University of Maryland Department of Computer Science TR-5011
University of Maryland Institute for Advanced Computer Studies TR-2012-09
June 2012

# REDUCED BASIS COLLOCATION METHODS FOR PARTIAL DIFFERENTIAL EQUATIONS WITH RANDOM COEFFICIENTS* 

HOWARD C. ELMAN ${ }^{\dagger}$ AND QIFENG LIAO ${ }^{\ddagger}$


#### Abstract

The sparse grid stochastic collocation method is a new method for solving partial differential equations with random coefficients. However, when the probability space has high dimensionality, the number of points required for accurate collocation solutions can be large, and it may be costly to construct the solution. We show that this process can be made more efficient by combining collocation with reduced-basis methods, in which a greedy algorithm is used to identify a reduced problem to which the collocation method can be applied. Because the reduced model is much smaller, costs are reduced significantly. We demonstrate with numerical experiments that this is achieved with essentially no loss of accuracy.


Key words. stochastic PDEs, collocation, reduced basis
AMS subject classifications. 35R60, 60H15, 65N22, 65N35

1. Introduction. Let $(\Omega, \Sigma, \mathcal{P})$ be a complete probability space, where $\Omega$ is the sample space, $\Sigma \in 2^{\Omega}$ the $\sigma$-algebra, and $\mathcal{P}: \Sigma \rightarrow[0,1]$ the probability measure. Let $D \subset \mathbb{R}^{d}(\mathrm{~d}=1,2,3)$ be a bounded and connected domain with a polygonal boundary $\partial D$. We consider the problem of finding a random function $u(\vec{x}, \omega) \in D \times \Omega \rightarrow \mathbb{R}$, such that $\mathcal{P}$-a.e. in $\Omega$,

$$
\begin{align*}
\mathcal{L}(\vec{x}, \omega ; u(\vec{x}, \omega))=f(\vec{x}) & \forall \vec{x} \in D,  \tag{1.1}\\
\mathfrak{b}(\vec{x}, \omega ; u(\vec{x}, \omega))=g(\vec{x}) & \forall \vec{x} \in \partial D, \tag{1.2}
\end{align*}
$$

where $\mathcal{L}$ is a partial differential operator and $\mathfrak{b}$ is a boundary operator, both of which can have random coefficients. In order to solve (1.1)-(1.2) numerically, the random coefficients in the operators should be represented by a finite number of random variables $\xi=\left[\xi_{1}(\omega), \ldots, \xi_{M}(\omega)\right]^{T}$. This could come from a variety of sources, for example, a truncated Karhunen-Loève (KL) expansion [2, 11], some partitioning of $D$ into subdomains, or uncertain boundary conditions. Letting $\Gamma_{i}:=\left[a_{i}, b_{i}\right]$ denote the image of $\xi_{i}(\omega)$ and $\Gamma:=\prod_{i=1}^{M} \Gamma_{i}$ the image of $\xi$, (1.1)-(1.2) can be rewritten as: find a function $u(\vec{x}, \xi) \in D \times \Gamma$, such that

$$
\begin{align*}
\mathcal{L}(\vec{x}, \xi ; u(\vec{x}, \xi))=f(\vec{x}) & \forall(\vec{x}, \xi) \in D \times \Gamma,  \tag{1.3}\\
\mathfrak{b}(\vec{x}, \xi ; u(\vec{x}, \xi))=g(\vec{x}) & \forall(\vec{x}, \xi) \in \partial D \times \Gamma, \tag{1.4}
\end{align*}
$$

where $\mathcal{L}$ and $\mathfrak{b}$ are assumed to be affinely dependent on $\xi$. It is of interest to identify moments and cumulative distributions associated with the solution $u(\vec{x}, \xi)$. Our aim

[^0]in this study is to develop a variant of the sparse grid collocation method for constructing solutions to (1.3)-(1.4), using reduced-basis methods to enhance efficiency.

The sparse grid collocation method [21, 32] for stochastic partial differential equations (PDEs) is an example of a spectral method, in which discrete solutions are constructed using polynomials in the random variable $\xi$ in combination with standard (e.g., finite element) spatial discretization. Collocation shares with Monte Carlo methods the feature that only solutions of a set of spatially discrete problems at a set of sample points $\left\{\xi^{(k)}\right\}$ are required, and it exhibits rapid convergence (nearly exponential in the number of sample points [21]), which makes it more efficient than Monte Carlo methods when the dimension $M$ of the sample space is modest, say, on the order of 100 or smaller $[18,19]$. However, the number of sample (collocation) points needed may still be large when accuracy requirements are strong, and if this is coupled with fine discretization for spatial accuracy, then the costs of stochastic collocation may be high. To address this issue, we combine collocation with reduced-basis methods $[4,13,20]$, in which parameterized discrete PDEs are projected into spaces of significantly smaller dimension. The reduced-basis methodology is designed to decrease the cost of simulation of parameter-dependent models; in this study, we show that this idea can be used with stochastic collocation to the same effect.

An outline of the rest of the paper is as follows. In Sections 2 and 3, we review the collocation and reduced basis methods, and in Section 4, we present our algorithm for combining them. In Sections 5 and 6, we demonstrate the efficiency of the combined method for solving stochastic versions of the diffusion equation and the incompressible Navier-Stokes equations. Finally, in Section 7 we make some concluding remarks.
2. Stochastic collocation on sparse grids. The main idea of stochastic collocation methods is to seek a numerical approximation to the exact solution of (1.3)-(1.4) in the form

$$
\begin{equation*}
u^{s c}(\vec{x}, \xi):=\sum_{\xi^{(k)} \in \Theta} u_{c}\left(\vec{x}, \xi^{(k)}\right) \tilde{L}_{\xi^{(k)}}(\xi) \tag{2.1}
\end{equation*}
$$

where $\Theta \subset \Gamma$ is a given sample set, $\left\{\tilde{L}_{\xi^{(k)}}(\xi)\right\}$ are some global interpolation polynomials defined in $\Gamma$ (e.g., Lagrange polynomials) and each coefficient function $u_{c}\left(\vec{x}, \xi^{(k)}\right)$ is the solution of a deterministic problem corresponding to a given realization $\xi^{(k)}$ of the random variable $\xi$,

$$
\begin{array}{rlrl}
\mathcal{L}\left(\vec{x}, \xi^{(k)} ; u\left(\vec{x}, \xi^{(k)}\right)\right) & =f(\vec{x}) & & \forall \vec{x} \in D \\
\mathfrak{b}\left(\vec{x}, \xi^{(k)} ; u\left(\vec{x}, \xi^{(k)}\right)\right) & =g(\vec{x}) & \forall \vec{x} \in \partial D \tag{2.3}
\end{array}
$$

We will be concerned with sparse grid collocation as described in Xiu and Hesthaven [32], which is based on the methodology of sparse grid interpolation. We begin with a brief review of this interpolation technique. Without loss of generality, the image of $\xi$ is assumed in this section to be $\Gamma^{*}:=[-1,1]^{M}$, since any finite $\Gamma=\prod_{i=1}^{M}\left[a_{i}, b_{i}\right]$ can be mapped to $\Gamma^{*}$. First, we consider a one-dimensional setting.

Introducing a level index $i \in \mathbb{N}$, let $\Theta_{1}^{i}=\left\{\xi_{j}^{i}, j=1: m_{i}\right\}$ be a partitioning of the interval $[-1,1]$, where $m_{i}$ is the number of partitioning points. For an arbitrary function $v(\xi) \in C([-1,1])$, its Lagrange interpolant is

$$
U^{i}(v)=\sum_{j=1}^{m_{i}} v\left(\xi_{j}^{i}\right) L_{j}^{i}(\xi)
$$

where

$$
L_{j}^{i}(\xi)=\prod_{k=1, k \neq j}^{m_{i}} \frac{\xi-\xi_{k}^{i}}{\xi_{k}^{j}-\xi_{k}^{i}} .
$$

A straightforward generalization for a function of $M$ variables $v(\xi) \in C\left(\Gamma^{*}\right)$ is the tensor-product interpolant

$$
\left(U^{i_{1}} \otimes \cdots \otimes U^{i_{M}}\right)(v)=\sum_{j_{1}=1}^{m_{i_{1}}} \cdots \sum_{j_{M}=1}^{m_{i_{M}}} v\left(\xi_{j_{1}}^{i_{1}} \cdots, \xi_{j_{M}}^{i_{M}}\right) L_{j_{1}}^{i_{1}}\left(\xi_{1}\right) \cdots L_{j_{1}}^{i_{M}}\left(\xi_{M}\right) .
$$

This requires the function values at $\prod_{i=1}^{M} m_{i}$ nodes, which is of exorbitant size for large $M$ and $m_{i}$. The number of nodes can be reduced dramatically using a sparse grid (Smolyak) operator [3],

$$
\begin{equation*}
A(q, M):=\sum_{q-M+1 \leq \mathbf{i} \mid \leq q}(-1)^{q-\mathbf{i} \mid}\binom{M-1}{q-|\mathbf{i}|}\left(U^{i_{1}} \otimes \cdots \otimes U^{i_{M}}\right), \tag{2.4}
\end{equation*}
$$

where $\mathbf{i} \in \mathbb{N}^{M},|\mathbf{i}|=i_{1}+\cdots+i_{M}$ and the index $q \geq M$ is called the sparse grid level. The sparse grid operator depends on function values at the sparse grid points

$$
\Theta_{q}:=\bigcup_{q-M+1 \leq \mathrm{i} \mid \leq q}\left(\Theta_{1}^{i_{1}} \otimes \cdots \otimes \Theta_{1}^{i_{M}}\right) .
$$

The size of the sample set (i.e., the number of sparse grid points) $\left|\Theta_{q}\right|$ can typically be chosen to be much smaller than $\prod_{i=1}^{M} m_{i}$ without significantly sacrificing interpolation accuracy [3].

Different choices of the one-dimensional partitioning sets $\left\{\Theta_{1}^{i}\right\}$ lead to different sparse grid operators, e.g., nested Clenshaw-Curtis abscissae $[3,32]$ and non-nested Gauss abscissae [8]. When nested abscissae are used, $A(q, M)(v)$ is an interpolant of $v(\xi)$ for any arbitrary $v(\xi) \in C(\Gamma)$. This is not true in general when non-nested abscissae are used.

In this study, we consider Clenshaw-Curtis interpolation, whose abscissae are the extrema of Chebyshev polynomials (see [32]), and $\Theta_{q}$ refers to the set consisting of Clenshaw-Curtis sparse grids. From [32], $\left|\Theta_{q}\right| \approx 2^{q-M} \frac{q!}{M!(q-M)!} \ll \prod_{i=1}^{M} m_{i}$. For this interpolation rule,

$$
A(q, M)(v) \in \sum_{|\mathrm{i}|=q}\left(P_{m_{i_{1}}-1} \otimes \cdots \otimes P_{m_{i_{M}}-1}\right),
$$

where $\boldsymbol{P}_{m_{i}-1}$ is the set of polynomials with degree at most $m_{i}-1$ and

$$
m_{i}=\left\{\begin{array}{cc}
1 & i=1 \\
2^{i-1}+1 & i>1
\end{array}\right.
$$

In addition, $A(q, M)(v)=v$ whenever $v \in \sum_{|\mathbf{i}|=q}\left(\boldsymbol{P}_{m_{i_{1}-1}} \otimes \cdots \otimes \boldsymbol{P}_{m_{i_{M}}-1}\right)$. Like the generic form of collocation solutions (2.1), the specific form for sparse grid sampling is denoted by

$$
\begin{equation*}
u_{q}^{s c}\left(\vec{x}, \xi^{(k)}\right):=\sum_{\xi^{(k)} \in \Theta_{q}} u_{c}\left(\vec{x}, \xi^{(k)}\right) L_{\xi^{(k)}}(\xi) \tag{2.5}
\end{equation*}
$$

where the interpolation polynomials $\left\{L_{\xi^{(k)}}\right\}$ come from the definition of the Smolyak operator (2.4). ${ }^{1}$

We can also use the sparse grid formulation to perform quadrature to approximate the moments of $u_{q}^{s c}$. For example, the Clenshaw-Curtis quadrature rule (see [22] for details) computes the mean of $u_{q}^{s c}(\vec{x}, \xi)$,

$$
\mathbb{E}\left(u_{q}^{s c}(\vec{x}, \xi)\right):=\int_{\Gamma} u_{q}^{s c}(\vec{x}, \xi) \rho(\xi) d \xi
$$

in the form

$$
\begin{equation*}
\tilde{\mathbb{E}}\left(u_{q}^{s c}(\vec{x}, \xi)\right):=\sum_{\xi^{(k)} \in \Theta_{q}} u_{c}\left(\vec{x}, \xi^{(k)}\right) \rho\left(\xi^{(k)}\right) w_{\xi^{(k)}} \tag{2.6}
\end{equation*}
$$

where $\left\{w_{\xi^{(k)}}\right\}$ are the weights of the Clenshaw-Curtis sparse grid quadrature. It can be seen that the evaluation of the mean function (2.6) does not entail evaluation of the interpolation polynomials $L_{\xi^{(k)}}$ in (2.5). An estimate for the variance can be computed in the same way.

Note that this Clenshaw-Curtis quadrature rule with level $q$ is exact for polynomials in the space

$$
\sum_{|\mathrm{i}|=q}\left(\boldsymbol{P}_{m_{i_{1}}} \otimes \cdots \otimes \boldsymbol{P}_{m_{i_{M}}}\right)
$$

This implies $\mathbb{E}\left(u_{q}^{s c}(\vec{x}, \xi)\right)=\tilde{\mathbb{E}}\left(u_{q}^{s c}(\vec{x}, \xi)\right)$ when $\xi$ is uniformly distributed, i.e., when the density function $\rho(\xi)$ is a constant.
3. Discretization and reduced basis approximation. In this section, we discuss finite element approximation and reduced basis methods for the problem (2.2)(2.3). To simplify the presentation in this section, we will assume the problem satisfies homogeneous Dirichlet conditions. It is straightforward to generalize the approach to nonhomogeneous conditions, which will be discussed in Section 6.

In general, we denote the weak form of the deterministic problem (2.2)-(2.3) corresponding to a given realization of $\xi$ by

$$
\mathfrak{B}_{\xi}(u(\cdot, \xi), v)=l(v) .
$$

[^1]Let $X^{h}$ be a spatial finite element approximation space (e.g., piecewise linear or quadratic polynomial spaces) of dimension $N_{h}$. A finite element formulation seeks an approximation $u_{h}(\cdot, \xi) \in X^{h}$ such that

$$
\begin{equation*}
\mathfrak{B}_{\xi}\left(u_{h}(\cdot, \xi), v\right)=l(v) \quad \forall v \in X^{h} . \tag{3.1}
\end{equation*}
$$

In the following, for any $\xi$, the finite element solution $u_{h}(\cdot, \xi)$ is referred to as a snapshot associated with $\xi$. Grouping solutions of (3.1) with respect to all $\xi \in \Gamma$ together, we define a set consisting of all snapshots

$$
S_{\Gamma}:=\left\{u_{h}(\cdot, \xi), \xi \in \Gamma\right\} .
$$

We will refer to this as the full snapshot set. Similarly, for a given finite set $\Theta \subset \Gamma$, we define a finite snapshot set

$$
S_{\Theta}:=\left\{u_{h}(\cdot, \xi), \xi \in \Theta\right\} .
$$

The matrix form of $S_{\Theta}$ is denoted by $\mathbf{S}_{\Theta} \in \mathbb{R}^{N_{h} \times|\Theta|}$, i.e., each column of $\mathbf{S}_{\Theta}$ is the vector of nodal point values of a finite element solution.

In this study, we assume that the spatial mesh is sufficiently fine so that the finite element discretization error is acceptable (we refer to standard a posteriori error estimation techniques, e.g., [1, 10, 29]). With this assumption, the Galerkin equation (3.1) typically has many degrees of freedom. On the other hand, the size of the sample set $\left|\Theta_{q}\right|$ approximately equals $2^{q-M} \frac{q!}{M!(q-M)!}$ [32]. Although this may be much smaller than the size of the tensor sample set, it may still be very large if high accuracy with respect to collocation error is needed. The combination of large-scale spatial discretization and large numbers of sample (collocation) points can cause the cost of sparse grid collocation to be unacceptably high. One aim hence is to show that these costs can be reduced through the use of reduced basis methods.

That is, suppose we have a set of basis functions $Q=\left\{q_{1}, \ldots, q_{N}\right\} \subset X^{h}$, where $N \ll N_{h}$. In the Galerkin formulation of the reduced basis method, we seek an approximate solution $u_{R}(\vec{x}, \xi) \in \operatorname{span}(Q)$ such that

$$
\begin{equation*}
\mathfrak{B}_{\xi}\left(u_{R}(\cdot, \xi), v\right)=l(v) \quad \forall v \in \operatorname{span}(Q) . \tag{3.2}
\end{equation*}
$$

The reduced problem (3.2) tends to be much smaller than the full problem (3.1) [4, 20, 30]. Because of this, it is also likely to be cheaper to solve, especially if the computation is done carefully using precomputed quantities when possible. In the following, we demonstrate this approach using two types of benchmark problems. For the first, we use the diffusion equation as a prototypical example of a linear problem; the approach considered is generally applicable to linear problems, such as the Stokes equations. For the second, we explore the methodology for quadratic problems as examplified by the Navier-Stokes equations. For highly nonlinear operators, we refer to [7].
3.1. Linear operators. When the operator $\mathcal{L}$ of (1.3) is linear as a spatial differential operator, the discrete weak formulation (3.1) leads to a linear system

$$
\begin{equation*}
\mathbf{A}_{\xi} \mathbf{u}_{\xi}=\mathbf{f} \tag{3.3}
\end{equation*}
$$

of order $N_{h}$. Since $\mathcal{L}$ is linear, $\mathbf{A}_{\xi}$ is independent of $\mathbf{u}_{\xi}$, and by assumption it is affinely dependent on the parameter $\xi \in \Gamma$. That is, it has the form

$$
\begin{equation*}
\mathbf{A}_{\xi}=\sum_{i=1}^{K} \phi_{i}(\xi) A_{i}, \tag{3.4}
\end{equation*}
$$

where $\left\{A_{i}\right\}$ are parameter-independent matrices and $\phi_{i}(\xi) \in \mathbb{R}$.
For $q_{i} \in Q$, let $\mathbf{q}_{i}$ be the vector of nodal coefficient values associated with $q_{i}$, and let $\mathbf{Q}=\left[\mathbf{q}_{1}, \ldots, \mathbf{q}_{N}\right] \in \mathbb{R}^{N_{h} \times N}$ be the matrix representation of $Q$. Then the linear system associated with the reduced problem (3.2) can be written as

$$
\begin{equation*}
\mathbf{Q}^{T} \mathbf{A}_{\xi} \mathbf{Q} \tilde{\mathbf{u}}_{\xi}=\mathbf{Q}^{T} \mathbf{f} \tag{3.5}
\end{equation*}
$$

With the expansion (3.4), (3.5) can be written as

$$
\begin{equation*}
(\sum_{i=1}^{K} \phi_{i}(\xi) \underbrace{\mathbf{Q}^{T} A_{i} \mathbf{Q}}_{A_{R, i}}) \tilde{\mathbf{u}}_{\xi}=\underbrace{\mathbf{Q}^{T} \mathbf{f}}_{\mathbf{f}_{R}} . \tag{3.6}
\end{equation*}
$$

Once the parameter-independent reduced matrices $\left\{A_{R, i}\right\}$ and vector $\mathbf{f}_{R}$ are precomputed, the reduced problem (3.5) for any arbitrary point $\xi \in \Gamma$ can be assembled with $O\left(N^{2}\right)$ operations. For $u_{R}$ defined by (3.2), we will also need an estimate for a norm of the error $e_{\xi}:=u_{h}(\cdot, \xi)-u_{R}(\cdot, \xi)$, which we would also like to compute with complexity dependent only on $N$ and not $N_{h}$. This means that standard a posteriori error estimators $[1,14,17]$ for finite element methods should be excluded, since they would incur a cost proportional to $N_{h}$. An effective alternative is the dual-based indicator developed in $[25,30,31]$, which depends on an estimate for the coercivity constant associated with the problem. If only a rough error estimate is required, we can use a simple residual indicator

$$
\begin{equation*}
\eta_{Q, \xi}:=\frac{\left\|\mathbf{A}_{\xi} \mathbf{Q} \tilde{\mathbf{u}}_{\xi}-\mathbf{f}\right\|_{2}}{\|\mathbf{f}\|_{2}} \tag{3.7}
\end{equation*}
$$

We will use this in our experiments. It can be computed efficiently using the relation

$$
\begin{align*}
\left\|\mathbf{A}_{\xi} \mathbf{Q} \tilde{\mathbf{u}}_{\xi}-\mathbf{f}\right\|_{2}^{2} & =\left(\mathbf{A}_{\xi} \mathbf{Q} \tilde{\mathbf{u}}_{\xi}-\mathbf{f}, \mathbf{A}_{\xi} \mathbf{Q} \tilde{\mathbf{u}}_{\xi}-\mathbf{f}\right)=\left(\mathbf{A}_{\xi} \mathbf{Q} \tilde{\mathbf{u}}_{\xi}, \mathbf{A}_{\xi} \mathbf{Q} \tilde{\mathbf{u}}_{\xi}\right)-2\left(\mathbf{A}_{\xi} \mathbf{Q} \tilde{\mathbf{u}}_{\xi}, \mathbf{f}\right)+(\mathbf{f}, \mathbf{f}) \\
& =\left(\left(\sum_{i=1}^{K} \phi_{i} A_{i}\right) \mathbf{Q} \tilde{\mathbf{u}}_{\xi},\left(\sum_{i=1}^{K} \phi_{i} A_{i}\right) \mathbf{Q} \tilde{\mathbf{u}}_{\xi}\right)-2\left(\left(\sum_{i=1}^{K} \phi_{i} A_{i}\right) \mathbf{Q} \tilde{\mathbf{u}}_{\xi}, \mathbf{f}\right)+(\mathbf{f}, \mathbf{f}) \\
3.8) \quad & =\tilde{\mathbf{u}}_{\xi}^{T}\left(\sum_{i=1}^{K} \sum_{j=1}^{K} \phi_{i} \phi_{j} \mathbf{Q}^{T} A_{i}^{T} A_{j} \mathbf{Q}\right) \tilde{\mathbf{u}}_{\xi}-2 \tilde{\mathbf{u}}_{\xi}^{T} \sum_{i=1}^{K}\left(\phi_{i} \mathbf{Q}^{T} A_{i}^{T} \mathbf{f}\right)+\mathbf{f}^{T} \mathbf{f} . \tag{3.8}
\end{align*}
$$

Once the matrices $\left\{\mathbf{Q}^{T} A_{i}^{T} A_{j} \mathbf{Q}\right\}$, vectors $\left\{\mathbf{Q}^{T} A_{i}^{T} \mathbf{f}\right\}$ and $\mathbf{f}^{T} \mathbf{f}$ are precomputed, the cost for computing $\eta_{Q, \xi}$ is $O\left(N^{2}\right)$. Similar economies can be achieved with dual-based estimators [30,31]. We refer to the reduced dense matrices and vectors $\left\{A_{R, i}\right\}, \mathbf{f}_{R}$ in (3.6), and $\left\{\mathbf{Q}^{T} A_{i}^{T} A_{j} \mathbf{Q}\right\},\left\{\mathbf{Q}^{T} A_{i}^{T} \mathbf{f}\right\}, \mathbf{f}^{T} \mathbf{f}$ in (3.8) as offline reduced matrices and vectors.
3.2. Quadratic operators. When $\mathcal{L}$ depends quadratically on the solution $u$, the algebraic form of the discrete problem (3.1) can be written as

$$
\begin{equation*}
\mathbf{A}_{\mathbf{u}, \xi} \mathbf{u}=\mathbf{f} \tag{3.9}
\end{equation*}
$$

where $\mathbf{A}_{\mathbf{u}, \xi}$ is linearly dependent on $\mathbf{u}$, i.e., if the solution takes the form $\mathbf{u}=\sum_{i} c_{i} \mathbf{u}_{i}$, then

$$
\begin{equation*}
\mathbf{A}_{\mathbf{u}, \xi}=\sum_{i} c_{i} \mathbf{A}_{\mathbf{u}_{i}, \xi} \tag{3.10}
\end{equation*}
$$

Solving (3.9) requires a nonlinear iteration that entails the solution of a linearized problem at every step, for example, as in Picard and Newton iterations [10, pp. 327329]. At iteration step $n$, the linearized problem for computing the solution at step $n+1$ can be written as

$$
\begin{equation*}
\mathbf{A}_{\mathbf{u}^{n}, \xi} \mathbf{u}^{n+1}=\mathbf{f} \tag{3.11}
\end{equation*}
$$

Like (3.5), the reduced version of (3.11) is

$$
\begin{equation*}
\mathbf{Q}^{T} \mathbf{A}_{\mathbf{u}^{n}, \xi} \mathbf{Q} \tilde{\mathbf{u}}^{n+1}=\mathbf{Q}^{T} \mathbf{f}, \tag{3.12}
\end{equation*}
$$

where $\mathbf{u}^{n}=\mathbf{Q} \tilde{\mathbf{u}}^{n}=\sum_{i=1}^{N} \tilde{u}_{i}^{n} \mathbf{q}_{i}$ with $\tilde{\mathbf{u}}^{n}=\left[\tilde{u}_{1}^{n}, \ldots, \tilde{u}_{N}^{n}\right]^{T}$ and $\mathbf{Q}=\left[\mathbf{q}_{1}, \ldots, \mathbf{q}_{N}\right]$. Using (3.10), we can rewrite (3.12) as

$$
\begin{equation*}
\left(\sum_{i=1}^{N} \tilde{u}_{i}^{n} \mathbf{Q}^{T} \mathbf{A}_{\mathbf{q}_{i}, \xi} \mathbf{Q}\right) \tilde{\mathbf{u}}^{n+1}=\mathbf{Q}^{T} \mathbf{f} \tag{3.13}
\end{equation*}
$$

Under the assumption that $\mathbf{A}_{\mathbf{q}_{i}, \xi}$ is affinely parameter-dependent, which implies

$$
\begin{equation*}
\mathbf{A}_{\mathbf{q}_{i}, \xi}=\sum_{j=1}^{K} \phi_{j}(\xi) A_{\mathbf{q}_{i}, j} \tag{3.14}
\end{equation*}
$$

the reduced linear system for quadratic operators can be finally stated as

$$
\begin{equation*}
(\sum_{i=1}^{N} \sum_{j=1}^{K} \phi_{j}(\xi) \tilde{u}_{i}^{n} \underbrace{\mathbf{Q}^{T} A_{\mathbf{q}_{i}, j} \mathbf{Q}}_{\tilde{A}_{i, j}}) \tilde{\mathbf{u}}^{n+1}=\mathbf{Q}^{T} \mathbf{f} \tag{3.15}
\end{equation*}
$$

Once the parameter-independent and solution-independent matrices $\left\{\tilde{A}_{i, j}\right\}$ are precomputed, the reduced system (3.15) can be assembled with a cost $O\left(N^{3}\right)$. Similarly to (3.8), we can also develop a reduced version of the residual indicator

$$
\begin{equation*}
\eta_{Q, \xi, n}:=\frac{\left\|\mathbf{A}_{\mathbf{u}^{n}, \xi} \mathbf{Q} \tilde{\mathbf{u}}^{n}-\mathbf{f}\right\|_{2}}{\|\mathbf{f}\|_{2}} \tag{3.16}
\end{equation*}
$$

With some precomputed offline reduced matrices and vectors as discussed in Section 3.1, it can be verified that computing the reduced residual indicator for quadratic operators also costs $O\left(N^{3}\right)$. Cf. [30] for efficient methods for computing dual-based error estimates.

In the next section, a systematic way for constructing the reduced basis $\mathbf{Q}$ together with computing the collocation solution is introduced.
4. Reduced basis collocation. We introduce a reduced basis collocation method for cheaply computing the collocation solution $u_{q}^{s c}(\cdot, \xi)$ in (2.5). Our main idea is to use $u_{R}$, the solution of the reduced problem (3.2), to serve in place of the collocation coefficient function $u_{c}(\cdot, \xi)$ in (2.5) wherever possible. That is, given a collocation point $\xi^{(k)}$, we compute $u_{R}\left(\cdot, \xi^{(k)}\right.$ ) and an error indicator such as (3.7) or (3.16). If this error indicator is smaller than some specified tolerance, we use $u_{R}\left(\cdot, \xi^{(k)}\right)$; if the error indicator is too large, then we compute the snapshot $u_{h}\left(\cdot, \xi^{(k)}\right)$ and use it as $u_{c}(\cdot, \xi)$ in (2.5). In the latter case, we also augment the reduced basis with this snapshot. Our strategy is described in detail as follows.

1. Starting with level $p=M$ (the set $\Theta_{M}$ has only one point, which is denoted by $\xi^{(0)}$ ), compute the snapshot $u_{h}\left(\cdot, \xi^{(0)}\right)$. Initialize the reduced basis $Q=$ $\left\{u_{h}\left(\cdot, \xi^{(0)}\right)\right\}$. In addition, use $u_{h}\left(\cdot, \xi^{(0)}\right)$ to serve as the coefficient function $u_{c}\left(\cdot, \xi^{(0)}\right)$ in (2.5).
2. Consider one higher level, i.e. $p+1$. Looping over sparse grid points in level $p+1$, for each point $\xi^{(k)}$, compute the reduced solution $u_{R}\left(\cdot, \xi^{(k)}\right.$ ) in (3.2) and estimate a norm of the error $e_{\xi^{(k)}}=u_{h}\left(\cdot, \xi^{(k)}\right)-u_{R}\left(\cdot, \xi^{(k)}\right)$.
(a) If the estimated error is smaller than a given tolerance, use $u_{R}\left(\cdot, \xi^{(k)}\right)$ to serve as the coefficient function $u_{c}\left(\cdot, \xi^{(k)}\right)$ in (2.5).
(b) If the estimated error is larger than the tolerance, compute the snapshot $u_{h}\left(\cdot, \xi^{(k)}\right)$ and augment the reduced basis $Q$ with it. Then, use $u_{h}\left(\cdot, \xi^{(k)}\right)$ to serve as the coefficient function $u_{c}\left(\cdot, \xi^{(k)}\right)$ in (2.5).
3. Update the sparse grid level (i.e. let $p=p+1$ ), and repeat step 2 , until the level $p$ reaches the given maximum level $q$.
This strategy is stated more formally in Algorithm 1 below. Here, tol stands for a given tolerance and the orthogonal complement $T\left(\mathbf{u}_{\xi^{(k)}}\right)$ is defined as

$$
T\left(\mathbf{u}_{\xi^{(k)}}\right)=\mathbf{u}_{\xi^{(k)}}-\Pi_{\mathbf{Q}}\left(\mathbf{u}_{\xi^{(k)}}\right),
$$

where $\Pi_{\mathbf{Q}}\left(\mathbf{u}_{\xi^{(k)}}\right)$ is the image of $\mathbf{u}_{\xi^{(k)}}$ under the $L^{2}$-projection from $\mathbb{R}^{N_{h}}$ to $\operatorname{span}\left\{\mathbf{q}_{i}\right\}_{i=1}^{N}$. $T\left(\mathbf{u}_{\xi^{(k)}}\right)$ can be computed using the Gram-Schmidt process, as implemented in the MATLAB function $\mathbf{q r}$. In the sequel, the sparse grid collocation solution associated with Algorithm 1 is denoted by $u_{q}^{r s c}$, and the full sparse grid collocation solution whose coefficient functions are standard finite element solutions, is denoted by $u_{q}^{\text {hsc }}$. In general, computing $u_{q}^{r s c}$ should be much less expensive than computing $u_{q}^{h s c}$, and as we will show in the following sections, the accuracy of the reduced solution is often comparable to that of the full solution.
5. Numerical study for diffusion problems. In this section, we consider diffusion problems, whose governing equations are

$$
\begin{array}{rlll}
-\nabla \cdot a(\cdot, \xi) \nabla u(\cdot, \xi)=1 & \text { in } & D \times \Gamma, \\
u(\cdot, \xi)=0 & & \text { on } & \partial D_{D} \times \Gamma, \\
\frac{\partial u(\cdot, \xi)}{\partial n}=0 & & \text { on } &  \tag{5.3}\\
\partial D_{N} \times \Gamma,
\end{array}
$$

```
Algorithm 1 Reduced basis collocation, homogeneous boundary conditions
    Compute \(S_{\Theta_{M}}\left(\Theta_{M}\right.\) contains only one point).
    Initialize the reduced basis matrix \(\mathbf{Q}:=\mathbf{S}_{\Theta_{M}} /\left\|\mathbf{S}_{\Theta_{M}}\right\|_{2}\).
    Construct the offline reduced matrices and vectors.
    for \(p=1: q\) do
        for \(k=1:\left|\Theta_{M+p}\right|\) do
            Compute \(\tilde{\mathbf{u}}_{\xi^{(k)}}\) by solving (3.6) and compute an error indicator, e.g., \(\eta_{Q, \xi^{(k)}}\) of
            (3.7).
            if \(\eta_{Q, \xi^{(k)}}<t o l\) then
                    Use the reduced solution derived from \(\tilde{\mathbf{u}}_{\xi^{(k)}}\) to serve as \(u_{c}\left(\cdot, \xi^{(k)}\right)\) in (2.5).
            else
                    Compute the full solution vector \(\mathbf{u}_{\xi^{(k)}}\) by solving (3.3).
                    Use the full solution derived from \(\mathbf{u}_{\xi^{(k)}}\) to serve as \(u_{c}\left(\cdot, \xi^{(k)}\right)\) in (2.5).
                    Compute \(T\left(\mathbf{u}_{\xi^{(k)}}\right)\), the orthogonal complement of \(\mathbf{u}_{\xi^{(k)}}\) with respect to \(\mathbf{Q}\).
                    Augment the reduced basis matrix \(\mathbf{Q}:=\left[\mathbf{Q}, T\left(\mathbf{u}_{\xi^{(k)}}\right)\right]\).
                    Reconstruct the offline reduced matrices and vectors.
            end if
        end for
    end for
```

where $\partial D=\partial D_{D} \cup \partial D_{N}$. The weak formulation is to find $u(\cdot, \xi) \in H_{0}^{1}(D)$ such that $(a \nabla u, \nabla v)=(1, v)$ for all $v \in H_{0}^{1}(D)$. We discretize in space using a bilinear $\left(Q_{1}\right)$ finite element approximation [5, 10].

To assess the accuracy of solutions obtained using the full and reduced basis collocation methods, we use the differences between the means of the solutions and that of a reference solution. In particular, we introduce the quantities

$$
\begin{align*}
& \epsilon_{h}:=\left\|\tilde{\mathbb{E}}\left(u_{q}^{h s c}\right)-\tilde{\mathbb{E}}\left(u_{r}^{h s c}\right)\right\|_{0} /\left\|\tilde{\mathbb{E}}\left(u_{r}^{h s c}\right)\right\|_{0},  \tag{5.4}\\
& \epsilon_{R}:=\left\|\tilde{\mathbb{E}}\left(u_{q}^{r s c}\right)-\tilde{\mathbb{E}}\left(u_{r}^{h s c}\right)\right\|_{0} /\left\|\tilde{\mathbb{E}}\left(u_{r}^{h s c}\right)\right\|_{0}, \tag{5.5}
\end{align*}
$$

where the reference collocation solution $u_{r}^{h s c}$ is a full collocation solution with a large grid level $r$ (we take $r \geq q+2$ ) and the norm is the functional $L^{2}$-norm. We will also examine the performance of Monte Carlo simulation using a sample set $\Theta_{m c}$ consisting of $\left|\Theta_{m c}\right|$ realizations of $\xi$, where the Monte Carlo error is measured as

$$
\begin{equation*}
\epsilon_{m c}:=\left\|\left(\frac{1}{\Theta_{m c}} \sum_{\xi \in \Theta_{m c}} u_{h}(\cdot, \xi)\right)-\tilde{\mathbb{E}}\left(u_{r}^{h s c}\right)\right\|_{0} /\left\|\tilde{\mathbb{E}}\left(u_{r}^{h s c}\right)\right\|_{0} \tag{5.6}
\end{equation*}
$$

We note that reduced basis methods can also be combined with Monte Carlo methods, as discussed in [4].
5.1. Test problem 1: piecewise constant diffusion coefficient. We consider the diffusion problem posed on the spatial domain $D=(-1,1) \times(-1,1)$, divided into $N_{D}$
equal-sized subdomains. A pure Dirichlet condition is applied (i.e., $\partial D_{N}=\emptyset$ ). Figure 5.1 illustrates two cases of domain partitionings. The permeability coefficient $a(\cdot, \xi)$ is defined to be constant on each subdomain, i.e.,

$$
\left.a(\cdot, \xi)\right|_{D_{k}}=\xi_{k}, \quad k=1: N_{D}
$$

where the random variable $\xi=\left[\xi_{1}, \ldots, \xi_{k}, \ldots, \xi_{N_{D}}\right]^{T}$ is independently and uniformly distributed in $\Gamma:=[0.01,1]^{N_{D}}$. It is also assumed that the flux $a(\cdot, \xi) \nabla u(\cdot, \xi)$ is continuous across subdomain interfaces. We consider two variants of this example, one where the domains consist of vertical strips, the other using squares (see Figure 5.1).

(a) case 1: $N_{D}$ subdomains

(b) case 2: $N_{D}=\tilde{N} \times \tilde{N}$ subdomains

FIG. 5.1. Domain partitionings.
It follows from (3.2) that the reduced basis method seeks a solution in the space $\operatorname{span}(Q)$. If each function in the full snapshot set $S_{\Gamma}$ can be approximated well by a linear combination of a finite set of linearly independent functions (referred to as a "basis" of $\left.S_{\Gamma}\right)$, then with this set as the reduced basis $Q$, the reduced solution $u_{R}(\cdot, \xi)$ is close to the finite element solution $u_{h}(\cdot, \xi)$. The size of this basis (we refer to it below as the "rank" of $S_{\Gamma}$ ) is then crucial. If the rank of $S_{\Gamma}$ is much smaller than $N_{h}$, then Algorithm 1 can cheaply compute an accurate reduced collocation solution (i.e., $\left.u_{q}^{r s c} \approx u_{q}^{h s c}\right)$.

For this test problem, we check the rank of $S_{\Gamma}$ as follows. We first generate a sample set $\Theta$ consisting of 3000 random points in $\Gamma$ and construct the corresponding snapshot set $S_{\Theta} .^{2}$ We use the MATLAB function rank to compute the rank of $\mathbf{S}_{\Theta}$ (the matrix of $S_{\Theta}$, see Section 3). Due to the large number of sample points, this rank can serve an estimate of the rank of the full snapshot set $S_{\Gamma}$.

Tables 5.1 and 5.2 show the estimated ranks for the two variants of the benchmark diffusion problems, using three different mesh sizes for the spatial discretization. It can be seen that the ranks tend to increase linearly with the number of subdomains, and they exhibit little dependence on the size of the grid. This suggests that the rank depends on properties of the underlying PDE and it indicates that fine-grid discrete problems can be projected into subspaces of significantly smaller dimensions without sacrificing accuracy. (Although for large $N_{D}$, in the cases of 36 or more square

[^2]Table 5.1
Estimated rank for the full snapshot set $S_{\Gamma}$ of test problem 1, case 1.

|  | $N_{D}$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Grid |  |  |  |  |  |  |  |  |  |
| $33^{2}=1089$ | 3 | 12 | 18 | 30 | 40 | 53 | 55 | 76 | 84 |
| $65^{2}=4225$ | 3 | 12 | 18 | 30 | 40 | 48 | 55 | 70 | 87 |
| $129^{2}=16641$ | 3 | 12 | 18 | 28 | 39 | 48 | 55 | 72 | 81 |

Table 5.2
Estimated rank for the full snapshot set $S_{\Gamma}$ of test problem 1, case 2.

|  | $N_{D}$ | 4 | 9 | 16 | 25 | 36 | 49 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Grid |  |  |  |  |  |  |  |
| $33^{2}=1089$ | 27 | 121 | 193 | 257 | 321 | 385 | 449 |
| $65^{2}=4225$ | 28 | 148 | 290 | 465 | 621 | 769 | 897 |
| $129^{2}=16641$ | 28 | 153 | 311 | 497 | 746 | 1016 | 1298 |

subdomains (see Table 5.2), the ranks are growing as the mesh is refined, they also appear to be tending to a limit with increasing number of grid points.) We also note that for some of these cases (of large $N_{D}$ ), the large ranks will make the reduced problems expensive to solve, although these costs would be smaller than those of the full system solves for fine enough spatial discretization.

Next, we use Algorithm 1 to compute the collocation solutions for a collection of examples of test problem 1. Tables 5.3-5.4 show results for case 1 (vertical subdomains) and Tables 5.5-5.6 for case 2 (square subdomains); all these experiments used a $65 \times 65$ discrete spatial grid. The tables show the number of full system solves $N_{\text {full solve }}$ and the size of the sample set $\left|\Theta_{q}\right|$. For example, in the case of $2 \times 2$ subdomains and tol $=10^{-4}$ (Table 5.5), there are seven full system solves at the sparse grid level $q=5$, which means that the residual error indicator is above the tolerance at seven sparse grid points among the total of nine sparse grid collocation points. This is not surprising, since the size of the reduced basis is very small at this stage (it grows from 1 to 8 ), and in this trivial case, there is no advantage for the reduced basis. For higher levels of sparse grids, however, the advantages become clear. At level $q=6$, there are 41 sparse grid points, and full system solves are needed at 12 of them, and at level $q=7$, full solves are needed at just 3 of 137 sparse grid points. For levels higher than 7, no full system solve is needed, which means that the reduced basis with size $N=23$ can provide as accurate (with respect to the error indicator) a solution as the full collocation solution. This trend holds for all the examples: the number of full system solves required to generate the reduced collocation solution is dramatically smaller than the number of collocation points. Moreover, the required number of full system solves needed for the the reduced basis is comparable to the ranks of the sets of full snapshots shown in Tables 5.1-5.2. For example, with four $(2 \times 2)$ square subdomains, the rank is 28 .

Table 5.3
Number of full system solves for test problem 1, case 1, with $5 \times 1$ subdomains $\left(N_{D}=5\right)$ and a $65 \times 65$ spatial grid.

| $q$ | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 16 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ${ }_{\text {tol }}\left\|\Theta_{q}\right\|$ | 1 | 11 | 61 | 241 | 801 | 2433 | 6993 | 19313 | 51713 | 869505 |
| $10^{-3}$ | 1 | 10 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $10^{-4}$ | 1 | 10 | 11 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| $10^{-5}$ | 1 | 10 | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Table 5.4
Number of full system solves for test problem 1, case 1, with $9 \times 1$ subdomains ( $N_{D}=9$ ), tol $=10^{-4}$ and a $65 \times 65$ spatial grid.

| $q$ | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\left\|\Theta_{q}\right\|$ | 1 | 19 | 181 | 1177 | 6001 | 26017 | 100897 | 361249 | 1218049 |
| $N_{\text {full solve }}$ | 1 | 18 | 34 | 2 | 1 | 1 | 0 | 0 | 0 |

For the cases of $5 \times 1$ and $2 \times 2$ subdomains, Figure 5.2 provides a more refined assessment of accuracy, using the relative mean function errors (5.5) (for tol $=10^{-4}$ ), where the reference levels are taken to be $r=18$ for five vertical subdomains and $r=17$ for four square subdomains. The figure also shows similar quantities for the Monte Carlo method, $\epsilon_{m c}$ of (5.6). The errors for the full collocation means ( $\epsilon_{h}$ of (5.4)) are also plotted, but there is no visual difference between $\epsilon_{h}$ and $\epsilon_{R}$. Thus, the reduced collocation solution is as accurate as the full collocation solution, and it is considerably more accurate than the Monte Carlo solution.
5.2. Test problem 2: truncated KL expansion coefficients. The spatial domain for this test problem is $D=(0,1) \times(0,1)$. Mixed boundary conditions are applied - the condition (5.2) is applied on the left $(x=0)$ and right $(x=1)$ boundaries, and (5.3) is applied on the top and bottom boundaries. The problem is discretized in space on a uniform $65 \times 65$ grid.

The diffusion coefficient for this test problem is assumed to be a random field with mean function $a_{0}(\vec{x})$, constant variance $\sigma$ and covariance function $C(\vec{x}, \vec{y})$,

$$
\begin{equation*}
C(\vec{x}, \vec{y})=\sigma \exp \left(-\frac{\left|x_{1}-y_{1}\right|}{c}-\frac{\left|x_{2}-y_{2}\right|}{c}\right), \tag{5.7}
\end{equation*}
$$

where $c$ is the correlation length. This random field can be approximated by a truncated Karhunen-Loève expansion $[2,8,11]$

$$
a(\vec{x}, \xi) \approx a_{0}(\vec{x})+\sum_{k=1}^{M} \sqrt{\lambda_{k}} a_{k}(\vec{x}) \xi_{k}
$$

where $a_{k}(\vec{x})$ and $\lambda_{k}$ are the eigenfunctions and eigenvalues of (5.7), and the random

Table 5.5
Number of full system solves for test problem 1, case 2, with $2 \times 2$ subdomains ( $N_{D}=4$ ) and a $65 \times 65$ spatial grid.

| $q$ | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 15 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\left\|\Theta_{q}\right\|$ | 1 | 9 | 41 | 137 | 401 | 1105 | 2929 | 7537 | 18945 | 271617 |
| $10^{-3}$ | 1 | 7 | 11 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| $10^{-4}$ | 1 | 7 | 12 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| $10^{-5}$ | 1 | 7 | 13 | 2 | 3 | 0 | 0 | 0 | 0 | 0 |

Table 5.6
Number of full system solves for test problem 1, case 2, with $4 \times 4$ subdomains ( $N_{D}=16$ ), tol $=10^{-4}$ and $a 65 \times 65$ spatial grid.

| $q$ | 16 | 17 | 18 | 19 | 20 | 21 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\left\|\Theta_{q}\right\|$ | 1 | 33 | 545 | 6049 | 51137 | 353729 |
| $N_{\text {full solve }}$ | 1 | 32 | 168 | 27 | 3 | 4 |

variables $\left\{\xi_{k}\right\}$ are assumed to be independently and uniformly distributed in $\Gamma:=$ $[-1,1]^{M}$.

The error associated with truncation of the Karhunen-Loève expansion depends on the amount of total variance captured, $\delta_{K L}:=\left(\sum_{k=1}^{M} \lambda_{k}\right) /\left(|D| \sigma^{2}\right)[6,26]$. We chose $M$ to be large enough so that $\delta_{K L}>95 \%$. The correlation length has an effect on this requirement - small $c$ leads to large $M$.

For our experiments, we set $a_{0}(\vec{x})=1$ and $\sigma=0.5$ and examine two values of the correlation length: $c=4$ with $M=5$ and $c=2.5$ with $M=8$. Tables 5.7 and 5.8 show the numbers of full system solves needed in Algorithm 1, for various choices of the level $q$ and tolerance tol. Figure 5.3 shows the relative errors $\epsilon_{h}, \epsilon_{R}$ and $\epsilon_{m c}$. The reference solutions correspond to reference levels $r=12$ for $M=5$ and $r=15$ for $M=8$.

The results for this example are consistent with those for problem 1. As the tolerance tol decreases or the sparse grid level $q$ increases, somewhat more full systems need to be solved, but as above the number of such solves is dramatically lower than is needed for full collocation. Moreover, the sizes of the reduced basis are very small, so the reduced system solves are inexpensive. In particular, for the extreme cases in the two examples, the full collocation method required 2433 and 15713 solves respectively, in contrast to at most 62 and 115 for reduced basis collocation. Figure 5.3 shows again that that there is little significant difference between the full collocation and reduced collocation solutions, and that mild tolerances for constructing the reduced basis can achieve acceptable accuracy in the reduced collocation solution.
5.3. Tolerance and error indicator in Algorithm 1. We discuss some issues related to the use of the residual error indicator (3.7) and the tolerance tol in Algorithm (1). For the error indicator, first, as observed above, an advantage of (3.7) is


Fig. 5.2. Test problem 1, $\epsilon_{R}\left(\right.$ for tol $\left.=10^{-4}\right), \epsilon_{h}$ and $\epsilon_{m c}$.
Table 5.7
Number of full system solves for $c=4$ with $M=5$ in test problem 2 .

| $q$ | 5 | 6 | 7 | 8 | 9 | 10 |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ${ }_{\text {tol }}\left\|\Theta_{q}\right\|$ | 1 | 11 | 61 | 241 | 801 | 2433 | Total |
| $10^{-5}$ | 1 | 9 | 8 | 3 | 0 | 2 | 23 |
| $10^{-6}$ | 1 | 10 | 12 | 10 | 2 | 1 | 36 |
| $10^{-7}$ | 1 | 10 | 21 | 9 | 7 | 2 | 50 |
| $10^{-8}$ | 1 | 10 | 26 | 18 | 5 | 2 | 62 |

that it can be computed at cost independent of $N_{h}$ using (3.8). However, this advantage holds only if the tolerance is not too small. In particular, since (3.8) requires a subtraction, the floating point computation will be accurate only if it not strongly affected by cancellation, which is true only when the square of the residual norm is significantly larger than the machine precision. Thus, we can use this (economical) strategy only if the tolerance is not too small, on the order of $10^{-7}$ or larger.

For the results shown in Tables 5.3, 5.7 and 5.8, we found (3.8) to be reliable for the tolerances above the dotted lines. For the results below these lines, we computed the residual norm directly, which incurs a cost proportional to $N_{h}$. Our expectation is that this cost can be avoided through use of a more effective error indicator such as the dual-based method of [25, 30].

Tables 5.3-5.8 together with Figures 5.2 and 5.3 show the impact of the $t o l$ on the performance of Algorithm 1. For test problem 1, Figure 5.2 shows that with a modest value tol $=10^{-4}$, the errors for reduced collocation are virtually identical to those for full collocation, and (for both methods) as the level $q$ is increased, the relative errors are reduced by approximately six orders of magnitude. We also found

Table 5.8
Number of full system solves for $c=2.5$ with $M=8$ in test problem 2.

| $q$ | 8 | 9 | 10 | 11 | 12 | 13 |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| ${ }_{\text {tol }}\left\|\Theta_{q}\right\|$ | 1 | 17 | 145 | 849 | 3937 | 15713 | Total |
| $10^{-5}$ | 1 | 14 | 16 | 4 | 1 | 1 | 37 |
| $10^{-6}$ | 1 | 16 | 27 | 6 | 10 | 2 | 62 |
| $10^{-7}$ | 1 | 16 | 50 | 9 | 8 | 2 | 86 |
| $10^{-8}$ | 1 | 16 | 69 | 16 | 5 | 8 | 115 |



Fig. 5.3. Test problem 2: $\epsilon_{h}$ and $\epsilon_{R}$.
there to be virtually no difference between these results and those for $t o l=10^{-3}$ or, for $2 \times 2$ subdomains, tol $=10^{-5}$. Performance for test problem 2 is more sensitive to tolerance (Figure 5.3). A tolerance of $10^{-6}$ produces solutions with approximately six digits of accuracy for fine enough sparse grid, although more stringent tolerance is needed for accuracy comparable to that of the full collocation method.
6. Numerical study for incompressible flow problems. We next consider a nonlinear example, the steady-state Navier-Stokes equations

$$
\begin{array}{rll}
-v(\cdot, \xi) \nabla^{2} \vec{u}(\cdot, \xi)+\vec{u}(\cdot, \xi) \cdot \nabla \vec{u}(\cdot, \xi)+\nabla p(\cdot, \xi)=0 & \text { in } & D \times \Gamma, \\
\nabla \cdot \vec{u}(\cdot, \xi)=0 & \text { in } D \times \Gamma, \\
\vec{u}(\cdot, \xi)=\vec{g}(\cdot, \xi) & \text { on } & \partial D \times \Gamma . \tag{6.3}
\end{array}
$$

The notation in (6.1)-(6.3) is standard: $\vec{u}(\cdot, \xi)$ is the flow velocity, $p(\cdot, \xi)$ is the scalar pressure and $v(\cdot, \xi)>0$ is the fluid viscosity parameter. We assume that there may be some uncertainty in the viscosity parameter $v(\cdot, \xi)$ (for example, in models of multiphase flows $[15,23,28]$ ) or the boundary data $\vec{g}(\cdot, \xi)$.
6.1. Specification of the problem. As discussed in Section 2, stochastic collocation methods solve a deterministic problem at each sample point $\xi \in \Theta_{q}$. With the
standard function space notation

$$
\begin{array}{ll}
\boldsymbol{H}^{k}:=H^{k}(\Omega)^{2}, \quad k \in \mathbb{N}, & \boldsymbol{H}_{E}^{1}:=\left\{\vec{u} \in \boldsymbol{H}^{1} \mid \vec{u}=\vec{g}(\cdot, \xi) \text { on } \partial \Omega\right\}, \\
\boldsymbol{H}_{0}^{1}:=\left\{\vec{u} \in \boldsymbol{H}^{1} \mid \vec{u}=\overrightarrow{0} \text { on } \partial \Omega\right\}, & L_{0}^{2}(\Omega):=\left\{q \in L^{2}(\Omega) \mid \int_{\Omega} q d \Omega=0\right\},
\end{array}
$$

the weak form of the deterministic problem associated with (6.1)-(6.3) is: find $\vec{u} \in$ $\boldsymbol{H}_{E}^{1}$ and $p \in L_{0}^{2}(D)$, such that

$$
\begin{align*}
(\nu \nabla \vec{u}, \nabla \vec{v})+(\vec{u} \cdot \nabla \vec{u}, \vec{v})-(p, \nabla \cdot \vec{v}) & =0 & & \forall \vec{v} \in \boldsymbol{H}_{0}^{1},  \tag{6.4}\\
(\nabla \cdot \vec{u}, q) & =0 & & \forall q \in L_{0}^{2}(D) . \tag{6.5}
\end{align*}
$$

Mixed finite element approximation of (6.4)-(6.5) is obtained by choosing finite dimensional subspaces $X_{E}^{h}, X_{0}^{h}$ and $M^{h}$ of $\boldsymbol{H}_{E}^{1}, \boldsymbol{H}_{0}^{1}$ and $M^{h}$, respectively. This leads to the discrete Galerkin formulation: find $\vec{u} \in X_{E}^{h}$ and $p \in M^{h}$ such that (6.4) holds for all $v \in X_{0}^{h}$ and (6.5) holds for all $q \in M^{h}$. We use the div-stable $Q_{2}-P_{-1}$ (biquadratic velocity - linear discontinuous pressure [10]) spatial discretization and denote the dimensions of $X_{0}^{h}$ and $M^{h}$, i.e., the numbers of velocity and pressure degrees of freedom, by $N_{h, u}$ and $N_{h, p}$ respectively.

To handle the quadratic term $(\vec{u} \cdot \nabla \vec{u}, \vec{v})$ in (6.4), we use a Picard iteration as discussed in [10, pp. 324-327] and implemented in the IFISS software package [9, 27]; it is straightforward to extend the results in this section to Newton iteration. To start the Picard iteration, we can solve a discrete Stokes problem to obtain an initial guess: find $\vec{u}^{0} \in X_{E}^{h}$ and $p^{0} \in M^{h}$ such that

$$
\begin{align*}
\left(\nabla \vec{u}^{0}, \nabla \vec{v}\right)-\left(p^{0}, \nabla \cdot \vec{v}\right) & =0 & \forall \vec{v} \in X_{0}^{h},  \tag{6.6}\\
\left(\nabla \cdot \vec{u}^{0}, q\right) & =0 & \forall q \in M^{h} . \tag{6.7}
\end{align*}
$$

The Picard iteration then computes a sequence of corrections at step $n$ : find $\delta \vec{u} \in X_{0}^{h}$ and $\delta p \in M^{h}$, such that

$$
\begin{align*}
(\nu \nabla \delta \vec{u}, \nabla \vec{v}) & +\left(\vec{u}^{n} \cdot \nabla \delta \vec{u}, \vec{v}\right)-(\delta p, \nabla \cdot \vec{v}) \\
& =-\left(\nu \nabla \vec{u}^{n}, \nabla \vec{v}\right)-\left(\vec{u}^{n} \cdot \nabla \vec{u}^{n}, \vec{v}\right)+\left(p^{n}, \nabla \cdot \vec{v}\right) \quad \forall \vec{v} \in X_{0}^{h},  \tag{6.8}\\
(\nabla \cdot \delta \vec{u}, q) & =-\left(\nabla \cdot \vec{u}^{n}, q\right) \quad \forall q \in M^{h} . \tag{6.9}
\end{align*}
$$

The velocity and pressure are then updated by

$$
\vec{u}^{n+1}=\vec{u}^{n}+\delta \vec{u}, \quad p^{n+1}=p^{n}+\delta p .
$$

Since the reduced basis methods discussed in Section 3 are built on homogeneous Dirichlet conditions, some care must be taken in treatment of inhomogeneous conditions. We use an approach described in [12], which is to first find a particular function $\vec{u}_{b c}^{0}$ that satisfies the Dirichlet boundary conditions and then write

$$
\vec{u}^{0}=\vec{u}_{b c}^{0}+\vec{u}_{i n}^{0},
$$

where $\vec{u}_{\text {in }}^{0}$ satisfies homogeneous boundary conditions. We refer to $\vec{u}_{\text {in }}^{0}$ as the interior part of the Stokes solution. Then, for the initial step (6.6)-(6.7), we solve a modified Stokes problem: find $\vec{u}_{i n}^{0} \in X_{0}^{h}$ and $p^{0} \in M^{h}$ such that

$$
\begin{align*}
\left(\nabla \vec{u}_{i n}^{0}, \nabla \vec{v}\right)-\left(p^{0}, \nabla \cdot \vec{v}\right) & =-\left(\nabla \vec{u}_{b c}^{0}, \nabla \vec{v}\right) & \forall \vec{v} \in X_{0}^{h},  \tag{6.10}\\
\left(\nabla \cdot \vec{u}_{i n}^{0}, q\right) & =-\left(\nabla \cdot \vec{u}_{b c}^{0}, q\right) & \forall q \in M^{h} . \tag{6.11}
\end{align*}
$$

We have some flexibility in the choice of the particular function $\vec{u}_{b c}^{0}$. We use a simple one in which $\vec{u}_{b c}^{0}$ is the interpolant of the boundary data $\vec{g}(\cdot, \xi)$; that is, we define $\vec{u}_{i n}^{0}$ and $\vec{u}_{b c}^{0}$ as

$$
\begin{gather*}
\vec{u}_{i n}^{0}= \begin{cases}\vec{u}^{0} & \text { on spatial grid points in } D, \\
0 & \text { on spatial grid points on } \partial D,\end{cases}  \tag{6.12}\\
\vec{u}_{b c}^{0}= \begin{cases}0 & \text { on spatial grid points in } D, \\
\vec{g}(\cdot, \xi) & \text { on spatial grid points on } \partial D .\end{cases} \tag{6.13}
\end{gather*}
$$

Then for the Picard iteration step, no special treatment of boundaries is needed, since the correction function $\delta \vec{u}$ satisfies a homogeneous boundary condition.

The algebraic equations associated with (6.10)-(6.11) can be written as

$$
\left[\begin{array}{cc}
A & B^{T} \\
B & 0
\end{array}\right]\left[\begin{array}{l}
\mathbf{u}^{0} \\
\mathbf{p}^{0}
\end{array}\right]=\left[\begin{array}{l}
\mathbf{f}_{\xi} \\
\mathbf{g}_{\xi}
\end{array}\right],
$$

and for the Picard step (6.8)-(6.9), the corresponding equations are

$$
\left[\begin{array}{cc}
A_{\xi}+N_{\mathbf{u}^{n}, \xi} & B^{T} \\
B & 0
\end{array}\right]\left[\begin{array}{l}
\delta \mathbf{u} \\
\delta \mathbf{p}
\end{array}\right]=\left[\begin{array}{c}
\mathbf{f}_{\mathbf{u}^{n}, \mathbf{p}^{n}, \xi}^{r} \\
\mathbf{g}_{\mathbf{u}^{r}, \mathbf{p}^{n}, \xi}
\end{array}\right],
$$

where $\left[\mathbf{u}^{n}, \mathbf{p}^{n}\right]^{T}$ is the solution vector at the most recent iteration step, and $N_{\mathbf{u}^{n}, \xi}$ is the quadratic term considered in (3.11).
6.2. Formulation of the reduced problem. We follow the development in Section 3 to define the reduced versions of (6.10)-(6.11) and (6.8)-(6.9). To begin, we introduce reduced bases $Q_{u}:=\left\{\vec{u}_{1}, \ldots, \vec{u}_{N_{u}}\right\} \subset X_{0}^{h}$ for velocity and $Q_{p}:=\left\{q_{1}, \ldots, q_{N_{p}}\right\} \subset$ $M^{h}$ for pressure with $N_{p}<N_{u} \ll N_{h, u}$. We then seek $\vec{u}_{R}^{0} \in \operatorname{span}\left\{Q_{u}\right\}$ and $p_{R}^{0} \in$ $\operatorname{span}\left\{Q_{p}\right\}$ such that

$$
\begin{align*}
\left(\nabla \vec{u}_{R}^{0}, \nabla \vec{v}\right)-\left(p_{R}^{0}, \nabla \cdot \vec{v}\right) & =-\left(\nabla \vec{u}_{b}^{0}, \nabla \vec{v}\right) & \forall \vec{v} \in \operatorname{span}\left\{Q_{u}\right\},  \tag{6.14}\\
\left(\nabla \cdot \vec{u}_{R}^{0}, q\right) & =-\left(\nabla \cdot \vec{u}_{b c}^{0}, q\right) & \forall q \in \operatorname{span}\left\{Q_{p}\right\} . \tag{6.15}
\end{align*}
$$

A Picard iteration step entails finding $\delta \vec{u}_{R} \in \operatorname{span}\left\{Q_{u}\right\}$ and $\delta p_{R} \in \operatorname{span}\left\{Q_{p}\right\}$ such that

$$
\begin{align*}
\left(\nu \nabla \delta \vec{u}_{R}, \nabla \vec{v}\right) & +\left(\vec{u}_{R}^{n} \cdot \nabla \delta \vec{u}_{R}, \vec{v}\right)-\left(\delta p_{R}, \nabla \cdot \vec{v}\right) \\
& =-\left(v \nabla \vec{u}_{R}^{n}, \nabla \vec{v}\right)-\left(\vec{u}_{R}^{n} \cdot \nabla \vec{u}_{R}^{n}, \vec{v}\right)+\left(p_{R}^{n}, \nabla \cdot \vec{v}\right) \quad \forall \vec{v} \in \operatorname{span}\left\{Q_{u}\right\},  \tag{6.16}\\
\left(\nabla \cdot \delta \vec{u}_{R}, q\right) & =-\left(\nabla \cdot \vec{u}_{R}^{n}, q\right) \quad \forall q \in \operatorname{span}\left\{Q_{p}\right\} .
\end{align*}
$$

With $\mathbf{Q}_{u}$ and $\mathbf{Q}_{p}$ representing the matrix form of the reduced velocity and pressure bases respectively, the linear system for the reduced problem (6.14)-(6.15) is

$$
\left[\begin{array}{cc}
\mathbf{Q}_{u}^{T} A \mathbf{Q}_{u} & \mathbf{Q}_{u}^{T} B^{T} \mathbf{Q}_{p}  \tag{6.18}\\
\mathbf{Q}_{p}^{T} B \mathbf{Q}_{u} & 0
\end{array}\right]\left[\begin{array}{c}
\tilde{\mathbf{u}}^{0} \\
\tilde{\mathbf{p}}^{0}
\end{array}\right]=\left[\begin{array}{c}
\mathbf{Q}_{u}^{T} \mathbf{f}_{\xi} \\
\mathbf{Q}_{p}^{T} \mathbf{g}_{\xi}
\end{array}\right] ;
$$

and for (6.16)-(6.17),

$$
\left[\begin{array}{cc}
\mathbf{Q}_{u}^{T}\left(A_{\xi}+N_{\tilde{\mathbf{u}}^{2}, \xi}\right) \mathbf{Q}_{u} & \mathbf{Q}_{u}^{T} B^{T} \mathbf{Q}_{p}  \tag{6.19}\\
\mathbf{Q}_{p}^{T} B \mathbf{Q}_{u} & 0
\end{array}\right]\left[\begin{array}{c}
\delta \tilde{\mathbf{u}} \\
\delta \tilde{\mathbf{p}}
\end{array}\right]=\left[\begin{array}{c}
\mathbf{Q}_{u}^{T} \mathbf{f}_{\tilde{u}^{n}, \tilde{\mathbf{p}}^{n}, \xi}^{r} \\
\mathbf{Q}_{p}^{T} \mathbf{g}_{\tilde{\mathbf{u}}^{n}, \tilde{\mathbf{p}}^{n}, \xi}
\end{array}\right] .
$$

where $\left[\tilde{\mathbf{u}}^{n}, \tilde{\mathbf{p}}^{n}\right]^{T}$ is the reduced solution vector at the most recent step. The residual error indicator is taken to be the discrete nonlinear residual associated with (6.4)(6.5),

$$
\eta_{Q, \xi, n}:=\left\|\left[\begin{array}{cc}
A_{\xi}+N_{\tilde{\mathbf{u}}^{n}, \xi} & B^{T}  \tag{6.20}\\
B & 0
\end{array}\right]\left[\begin{array}{c}
Q_{u} \tilde{\mathbf{u}}^{n} \\
Q_{p} \tilde{\mathbf{p}}^{n}
\end{array}\right]\right\|_{2} /\left\|\left[\begin{array}{c}
\mathbf{f}_{\xi} \\
\mathbf{g}_{\xi}
\end{array}\right]\right\|_{2}
$$

Using the techniques introduced in Section 3.1 and 3.2, once offline reduced matrices and vectors are precomputed, the reduced linear systems (6.18) and (6.19) can be assembled with costs $O\left(N_{u}^{2}\right)$ and $O\left(N_{u}^{3}\right)$ respectively, and for modest tolerances, the residual indicator can be evaluated with a cost $\left.O\left(N_{u}^{3}\right)\right)^{3}$

The details of reduced basis collocation for the steady-state Navier-Stokes equations are presented in Algorithm 2 below. In the algorithm, $\vec{u}_{i n}\left(\cdot, \xi^{(k)}\right)$ is the interior part of $\vec{u}\left(\cdot, \xi^{(k)}\right)$ (see (6.12)). For the reduced basis, we note that the spaces generated by a set of snapshots

$$
\left\{\binom{\vec{u}_{\text {in }}\left(\cdot, \xi^{(1)}\right)}{p\left(\cdot, \xi^{(1)}\right)}, \ldots,\binom{\vec{u}_{\text {in }}\left(\cdot, \xi^{(N)}\right)}{p\left(\cdot, \xi^{(N)}\right)}\right\}
$$

do not automatically satisfy an inf-sup condition

$$
\gamma_{R}:=\min _{0 \neq q_{R} \in \operatorname{span}\left\{Q_{p}\right\}} \max _{\vec{O} \neq \vec{v}_{R} \in \operatorname{span}\left\{Q_{u}\right\}} \frac{\left(q_{R}, \nabla \cdot \vec{v}_{R}\right)}{\left|\vec{v}_{R}\right|_{1}\left\|q_{R}\right\|_{0}} \geq \gamma^{*}>0
$$

for $\gamma^{*}$ independent of $Q_{u}$ and $Q_{p}$. To ensure stability in this sense, we use an approach described in [24], which enriches the set of velocity snapshots with $\left\{\vec{r}\left(\cdot, \xi^{(k)}\right)\right\}_{k=1}^{N}$ satisfying

$$
\begin{equation*}
\left(\nabla \vec{r}\left(\cdot, \xi^{(k)}\right), \nabla \vec{v}\right)=\left(p\left(\cdot, \xi^{(k)}\right), \nabla \cdot \vec{v}\right) \quad \forall \vec{v} \in X_{0}^{h} \tag{6.21}
\end{equation*}
$$

These enriching functions are supremizers that satisfy [24]

$$
\vec{r}\left(\cdot, \xi^{(k)}\right)=\underset{\vec{v} \in X_{0}^{h}}{\arg \sup } \frac{\left(p\left(\cdot, \xi^{(k)}\right), \nabla \cdot \vec{v}\right)}{|\vec{v}|_{1}} .
$$

[^3]It follows that the reduced bases generated in Algorithm 2 are stable in the sense that

$$
\begin{equation*}
\gamma_{R} \geq \gamma_{h}:=\min _{0 \neq q \in M^{h}} \max _{\overrightarrow{0} \neq \vec{v} \in X_{0}^{h}} \frac{(q, \nabla \cdot \vec{v})}{|\vec{v}|_{1}\|q\|_{0}} . \tag{6.22}
\end{equation*}
$$

In addition, it is clear that $N_{u}=2 N_{p}$ for $Q_{u}$ and $Q_{P}$ generated in Algorithm 2.

```
Algorithm 2 Reduced basis collocation for Navier-Stokes problems
    Start with level \(M\left(\Theta_{M}=\left\{\xi^{(0)}\right\}\right)\), and compute \(\vec{u}\left(\cdot, \xi^{(0)}\right)\) and \(p\left(\cdot, \xi^{(0)}\right)\).
    Compute the supremizer function \(r\left(\cdot, \xi^{(0)}\right)\) of (6.21).
    Initialize the reduced basis \(Q_{u}=\left\{\vec{u}_{i n}\left(\cdot, \xi^{(0)}\right), \vec{r}\left(\cdot, \xi^{(0)}\right)\right\}\) and \(Q_{p}=\left\{p\left(\cdot, \xi^{(0)}\right)\right\}\).
    Construct the offline reduced matrices and vectors.
    for \(p=1: q\) do
        for \(k=1:\left|\Theta_{M+p}\right|\) do
            Compute the reduced solution and an error indicator \(\eta_{Q, \xi^{(k)}, n}\)
            if \(\eta_{Q, \xi^{(k)}, n}<t o l\) then
                Use the reduced solution to serve as \(u_{c}\left(\cdot, \xi^{(k)}\right)\) in (2.5).
            else
            - Compute the full solutions \(\vec{u}\left(\cdot, \xi^{(k)}\right)\) and \(p\left(\cdot, \xi^{(k)}\right)\).
            - Use the full solution to serve as \(u_{c}\left(\cdot, \xi^{(k)}\right)\) in (2.5).
            - Compute the supremizer function \(r\left(\cdot, \xi^{(k)}\right)\) of (6.21).
            - Augment \(Q_{u}\) with \(\left\{\vec{u}_{i n}\left(\cdot, \xi^{(k)}\right), \vec{r}\left(\cdot, \xi^{(k)}\right)\right\}\) and \(Q_{p}\) with \(\left\{p\left(\cdot, \xi^{(k)}\right)\right\}\),
                by Gram-Schmidt orthogonalization.
            - Reconstruct the offline reduced matrices and vectors.
            end if
        end for
    end for
```

6.3. Test problem 3: driven cavity flow with uncertainty in viscosity. The flow domain here is the square $D=(-1,1) \times(-1,1)$. The velocity profile

$$
\begin{equation*}
u=1-x^{4}, \quad v=0, \tag{6.23}
\end{equation*}
$$

is imposed on the top boundary $(y=1)$, and all other boundaries are no-slip and no-penetration so that $\vec{u}=(0,0)$. As in test problem 1 , we divide the square domain into $N_{D}$ subdomains and the viscosity is defined to be constant on each subdomain, $\left.v(\cdot, \xi)\right|_{D_{k}}=\xi_{k}, k=1: N_{D}$, where the random variable $\xi=\left[\xi_{1}, \ldots, \xi_{N_{D}}\right]^{T}$ is uniformly distributed in $\Gamma=[0.01,1]^{N_{D}}$. Two examples are shown in Figure 6.1. In case 1, the square domain is equally divided into two parts, and in case 2 , the domain is subdivided into an interior square centered at $(0,0)$ and two square annuli. Each of the subdomains has width 0.4 . Results for uniform $33 \times 33$ and $65 \times 65$ spatial grids are reported below.

The number of full system solves are shown in Tables 6.1 and 6.2 for domain case 1 and case 2 respectively, where two tolerance values ( $10^{-4}$ and $10^{-5}$ ) are tested.


Fig. 6.1. Domain partitionings for driven cavity flow.

Exactly as for the diffusion equations (5.4)-(5.6), we compute the mean function errors for the velocity and pressure solutions. Figures 6.2 and 6.3 show the errors, where we used the reference level $r=11$ for both types of domain.

Table 6.1
Number of full system solves for test problem 3, domain case 1.

| $q$ |  |  | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| :--- | :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| tol | Grids | $\left\|\Theta_{q}\right\|$ | 1 | 5 | 13 | 29 | 65 | 145 | 321 | 705 |
| $10^{-4}$ | $33 \times 33$ |  | 1 | 4 | 5 | 6 | 6 | 7 | 6 | 3 |
| $10^{-4}$ | $65 \times 65$ | 1 | 4 | 5 | 5 | 5 | 5 | 5 | 2 | 32 |
| $10^{-5}$ | $33 \times 33$ |  | 1 | 4 | 8 | 7 | 8 | 10 | 11 | 5 |
| $10^{-5}$ | $65 \times 65$ |  | 1 | 4 | 8 | 7 | 8 | 9 | 9 | 3 |

Table 6.2
Number of full system solves for test problem 3, domain case 2.

| $q$ |  |  | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| tol | Grids | $\left\|\Theta_{q}\right\|$ | 1 | 7 | 25 | 69 | 177 | 441 | 1073 | Total

The reduced inf-sup constants $\gamma_{R}$ for domain case 2 discretized on a $65 \times 65$ spatial grid are shown in Table. 6.3. The square of the discrete inf-sup constant for this element and mesh is known to be $\gamma_{h}^{2}=0.2137$ [10, p. 271]. It is evident from Table 6.3 that $\gamma_{R}^{2}$ is bounded below by 0.2137 , which is consistent with (6.22). As the size of the reduced basis increases, $\gamma_{R}^{2}$ becomes closer to $\gamma_{h}^{2}$.


Fig. 6.2. Mean function errors of test problem 3 with domain case $1,65 \times 65$ grid.


Fig. 6.3. Mean function errors of test problem 3 with domain case $2,65 \times 65$ grid.

Table 6.3
Inf-sup constants of reduced basis for test problem 3, domain case 2.

| $N_{u}$ | 2 | 4 | 20 | 50 | 100 | 200 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\gamma_{R}^{2}$ | 0.2431 | 0.2430 | 0.2374 | 0.2359 | 0.2327 | 0.2292 |

6.4. Test problem 4: driven cavity flow with uncertain boundary conditions. The flow domain in this section is also the square $D=(-1,1) \times(-1,1)$, and the boundary condition (6.23) is specified at the top boundary. Unlike Section 6.3 where the other boundaries are assumed to be non-slip, we now assume there is some un-


Fig. 6.4. Meshes for test problem 4.
certainty on these boundaries:

$$
\begin{array}{lll}
u=0, & v=\xi_{1}\left(1-y^{4}\right), & \text { for } x=1, \\
u=\xi_{2}\left(1-x^{4}\right), & v=0, & \text { for } y=-1, \\
u=0, & v=\xi_{3}\left(1-y^{4}\right), & \text { for } x=-1,
\end{array}
$$

where $\xi=\left[\xi_{1}, \xi_{2}, \xi_{3}\right]^{T}$ is assumed to be independently and uniformly distributed in $[-0.1,0.1]^{3}$. The viscosity is taken to be a deterministic constant, $v=1 / 500$ here. Instead of the uniform meshes used in previous examples, stretched meshes of sizes $33 \times 33$ and $65 \times 65$ are used to discretize $D$ (see Figure 6.4). ${ }^{4}$ Table 6.4 shows the number of full system solves for this test problem for a range of tolerances, and Figure 6.5 shows the error in mean functions, where the reference level is $r=11$.

Table 6.4
Number of full system solves for test problem 4.

| $q$ |  |  | 3 | 4 | 5 | 6 | 7 | 8 | 9 |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| tol | Grids | $\left\|\Theta_{q}\right\|$ | 1 | 7 | 25 | 69 | 177 | 441 | 1073 | total |
| $10^{-4}$ | $33 \times 33$ |  | 1 | 3 | 1 | 1 | 0 | 0 | 0 | 6 |
| $10^{-4}$ | $65 \times 65$ |  | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 4 |
| $10^{-5}$ | $33 \times 33$ |  | 1 | 6 | 7 | 4 | 1 | 0 | 0 | 19 |
| $10^{-5}$ | $65 \times 65$ |  | 1 | 4 | 4 | 3 | 1 | 0 | 0 | 13 |
| $10^{-6}$ | $33 \times 33$ |  | 1 | 6 | 15 | 10 | 8 | 2 | 0 | 42 |
| $10^{-6}$ | $65 \times 65$ |  | 1 | 6 | 10 | 5 | 2 | 0 | 0 | 24 |
| $10^{-7}$ | $33 \times 33$ |  | 1 | 6 | 17 | 27 | 18 | 10 | 5 | 84 |
| $10^{-7}$ | $65 \times 65$ |  | 1 | 6 | 16 | 13 | 9 | 4 | 1 | 50 |

[^4]

Fig. 6.5. Mean function errors of test problem 4, stretched $65 \times 65$ grid.

The trends for all the examples in this section (test problem 4) and the preceding one (test problem 3) are consistent with those for the diffusion equation. In particular, the number of full system solves required for reduced basis collocation does not increase as the number of the spatial degrees of freedom increases, these numbers are significantly smaller than what is needed for full collocation, and with moderate values of $\operatorname{tol}\left(10^{-4}\right.$ for problem 3 and $10^{-6}$ for problem 4) the reduced solutions are as accurate as those obtained from full collocation.
7. Concluding remarks. We conclude with a brief summary of our observations from this study. The main one, seen in all the examples considered, is that the reduced basis method can be used to significantly reduce the dimension of the discrete problems that need to be solved to construct collocation solutions of stochastic partial differential equations. Moreover, the computational results indicate that the dimensions of the reduced bases do not depend on the sizes of the discrete spatial problems that the reduced problems approximate. This suggests that the reduced dimensions depend on properties of the underlying partial differential equations and that the combined reduced basis collocation method is of potential benefit whenever spatial accuracy is of importance.

## REFERENCES

[1] M. Ainsworth, A posteriori error estimation for lowest order Raviart-Thomas mixed finite elements, SIAM J. Sci. Comput., 30 (2007), pp. 189-204.
[2] I. Babuška, F. Nobile, and R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, SIAM J. Numer. Anal., 45 (2007), pp. 1005-1034.
[3] V. Barthelmann, E. Novak, and K. Ritter, High dimensional polynomial interpolation on sparse grids, Adv. Comput. Math., 12 (2000), pp. 273-288.
[4] S. Boyaval, C. L. Bris, T. Lelièvre, Y. Maday, N. Nguyen, and A. Patera, Reduced basis techniques for stochastic problems, Arch. Comput. Methods Engrg., 17 (2010), pp. 1-20.
[5] D. Braess, Finite Elements, Cambridge University Press, London, 1997.
[6] J. L. Brown, Jr., Mean square truncation error in series expansions of random functions, J. Soc. Indust. Appl. Math., 8 (1960), pp. 28-32.
[7] S. Chaturantabut and D. Sorensen, Nonlinear model reduction via discrete empirical interpolation, SIAM J. Sci. Comput., 32 (2010), pp. 2737-2764.
[8] H. C. Elman, C. W. Miller, E. T. Phipps, and R. S. Tuminaro, Assessment of collocation and Galerkin approaches to linear diffusion equations with random data, Int. Journal for Uncertainty Quantification, 1 (2011), pp. 19-34.
[9] H. C. Elman, A. Ramage, and D. J. Silvester, Algorithm 866: IFiSS, a Matlab toolbox for modelling incompressible flow, ACM Trans. Math. Soft., 33 (2007), pp. 14:1-14:18.
[10] H. C. Elman, D. J. Silvester, and A. J. Wathen, Finite Elements and Fast Iterative Solvers, Oxford University Press, New York, 2005.
[11] R. Ghanem and P. Spanos, Stochastic Finite Elements: A Spectral Approach, Dover Publications, New York, 2003.
[12] M. Gunzburger, J. Peterson, and J. Shadid, Reduced-order modeling of time-dependent PDEs with multiple parameters in the boundary data, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 1030-1047.
[13] B. Hassdonk and M. Ohlberger, Reduced basis method for finite volume approximations of parametrized linear evolution equations, M2AN, 42 (2008), pp. 277-301.
[14] D. Kay and D. Silvester, A posteriori error estimation for stabilized mixed approximations of the Stokes equations, SIAM J. Sci. Comput., 21 (1999), pp. 1321-1336.
[15] J. Kim, Phase field computations for ternary fluid flows, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 4779-4788.
[16] A. Klimke, Sparse Grid Interpolation Toolbox - User's Guide, Tech. Rep. IANS report 2007/017, University of Stuttgart, 2007.
[17] Q. Liao and D. Silvester, A simple yet effective a posteriori estimator for classical mixed approximation of Stokes equations, Appl. Numer. Math., 62 (2012), pp. 1242-1256.
[18] X. Ma and N. Zabaras, An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations, J. Comput. Phys., 228 (2009), pp. 3084-3113.
[19] _-, High-dimensional stochastic model representation technique for the solution of stochastic pdes, J. Comput. Phys., 229 (2010), pp. 3884-3915.
[20] N. C. Nguyen, K. Veroy, and A. T. Patera, Certified real-time solution of parametrized partial differential equations, in Handbook of Materials Modeling, S. Yip, ed., Kluwer Academic Publishing: Dordrecht, 2005, pp. 1523-1558.
[21] F. Nobile, R. Tempone, and C. G. Webster, A sparse grid stochastic collocation method for partial differential equations with random input data, SIAM J. Numer. Anal., 46 (2008), pp. 2309-2345.
[22] E. Novak and K. Ritter, High dimensional integration of smooth functions over cubes, Numer. Math., 75 (1996), pp. 79-97.
[23] M. A. Olshanskii and A. Reusken, Analysis of a Stokes interface problem, Numer. Math., 103 (2006), pp. 129-149.
[24] A. Quarteroni and G. Rozza, Numerical solution of parametrized Navier-Stokes equations by reduced basis methods, Numer. Meth. Part. Diff. Equations, 23 (2007), pp. 923-948.
[25] G. Rozza, D. Huynh, and A. Patera, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics, Arch. Comput. Methods Eng., 15 (2008), pp. 229-275.
[26] C. Schwab and R. A. Todor, Karhunen-Loève approximation of random fields by generalized fast multipole methods, J. Comput. Phys., 217 (2006), pp. 100-122.
[27] D. J. Silvester, H. C. Elman, and A. Ramage, Incompressible Flow and Iterative Solver Software (IFISS) version 3.2, May 2012. http://www.manchester.ac.uk/ifiss/.
[28] Z. Tan, D. V. Le, Z. Li, K. M. Lim, and B. C. Khoo, An immersed interface method for solving incompressible viscous flows with piecewise constant viscosity across a moving elastic membrane, J. Comput. Phys., 227 (2008), pp. 9955-9983.
[29] R. Verfürth, A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Tech-
niques, Wiley-Teubner, Chichester, 1996.
[30] K. Veroy and A. Patera, Certified real-time solution of the parametrized steady incompressible Navier-Stokes equations: rigorous reduced-basis a posteriori error bounds, Int. J. Numer. Meth. Fluids, 47 (2005), pp. 773-788.
[31] K. Veroy, A. Patera, and D. V. Rovas, A posteriori error estimation for reduced-basis approximation of parameterized elliptic partial differential equations: "convex inverse" bound conditioners, Control, Optim. Calculus Var., 8 (2002), pp. 1007-1028.
[32] D. Xiu and J. Hesthaven, High-order collocation methods for differential equations with random inputs, SIAM J. Sci. Comput., 27 (2005), pp. 1118-1139.


[^0]:    *This work was supported in part by the U. S. Department of Energy under grant DEFG0204ER25619 and by the U. S. National Science Foundation under grant DMS1115317.
    ${ }^{\dagger}$ Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742, USA. (Email: elman@cs.umd.edu).
    ${ }^{\ddagger}$ Department of Computer Science, University of Maryland, College Park, MD 20742, USA. (Email: qliao@umd.edu).

[^1]:    ${ }^{1}$ We refer to the MATLAB toolbox SPINTERP [16] for evaluating $L_{\boldsymbol{\xi}}{ }^{(k)}(\xi)$.

[^2]:    ${ }^{2}$ We have repeated this test more than ten times for different random sets, and no significantly different results were found.

[^3]:    ${ }^{3}$ In the experiments described below, the error indicator was computed directly at cost $O\left(N_{h}\right)$.

[^4]:    ${ }^{4}$ These are generated by IFISS [27] using the default setting for mesh generation.

