optimal control problem with a random coefficient and consider a deterministic objective. In Sect. 4, we discuss the discretization of the stochastic parameter space using sparse-grids collocation and describe the Smolyak scheme. In Sect. 5, a collective-smoothing multigrid scheme for nonlinear elliptic optimal control problems is discussed. Section 6 is dedicated to numerical experiments for validating the numerical performance of the CSMG multigrid scheme combined with sparse-grids collocation techniques. Typical multigrid convergence rates and robustness with respect to a large choice of optimization parameters is obtained. Results of computation of stochastic optimal control solutions are reported with a focus on the moments of the tracking ability of the optimization scheme. We investigate the construction of a robust control obtained as the mean of the controls resulting from different realizations of reaction fields and demonstrate that this control represents an improvement compared with the control obtained considering a mean field reaction coefficient. A conclusion section completes the exposition of our work.

2 Random coefficients

In the modeling of application systems, material properties are only known at a statistical level. On the other hand, observation (or preparation) of physical configurations is subject to random noise. Both sources of randomness result in uncertainties in model and process conditions that can be accommodated in the framework described below.

Consider the following nonlinear elliptic equation

$$\begin{cases} \Delta y + G_\delta(y) = f & \text{in } \Omega \\ y = g & \text{on } \partial\Omega, \end{cases} \tag{1}$$

where $\Omega \subset \mathbb{R}^d$, $d \geq 1$ is the spatial dimension. The nonlinear term $G_\delta(y)$ models the reaction kinetics for the state y; δ is a reaction coefficient that is considered subject to randomness. We assume that the coefficient δ is described by a spatial random field. That is, we assume that $\delta = \delta(x, \omega)$ where $x \in \Omega$ and $\omega \in \mathcal{O}$, where the triple $(\mathcal{O}, \mathbf{A}, \mathcal{P})$ denotes a probability space where $\mathcal{O}$ is the space of elementary events, $\mathbf{A}$ is the sigma-algebra of subsets of $\mathcal{O}$, and $\mathcal{P}$ the probability measure on $\mathcal{O}$. The values of the stochastic function $\delta(x, \omega)$ are usually spatially correlated in a way characterized by a covariance structure. Clearly, we cannot model numerically the resulting infinite-dimensional coefficient space and a suitable finite-dimensional approximation must be introduced. For this purpose, a common approach is to use the Karhunen–Loève (KL) expansion [21] of the random field $\delta(x, \omega)$, that is based on a spectral decomposition of the covariance kernel of the stochastic process. We assume that the mean and the covariance of $\delta(x, \omega)$ are known respectively as

$$\delta_0(x) = \mathbb{E}(\delta)(x) := \int_{\mathcal{O}} \delta(x, \omega) d\mathcal{P}(\omega) \tag{2}$$

and

$$C_\delta(x, x') = \int_{\mathcal{O}} (\delta(x, \omega) - \delta_0(x))(\delta(x', \omega) - \delta_0(x')) d\mathcal{P}(\omega). \tag{3}$$

We see that C_δ defines the kernel of a compact, positive, and self-adjoint operator. Denote with λ_j the real positive eigenvalues and with $z_j(x)$ the corresponding orthonormal eigenfunctions of C_δ as follows

$$\int_\Omega C_\delta(x, x') z_j(x') dx' = \lambda_j z_j(x), \quad x \in \Omega,$$

where we assume to order the eigenvalues in decreasing order. Having the eigenpairs, we can define the following uncorrelated random variables

$$Y_j(\omega) = \frac{1}{\sqrt{\lambda_j}} \int_\Omega (\delta(x, \omega) - \delta_0(x)) z_j(x) dx, \quad j=1, 2, \ldots$$

with zero mean and unit variance, i.e. $\int_{\mathcal{O}} Y_i(\omega) Y_j(\omega) d\mathcal{P}(\omega) = \delta_{ij}$. We assume that the random variables Y_j are independent.

Now, the truncated KL expansion of $\delta(x, \omega)$ is given by

$$\delta_N(x, \omega) = \delta_0(x) + \sum_{j=1}^{N} \sqrt{\lambda_j} z_j(x) Y_j(\omega), \tag{4}$$

where N denotes the number of terms in the truncation. We see that $\delta_N(x, \omega)$ may provide a suitable approximation to $\delta(x, \omega)$ assuming that the eigenvalues decay sufficiently fast and N is sufficiently large.

We consider random fields characterized by squared exponential covariance

$$C_\delta(x, x') = s^2 \exp\left(-\frac{|x - x'|^2}{\ell^2}\right), \qquad x, x' \in [a, b],$$

with mean δ_0 and variance s^2. The degree of variability of this random field can be characterized by the ratio s/δ_0, while the frequency of variation of this field is related to the ratio L/ℓ, where L is the characteristic length of the physical domain and ℓ represents the physical correlation length. One can show that the eigenvalues have exponential decay as $\lambda \sim c_1 \ell \exp(-c_2 \ell^2)$ for some positive constants c_1 and c_2; see, e.g., [16,23].

We consider random fields of the type given above that can be approximated by (4) with good accuracy for moderate values of N. Therefore, assuming finite dimensional random fields, we can write the explicit dependence of δ on the random variables $[Y_1, \ldots, Y_N]$ as follows

$$\delta(x, \omega) \approx \delta(x, Y_1, \ldots, Y_N).$$

With this setting, we have that the solution to (1) can be described in terms of spatial coordinates and the set of random variables as follows

$$y(x, \omega) \approx y(x, Y_1, \ldots, Y_N).$$

We see that solving a stochastic PDE problem requires to solve for the set of all deterministic solutions corresponding to all possible (infinite) events. An approximation to the stochastic solution can be obtained considering a subset of events in the probabilistic space.

3 An optimal control problem with random coefficients

We consider a reaction diffusion processes controlled through source terms with the purpose of attaining a desired target profile given by $y_d(x)$. We assume that the target is deterministic and assume that the random parameter field is sufficiently regular such that moments of the solution up to a given degree exist; see [26]. We denote with $Y = [Y_1, \ldots, Y_N]$ and omit the fact that all dependent variables representing the solution to the optimal control problem are functions of the event ω. Thus, for the state we write $y(x)$ instead of $y(x, \omega)$, etc..

In the following, we formulate a nonlinear distributed optimal control problem with a random coefficient

$$\begin{cases} \min_{u \in L^2(\Omega;\mathcal{O})} J(y, u) \\ \Delta y + G_{\delta(Y(\omega))}(y) = f + u & \text{in } \Omega \times \mathcal{O}, \\ y = g & \text{on } \partial\Omega, \end{cases} \tag{5}$$

where we assume deterministic boundary conditions and forcing term. A representative singular control problem is obtained choosing

$$G_\delta(y) = \delta\, e^y,$$

which models chemically generated heat in a solid fuel ignition model [2], where Ω represents a solid fuel, y its temperature distribution, and Δy models heat transfer due to conduction. If Ω is a unit square and $f = 0$, then there exists a critical value $\delta^* \approx 6.808$ such that there is no solution to the governing equation if $\delta > \delta^*$; see [2,6]. However, in the presence of control, by adding or subtracting thermal energy to the fuel through a source term, we can drive the system to a given target configuration.

To attain a desired target given by $y_d(x) \in L^2(\Omega)$, we introduce the following deterministic cost functional

$$J(y, u) = \frac{1}{2}||y - y_d||^2_{L^2(\Omega)} + \frac{\nu}{2}||u||^2_{L^2(\Omega)}. \tag{6}$$

Here, $\nu > 0$ is the weight of the cost of the control, i.e. the optimization parameter.

The optimal control problem (5)–(6) is stochastic in the sense that for any event ω a different control results. Each single event corresponds to the solution of a deterministic optimal control problem; see, e.g., [6]. For ω, we have the occurrence $\delta = \delta(x, \omega)$. With this specification and choosing $f = 0$ and $g = 0$, we have that the solution to (5)–(6) is characterized by the following first-order optimality system

$$\begin{array}{ll} \Delta y + \delta\, e^y = u & \text{in } \Omega, \\ \Delta p + \delta\, e^y\, p + (y - y_d) = 0 & \text{in } \Omega, \\ \nu u - p = 0 & \text{in } \Omega, \\ y = 0, \quad p = 0 & \text{on } \partial\Omega. \end{array} \tag{7}$$

At this point, we can introduce the reduced cost functional $\hat{J}$ given by

$$\hat{J}(u) = J(y(u), u), \tag{8}$$

where $y(u)$ denotes the solution to the state equation for a given u. Notice that for $\delta > \delta^*$, the governing equation may have multiple solutions. We assume that $y(u)$ represents the solution of the lower branch [6]. The gradient of $\hat{J}$ with respect to u is given by $\nabla \hat{J}(u) = \nu u - p(u)$ where $p(u)$ is the solution of the adjoint equation for the given $y(u)$.

The solution of optimal control problems with random coefficients requires to compute a control for each coefficient configuration. The space of these controls represents the solution of the stochastic control problem. However, an application task is to deliver a unique control for the governing random PDE model. This control should be at least sub-optimal for all configurations of the coefficient and ideally should be such to minimize the mean of the objective while spanning the stochastic coefficient space. That is, one could consider the following

$$\min_{u \in L^2(\Omega)} \mathbb{E}(\hat{J})(u) = \min_{u \in L^2(\Omega)} \int_{\mathcal{O}} \hat{J}_\omega(u)\, d\mathcal{P}(\omega) \tag{9}$$

where $\hat{J}_\omega(u) = \hat{J}(y_\omega(u), u)$. This formulation requires to solve an optimization problem with infinitely many terms in the objective (the integral in $\mathcal{O}$) and infinitely many PDE-constraints (that characterize $y_\omega(u)$).

In order to define a reasonable approximation to this problem, it is important to consider the following fact discussed in [24,25]. Given a functional F depending on a random variable δ with mean $\mathbb{E}(\delta)$ and standard deviations s_δ we have the following second-order Taylor expansion

$$\mathbb{E}(F(\delta)) = F(\mathbb{E}(\delta)) + \frac{1}{2}\frac{\partial^2 F}{\partial \delta^2}\, s_\delta^2 + O(s^4)$$

where the second derivative is evaluated at $\mathbb{E}(\delta)$. From this formula we conclude that a first-order approximation to (9) corresponds to $\min_{u\in L^2(\Omega)} \hat{J}(u)$ where the state equation has mean field $\mathbb{E}(\delta)$ as reaction coefficient.

Notice that the state equation is nonlinear and therefore the control obtained using the mean reaction coefficient, $u_{\mathbb{E}(\delta)}$, is not the same as the mean of the controls, $\mathbb{E}(u)$, obtained by averaging over all δ configurations. Therefore, while the Taylor series expansion above shows that $u_{\mathbb{E}(\delta)}$ is a first-order approximate solution to (9), it appears reasonable to ask whether or not $\tilde{u} = \mathbb{E}(u)$ can be also an approximate minimizer. This is not clear since in general we expect

$$J(y(u_{\mathbb{E}(\delta)}), u_{\mathbb{E}(\delta)}) \neq J(y(\mathbb{E}(u)), \mathbb{E}(u))$$

In the section of numerical experiments we compare these two approximation strategies in the effort of designing a robust optimal control. It turns out that $\tilde{u}$ represents an improvement upon $u_{\mathbb{E}(\delta)}$.

4 Discretization of the probabilistic space

We recall that the solution of a stochastic PDE problem is the set of all deterministic solutions corresponding to all possible events. To approximate to the set of all deterministic optimization solutions one can choose to represent the stochastic PDE problem using the so called generalized polynomial chaos expansion (GPCE) [11,30]. But this approach is difficult to implement and requires to solve very large-sized coupled systems.

A viable approach is by collocation on a grid in the probability space and the physical domain is approximated using a classical discretization scheme. On this grid we need to solve (5)–(6) for each grid point and all statistical quantities related to these solutions are obtained by integration on this grid. However, a cartesian grid of a N-dimensional space will require a number of function evaluations that grows exponentially with N. This problem can be alleviated in part using a set of nodes with improved interpolation properties (low Lebesgue number) as the Chebyshev nodes given by

$$z_i^j = -\cos\frac{\pi(i-1)}{m_j - 1}, i = 1, \ldots, m_j,$$

where $m_j = 2^{j-1} + 1$ ($m_1 = 1$) denotes the number of nodes of a j-order (or level) grid. One notices that the set of Chebyshev nodes for all j is nested. Let $\{z^j\} = \{z_1^j, \ldots, z_{m_j}^j\}$ then the nesting property gives that $\{z^j\} \subset \{z^{j+1}\}$. Correspondingly, let $\{w^j\}$ be the set the quadrature weights in one dimension which correspond to the nodes $\{z^j\}$. Then, on a tensor-product grid we have that the mean value of a function $f : \mathbb{R}^N \to \mathbb{R}$ is given by

$$\mathbb{E}(f) = \sum_{i_1=1}^{m_1} \cdots \sum_{i_N=1}^{m_N} f(z_{i_1}^{j_1}, \ldots, z_{i_N}^{j_N})\,(w^{j_1} \otimes \cdots \otimes w^{j_N}).$$

To reduce the $m_1 \times \cdots \times m_N$-number of function evaluations, we use a Smolyak scheme [12,23,29] that constructs interpolation as linear combination of tensor-product formulas on a minimal number of nodes of the multidimensional space. A full N-dimensional grid of order $\mathcal{J}$, with $j_1 = j_2 = \cdots j_N = \mathcal{J}$ in each dimension, is formed using a tensor product of the constituent one-dimensional grids and has total order $j_1 + j_2 + \cdots + j_N = N\,\mathcal{J}$. On the other hand, the sparse grid of order $\mathcal{J}$ is composed of a strict subset of full grids where $\mathcal{J}$ is the order of the largest allowed grid and the orders of the constituent grids add up to a total order given by $\mathcal{J} + N - 1$. For example, consider $N = 4$ and a sparse grid of order $\mathcal{J} = 2$. This grid results in $\{z^1 \otimes z^4\} \bigcup \{z^2 \otimes z^3\} \bigcup \{z^3 \otimes z^2\} \bigcup \{z^4 \otimes z^1\}$.

To define the weights for sparse-grid integration, we need to define the difference weights of order j, which are $\eta^j = w^j - w^{j-1}$ on the odd numbered nodes. On the sparse grid of dimension N and total order $\mathcal{J}$, the weights are given by

$$w^{\mathcal{J}} = \sum_{\ell=1}^{N+\mathcal{J}-1} \sum_{|\mathbf{j}|=\ell} (\eta^{j_1} \otimes \cdots \otimes \eta^{j_N}).$$

For more details on sparse grids see [10,18].

5 A collective smoothing multigrid scheme

An advantage of the stochastic collocation approach on sparse grids with respect to, e.g., Monte Carlo simulation, is to greatly reduce the number of solver calls. Nevertheless, the number of optimality system solves remains considerably large thus requiring an efficient and robust solution strategy for each coefficient configuration. For this purpose we use the collective-smoothing multigrid strategy [6,7] that has been proven to solve (5)–(6) with textbook multigrid efficiency independently of the choice of optimization parameter's values.

To illustrate the collective-smoothing multigrid scheme, consider L grid levels indexed by $k = 1, \ldots, L$, where $k = L$ refers to the finest grid. The mesh of level k is denoted by Ω_k where $h_k = h_1/2^{k-1}$. Any operator and variable defined on the discrete space Ω_k is indexed by k. The discretized optimality system at level k with homogeneous Dirichlet boundary conditions is given by

$$\begin{aligned} &\Delta_h y_h + \delta \exp(y_h) = f_h + u_h, \\ &\Delta_h p_h + \delta \exp(y_h)\, p_h + (y_h - y_{dh}) = 0, \\ &\nu u_h - p_h = 0, \end{aligned} \tag{10}$$

where Δ_h represents the five-point finite differences Laplacian with homogeneous Dirichlet boundary conditions. We denote this nonlinear algebraic problem as the following

$$A_k(w_k) = f_k, \quad w_k = (y_k, u_k, p_k). \tag{11}$$

As well known [9,15,28], the multigrid strategy combines two complementary schemes. The high-frequency components of the solution error are reduced by a smoothing iteration, denoted by S_k, while the low-frequency error components are effectively reduced by a coarse-grid correction method as defined below.

The action of one multigrid cycle applied to (11) can be expressed in terms of a (nonlinear) multigrid iteration operator B_k. Starting with an initial approximation $w_k^{(0)}$ the result of one $V(\nu_1, \nu_2)$-cycle is then denoted by $w_k = B_k(w_k^{(0)})\, f_k$.

Algorithm 1 *Multigrid $V(\nu_1, \nu_2)$-Cycle*

Set $B_1(w_1^{(0)}) \approx A_1^{-1}$ (e.g., iterating with S_1 starting with $w_1^{(0)}$). For $k = 2, \ldots, L$ define B_k in terms of B_{k-1} as follows.

1. *Set the starting approximation $w_k^{(0)}$.*
2. *Pre-smoothing. Define $w_k^{(l)}$ for $l = 1, \ldots, \nu_1$, by*

$$w_k^{(l)} = S_k(w_k^{(l-1)}, f_k).$$

3. *Coarse-grid correction. Set $w_k^{(\nu_1+1)} = w_k^{(\nu_1)} + I_{k-1}^k (w_{k-1} - \hat{I}_k^{k-1} w_k^{(\nu_1)})$ where*

$$w_{k-1} = B_{k-1}\left(\hat{I}_k^{k-1} w_k^{(\nu_1)}\right)\left[I_k^{k-1}\left(f_k - A_k\left(w_k^{(\nu_1)}\right)\right) + A_{k-1}\left(\hat{I}_k^{k-1} w_k^{(\nu_1)}\right)\right].$$

4. *Post-smoothing. Define $w_k^{(l)}$ for $l = \nu_1 + 2, \ldots, \nu_1 + \nu_2 + 1$, by*

$$w_k^{(l)} = S_k\left(w_k^{(l-1)}, f_k\right).$$

5. *Set $B_k(w_k^{(0)})\, f_k = w_k^{(\nu_1+\nu_2+1)}$.*

We choose I_k^{k-1} to be the full-weighted restriction operator [15,28]. The prolongation I_{k-1}^k is defined by bilinear interpolation. We choose $\hat{I}_k^{k-1}$ to be straight injection.

We remark that a key component in the development of multigrid solvers for optimality systems is the design of robust collective smoothing schemes [7]. In the case of singular optimal control problems the construction of a CSMG scheme is presented in [6]. To illustrate the construction of this scheme in the present case, let $\mathbf{x} \in \Omega_k$, where $\mathbf{x} = (ih, jh)$ and i, j index the grid points, e.g., lexicographically, and define the stencil-set $\omega_{ij} = \{0, 0; 1, 0; -1, 0; 0, 1; 0, -1\}$. We can express the action of Δ_h on the function v_h in i, j in the following compact form

$$\Delta_h v_h|_{ij} = \frac{1}{h^2}\left(\sum_{s,t\in\omega_{ij},\, s,t\neq i,j} v_{st} - 4v_{ij}\right).$$

define

$$A_{ij} = \sum_{s,t\in\omega_{ij},\, s,t\neq i,j} y_{st} - h^2 f_{ij} \quad \text{and} \quad B_{ij} = \sum_{s,t\in\omega_{ij},\, s,t\neq i,j} p_{st} - h^2 y_{dhij}.$$

Here A_{ij} and B_{ij} are considered constant during the update of the variables at i, j. With this setting the optimality system at i, j becomes

$$A_{ij} - 4\, y_{ij} + h^2\, e^{y_{ij}} - h^2\, p_{ij}/\nu = 0, \tag{12}$$

$$B_{ij} - 4\, p_{ij} + h^2\, e^{y_{ij}}\, p_{ij} + h^2\, y_{ij} = 0, \tag{13}$$

where we replaced $u_{ij} = p_{ij}/\nu$.

To approximately solve this system we apply a local Newton update as in [6]. Consider (12)–(13). The inverse of the Jacobian of these two equations is given by

$$M_{ij}^{-1} = \frac{1}{\det M_{ij}} \begin{pmatrix} -4 + h^2\, e^{y_{ij}} & h^2/\nu \\ -h^2\,(1 + e^{y_{ij}}\, p_{ij}) & -4 + h^2\, e^{y_{ij}} \end{pmatrix}, \tag{14}$$

where $\det M_{ij} = (-4 + h^2\, e^{y_{ij}})(-4 + h^2\, e^{y_{ij}})$. Notice that h can be chosen sufficiently small to guarantee that $\det M_{ij} \neq 0$.

We have the following smoothing scheme

Algorithm 2 *Collective Gauss-Seidel Iteration*

1. *Set the starting approximation.*
2. *For ij in, e.g., lexicographic order do*

$$\begin{pmatrix} y \\ p \end{pmatrix}_{ij}^{(1)} = \begin{pmatrix} y \\ p \end{pmatrix}_{ij}^{(0)} + M^{-1} \begin{pmatrix} r_y \\ r_p \end{pmatrix}_{ij};$$

3. *end.*

where $r_{ij}^y = -(A_{ij} - 4\, y_{ij} + h^2\, e^{y_{ij}} - h^2\, p_{ij}/\nu)$ and $r_{ij}^p = -(B_{ij} - 4\, p_{ij} + h^2\, e^{y_{ij}}\, p_{ij} + h^2\, y_{ij})$, denote the residuals of state and adjoint equations at i, j prior to the update.

6 Numerical experiments

We validate the collective-smoothing multigrid scheme and sparse-grids collocation techniques with our elliptic optimal control problem with random coefficient benchmark. We investigate two robust control strategies. The first one is obtained by the mean of the controls resulting from different realizations of the reaction coefficient field. The other

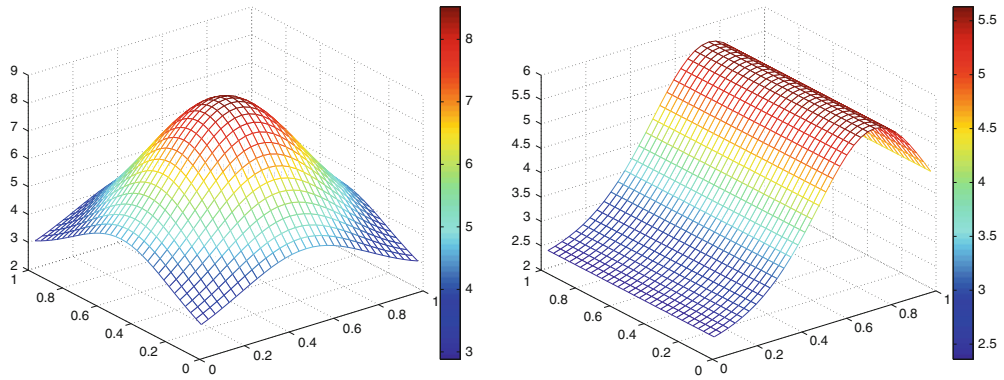

Fig. 1 Two realization of $\delta(x_1, x_2, \omega)$. *Left* $Y_1 = 0.0$, $Y_2 = 0.0$, $Y_3 = 1.0$, $Y_4 = 1.0$. *Right* $Y_1 = 0.0$, $Y_2 = -1.0$, $Y_3 = 0.0$, $Y_4 = 1.0$

control is obtained by using a reaction diffusion model with an averaged reaction field.

We consider a random reaction coefficient given by [23]

$$\begin{aligned}\delta(x_1, x_2, \omega) = \delta_0 + \exp\{&[Y_1(\omega)\cos(\pi x_2) \\ &+ Y_3(\omega)\sin(\pi x_2)]\, e^{-1/8} + [Y_2(\omega)\cos(\pi x_1) \\ &+ Y_4(\omega)\sin(\pi x_1)]\, e^{-1/8}\}\end{aligned}$$

where $\delta_0 = 2$ and $Y_j \in [-1, 1]$, $j = 1, 2, 3, 4$. This field is characterized by a squared exponential covariance typical of Gaussian processes. Two realization of $\delta(x_1, x_2, \omega)$ are depicted in Fig. 1 where we see that there are realizations of the reaction coefficient function such that $\delta(x_1, x_2) > \delta^*$ in a large part of the computational domain. On the other hand, we can see in Fig. 2 that the mean field coefficient is bounded $\mathbb{E}(\delta) \leq 3.84$. Because source terms enter linearly into the problem, we do not expect large effects on the optimization solution due to their randomness and we take $f = 0$.

The desired target configuration is given by

$$y_d(x_1, x_2) = 1 - \sin(\pi\, x_1)\sin(\pi\, x_2).$$

Because of homogeneous Dirichlet boundary conditions, this target is not attainable by any control.

We use two pre- and one post-smoothing steps ($\nu_1 = \nu_2 = 2$)
and $h = 1/4$ is the coarsest space mesh size. We take $\Omega = (0, 1) \times (0, 1)$. Two different grids $N_x \times N_y$ are considered: 64×64 and 128×128. For the discretization of the stochastic space we use a four-dimensional sparse grid and present results with order $\mathcal{J} = 2$ and $\mathcal{J} = 3$. These correspond to a total number of sparse-grid points of $\chi = 41$ and $\chi = 137$, which appear to be a sufficiently large number to capture the first few moments; compare with [31].

To describe the results of the experiments we report mean values of the observed multigrid convergence factors ρ_{obs}, which are defined as the asymptotic value of the ratio of the norm of the dynamic residuals given by $||r_y|| + ||r_p||/\nu$ resulting from two successive multigrid cycles. In all experiments, we iterate the multigrid algorithm to a specified stopping criteria given by $||r_y|| + ||r_p||/\nu < 10^{-10}$.

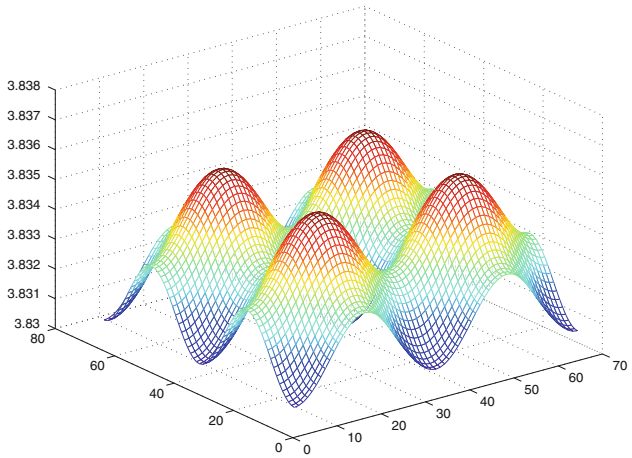

Fig. 2 The mean field reaction coefficient $\mathbb{E}(\delta) \approx 3.83$

The tracking ability of the CSMG algorithm will be expressed in terms of the mean values of the norms of the tracking error $\mathbb{E}(\|y - y_d\|)$. Other important statistical observable are the variance of the tracking error $Var\,(\|y - y_d\|) = \mathbb{E}((\|y - y_d\| - \mathbb{E}(\|y - y_d\|))^2)$, and its skewness $Skew\,(\|y - y_d\|) = \mathbb{E}((\|y - y_d\| - \mathbb{E}(\|y - y_d\|))^3)$.

Next, we report results of experiments. The results reported in Table 1 illustrate efficiency and robustness of the proposed multigrid solvers. These results demonstrate usual multigrid convergence speeds which appear to be independent of the value of ν.

In Table 1, we see that CPU times scale linearly with the number of spatial grid points. We also see that CPU times scale linearly with the number of evaluation points of the stochastic sparse grid. Altogether, we obtain optimal computational complexity in CPU time. We find less improvement of the mean and of the other moments of the objective when we refine the sparse-grid mesh.

Table 1 Results with the CSMG sparse-grids scheme: Denote $\|\Delta y\| = \|y - y_d\|$

ν	$N_x \times N_y$	CPU(s)	$\mathbb{E}(\|\Delta y\|)$	$Var\ (\|\Delta y\|)$	$Skew\ (\|\Delta y\|)$	$\mathbb{E}(\rho_{obs})$
$\mathcal{J} = 2, \chi = 41$						
10^{-3}	64×64	10.5	5.22×10^{-1}	3.85×10^{-6}	-4.85×10^{-9}	0.07
10^{-6}	64×64	11.5	1.97×10^{-1}	7.90×10^{-9}	-6.76×10^{-13}	0.05
10^{-3}	128×128	42.8	5.36×10^{-1}	3.67×10^{-6}	-4.52×10^{-9}	0.07
10^{-6}	128×128	51.4	2.28×10^{-1}	6.50×10^{-9}	-5.05×10^{-13}	0.05
$\mathcal{J} = 3, \chi = 137$						
10^{-3}	64×64	35.5	5.22×10^{-1}	3.46×10^{-6}	-2.15×10^{-9}	0.07
10^{-6}	64×64	38.6	1.97×10^{-1}	7.15×10^{-9}	-5.36×10^{-13}	0.05
10^{-3}	128×128	145.4	5.36×10^{-1}	3.30×10^{-6}	-2.00×10^{-9}	0.07
10^{-6}	128×128	158.3	2.28×10^{-1}	5.88×10^{-9}	-4.00×10^{-6}	0.05

Table 2 Results with a given control (left) and moments of tracking ability for controls corresponding to different realization of the reaction coefficient (as in Table 1); $N_x \times N_y = 128 \times 128$, $\chi = 137$

	ν	$\mathbb{E}(\|y - y_d\|)$	$Var\ (\|y - y_d\|)$	$\mathbb{E}(\|\Delta y\|)$	$Var\ (\|\Delta y\|)$
$y_{\mathbb{E}(u)}$	10^{-4}	4.44×10^{-1}	2.91×10^{-3}	4.34×10^{-1}	1.60×10^{-6}
$y_{u_{\mathbb{E}(\delta)}}$	10^{-4}	5.97×10^{-1}	2.82×10^{-4}		
$y_{\mathbb{E}(u)}$	10^{-6}	2.42×10^{-1}	3.07×10^{-3}	2.28×10^{-1}	5.88×10^{-9}
$y_{u_{\mathbb{E}(\delta)}}$	10^{-6}	5.76×10^{-1}	3.04×10^{-4}		
$y_{\mathbb{E}(u)}$	10^{-8}	1.16×10^{-1}	4.36×10^{-3}	9.17×10^{-2}	2.75×10^{-11}
$y_{u_{\mathbb{E}(\delta)}}$	10^{-8}	5.67×10^{-1}	3.38×10^{-4}		

We now discuss the problem of defining a unique control that provides good tracking features for all configurations of the reaction coefficient. As a criteria to evaluate the confidence level for this control, for any given value of the control weight ν, we compare with the mean and other moments of the tracking functional resulting from the optimal controls specific for each δ configuration as given in Table 1.

As possible control functions, we first consider the function $\tilde{u} = \mathbb{E}(u)$, i.e. the mean function of the optimal controls corresponding to each point of the configuration space. Another control function is obtained considering a mean field coefficient and computing the corresponding control. With these candidate control functions, we solve the forward problem corresponding to each configuration of the reaction random coefficient. Thus, we obtain $y_{\mathbb{E}(u)}$ and $y_{\mathbb{E}(\delta)}$ and evaluate the corresponding tracking properties. Notice that for some δ the operator equation is nonmonotone and solving for y with a given u is a difficult task. However, we can compute $y(u)$ using $\mathbb{E}(y)$ as initial approximation for a nonlinear multigrid iteration for the forward problem.

To compare the ability of $y_{\mathbb{E}(u)}$ and $y_{\mathbb{E}(\delta)}$ to track the desired target for each different coefficient configuration, we consider different control weights. The results of these experiments are given in Table 2. We see that the controlled state $y_{\mathbb{E}(\delta)}$ obtained using a mean field coefficient (see Fig. 2) has a limited tracking ability which results less sensitive to the value of ν. On the other hand, the controlled state $y_{\mathbb{E}(u)}$ corresponding to a mean control function provides better tracking performance than $y_{\mathbb{E}(\delta)}$ and this performance improves as we consider smaller weights. In addition, we have the desirable property that the mean tracking obtained with $y_{\mathbb{E}(u)}$ is close to the mean tracking error obtained considering different realization of the reaction coefficient.

7 Conclusions

A singular elliptic optimal control problem with a random reaction coefficient was formulated and a collective-smoothing multigrid method combined with sparse-grids collocation techniques was implemented to robustly and efficiently solve these optimization problems. Results of computation of stochastic optimal control solutions were reported with a focus on the moments of the tracking ability of the optimization schemes. The construction of control strategies was discussed and it was found that the mean control function is robust and accurate for all configuration of the reaction coefficient.

Acknowledgments I would like to gratefully acknowledge the help of Greg von Winckel and many fruitful discussion with Volker Schulz and Claudia Schillings.

References

1. Babuška, I., Nobile, F., Tempone, R.: A stochastic collocation method for elliptic partial differential equations with random input data. SIAM J. Numer. Anal. **45**(3), 1005–1034 (2007)
2. Bebernes, J., Eberly, D.: Mathematical Problems from Combustion Theory. Springer, New York (1988)
3. Borzì, A.: High-order discretization and multigrid solution of elliptic nonlinear constrained optimal control problems. J. Comput. Appl. Math. **200**, 67–85 (2007)
4. Borzì, A.: Smoothers for control- and state-constrained optimal control problems. Comput. Vis. Sci. **11**, 59–66 (2008)
5. Borzì, A., Borzì, G.: An efficient algebraic multigrid method for solving optimality systems. Comput. Vis. Sci. **7**(3/4), 183–188 (2004)
6. Borzì, A., Kunisch, K.: The numerical solution of the steady state solid fuel ignition model and its optimal control. SIAM J. Sci. Comput. **22**(1), 263–284 (2000)
7. Borzì, A., Schulz, V.: Multigrid methods for PDE optimization. SIAM Rev. **51**, 361–395 (2009)
8. Borzì, A., von Winckel, G.: Multigrid methods and sparse-grid collocation techniques for parabolic optimal control problems with random coefficients. SIAM J. Sci. Comput. **31**, 2172–2192 (2009)
9. Brandt, A.: Multi-level adaptive solutions to boundary-value problems. Math. Comput. **31**, 333–390 (1977)
10. Bungartz, H.-J., Griebel, M.: Sparse grids. Acta Numerica **13**, 147–269 (2004)
11. Elman, H., Furnival, D.: Solving the stochastic steady-state diffusion problem using multigrid. IMA J. Numer. Anal. **27**, 675–688 (2007)
12. Ganapathysubramanian, B., Zabaras, N.: Sparse grid collocation schemes for stochastic natural convection problems. J. Comput. Phys. **225**, 652–685 (2007)
13. Gasca, M., Sauer, T.: Polynomial interpolation in several variables. Adv. Comput. Math. **12**, 377–410 (2000)
14. Ghanem, R.G., Spanos, P.D.: Stochastic Finite Elements: A Spectral Approach. Springer, New York (1991)
15. Hackbusch, W.: Multi-grid Methods and Applications. Springer, New York (1985)
16. Huang, S.P., Mahadevan, S., Rebba, R.: Collocation-based stochastic finite element analysis for random field problems. Probabilistic Eng. Mech. **22**, 194–205 (2007)
17. Ito, K., Kunisch, K.: Lagrange Multiplier Approach to Variational Problems and Applications. SIAM, Philadelphia (2008)
18. Klimke, A., Wohlmuth, B.: Algorithm 847: Spinterp: piecewise multilinear hierarchical sparse grid interpolation in MATLAB. ACM Trans. Math. Softw. **31**, 561–579 (2005)
19. Lions, J.L.: Optimal Control of Systems Governed by Partial Differential Equations. Springer, Berlin (1971)
20. Lions, J.-L.: Control of Distributed Singular Systems. Gauthier-Villars, Paris (1985)
21. Loeve, M.: Probability Theory. Vols. I & II, IV edn. Springer, New York (1978)
22. Marzouk, Y.M., Najm, H.N., Rahn, L.A.: Stochastic spectral methods for efficient Bayesian solution of inverse problems. J. Comput. Phys. **224**, 560–586 (2007)
23. Nobile, F., Tempone, R., Webster, C.: A sparse grid stochastic collocation method for elliptic partial differential equations with random input data. SIAM J. Numer. Anal. **46**, 2309–2345 (2008)
24. Putko, M.M., Newman, P.A., Taylor, A.C. III., Green, L.L.: Approach for uncertainty propagation and robust design in CFD using sensitivity derivatives. J. Fluids Eng. **124**, 60–69 (2002)
25. Schulz, V., Schillings, C.: On the nature and treatment of uncertainties in aerodynamic design. AIAA J. **47**, 646–654 (2009)
26. Schwab, Ch, Todor, R.A.: Sparse finite elements for stochastic elliptic problems higher order moments. Computing **71**, 43–63 (2003)
27. Seynaeve, B., Rosseel, E., Nicolaï, B., Vandewalle, S.: Fourier mode analysis of multigrid methods for partial differential equations with random coefficients. J. Comput. Phys. **224**, 132–149 (2007)
28. Trottenberg, U., Oosterlee, C., Schüller, A.: Multigrid. Academic Press, London (2001)
29. Xiu, D., Hesthaven, J.S.: High-order collocation methods for differential equations with random inputs. SIAM J. Sci. Comput. **27**(3), 1118–1139 (2005)
30. Xiu, D., Karniadakis, G.E.: The Wiener-Askey polynomial chaos for stochastic differential equations. SIAM J. Sci. Comput. **24**, 619–644 (2002)
31. Zabaras, N., Ganapathysubramanian, B.: A scalable framework for the solution of stochastic inverse problems using a sparse grid collocation approach. J. Comput. Phys. **227**, 4697–4735 (2008)