PARALLEL DOMAIN DECOMPOSITION METHODS FOR STOCHASTIC ELLIPTIC EQUATIONS[*]

CHAO JIN[†], XIAO-CHUAN CAI[‡], AND CONGMING LI[§]

Abstract. We present parallel Schwarz type domain decomposition preconditioned recycling Krylov subspace methods for the numerical solution of stochastic elliptic problems, whose coefficients are assumed to be a random field with finite variance. Karhunen-Loève (KL) expansion and double orthogonal polynomials are used to reformulate the stochastic elliptic problem into a large number of related, but uncoupled deterministic equations. The key to an efficient algorithm lies in "recycling computed subspaces". Based on a careful analysis of the KL expansion we propose and test a grouping algorithm that tells us when to recycle and when to recompute some components of the expensive computation. We show theoretically and experimentally that the Schwarz preconditioned recycling GMRES method is optimal for the entire family of linear systems. A fully parallel implementation is provided and scalability results are reported in the paper.

Key words. Stochastic elliptic equations, domain decomposition, recycling Krylov subspace method, parallel scalability

AMS subject classifications. 35R60, 60H15, 60H35, 65C30, 47B80, 65N55, 65M55, 65M60, 65M70

1. Introduction. Many physical phenomena are modeled by partial differential equations whose coefficients are measured through experiments which often contain some levels of uncertainty. To understand the model accurately, the randomness must be considered in the differential equations. In the last decade, many researchers have studied the so-called stochastic partial differential equation (SPDE). One of the approaches for numerically solving SPDEs is known as the stochastic Galerkin method, see for example [1, 2, 11, 17, 22, 25, 30] and references therein. The method expands the random field in the equation. Several such expansions are available, for example, [1, 2, 9, 13, 20] use the Karhunen-Loève expansion introduced in [21], [17, 29, 30] use the Wiener Chaos expansion [31], and [29] uses the generalized Chaos expansion. With the expansion, the SPDE is reduced to a high dimensional deterministic equation, which can be solved by different numerical methods. Based on the solution, the statistics of the physical solution of the original stochastic partial differential equation, such as the mean and the variance, can be derived.

In this paper, we focus on elliptic problems with stochastic diffusion coefficients. For this type of SPDEs, we use the double orthogonal basis [3, 16] to decouple the high dimensional equation in the probability space and produce a sequence of independent equations. After the discretization, we obtain a large number of linear systems:

$$(1.1) \qquad A_i x_i = b_i, \quad i = 1, 2, \cdots$$

where the matrices A_i and right-hand sides b_i are closely related but independent from each other. There are a number of techniques for solving a sequence of related linear

[†]Department of Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309 (chao.jin@colorado.edu)

[‡]Department of Computer Science, University of Colorado at Boulder, Boulder, CO 80309 (cai@cs.colorado.edu)

[§]Department of Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309 (cli@colorado.edu)

systems of equations. In this paper, we focus on the recently introduced recycling Krylov subspace method [24], which retains a Krylov subspace generated while solving a previous system with GMRES and uses it to reduce the cost of solving the current system. In our implementation, we extend the original algorithm in [24] to the flexible version of GMRES, which allows the change of preconditioner during the iteration.

Mathematically we prove, under certain assumptions, that all these linear systems from the KL expansion are spectrally equivalent, however, we also show numerically that it is generally not a good idea to recycle or reuse the same Krylov subspace and/or preconditioner for the entire sequence of linear systems since some of the spectrally equivalent systems are quite different. Based on some analysis of the KL expansion and the lower bounds of the elliptic coefficients we propose a grouping algorithm which partitions the large number of systems into several groups and our tests show that within each group, recycling Krylov subspaces and recycling preconditioners can be safely and efficiently applied.

For preconditioning, we use an overlapping additive Schwarz domain decomposition method [27]. Our parallel implementation is based on the Portable Extensible Toolkit for Scientific computation (PETSc) package from Argonne National Laboratory [4]. The scalability and parallel performance of the algorithm are studied theoretically and experimentally.

The rest of the paper is organized as follows. In Section 2, we describe the stochastic Galerkin method including the stochastic weak formulation, the Karhunen-Loève expansion, the double orthogonal basis, and the discretization. Section 3 presents the additive Schwarz preconditioned recycling Krylov subspace method. A grouping algorithm for efficiently recycling the Krylov subspace and the preconditioner is also introduced. Some experimental results are reported in Section 4 and some concluding remarks are given in Section 5.

2. Stochastic Galerkin method. In this section, we briefly describe how to transform a stochastic elliptic partial differential equation to a sequence of independent deterministic partial differential equations. We use an elliptic equation with a stochastic diffusion coefficient as a model problem. See [16] for a complete description of the methodology.

2.1. The stochastic weak formulation. We begin with a brief review of notations. Given a probability space $(\Omega, \mathcal{A}, \mathcal{P})$ with sample space Ω, σ-algebra $\mathcal{A}$ and probability measure $\mathcal{P}$, a real-valued random variable is a function $\xi(\omega)\colon \Omega \to \mathbf{R}$. The probability distribution measure of ξ is defined on the Borel set $\mathcal{B}$ as $\mu(\mathcal{B}) = \mathcal{P}(\xi^{-\infty}(\mathcal{B}))$. The mean, or expected value of $\xi(\omega)$, is

$$(2.1) \qquad \langle \xi \rangle = \int_\Omega \xi(\omega) d\mathcal{P}(\omega) = \int_{\mathbf{R}} z d\mu(z) = \int_{\mathbf{R}} z\rho(z) dz,$$

where ρ is the probability density function of ξ. We also define the space

$$L^2(\Omega) = \left\{ \xi(\omega) \left| \int_\Omega |\xi|^2 d\mathcal{P}(\omega) < \infty \right. \right\}.$$

Let $D \subset R^2$ be the domain of the variable x. A random field $a(x, \omega)\colon D \times \Omega \to \mathbf{R}$ is a real-valued function jointly measurable with respect to the Lebesgue measure on D and the probability measure $\mathcal{P}$ on Ω. Define the space

$$L^2(D \times \Omega) = \{u(x, \omega)| \ \langle \|u(x, \omega)\|_{L^2(D)} \rangle < \infty\}.$$

The stochastic Sobolev space $H_0^1(D \times \Omega)$ is defined analogously. Now recall that for the classical deterministic 2D elliptic equation:

$$(2.2) \qquad \begin{cases} -\nabla \cdot (a(x)\nabla u(x)) &=& f(x) & \text{in } D \\ u(x) &=& 0 & \text{on } \partial D, \end{cases}$$

the weak form of the problem is to find $u(x) \in H_0^1(D)$ such that

$$(2.3) \qquad B[u, v] = (f, v) \quad \forall v \in H_0^1(D),$$

where

$$B[u, v] = \int_D a(x)\nabla u(x) \cdot \nabla v(x)dx, \quad (f, v) = \int_D f(x)v(x)dx.$$

If the diffusion coefficient is a random field $a(x, \omega) \in L^2(D \times \Omega)$, so is the solution u. Consequently, we have the stochastic elliptic equation

$$(2.4) \qquad \begin{cases} -\nabla \cdot (a(x, \omega)\nabla u(x, \omega)) &=& f(x) & x \in D, \ \omega \in \Omega \\ u(x, \omega) &=& 0 & x \in \partial D, \ \omega \in \Omega. \end{cases}$$

The weak form of (2.4) is to find $u(x, \omega) \in H_0^1(D \times \Omega)$ such that

$$(2.5) \qquad \langle B[u, v] \rangle = \langle (f, v) \rangle \quad \forall v \in H_0^1(D \times \Omega).$$

We assume that $a(x, \omega) \in L^\infty(D \times \Omega)$ is strictly positive, with lower and upper bounds α and β respectively,

$$(2.6) \qquad 0 < \alpha \le a(x, \omega) \le \beta.$$

Under this assumption, the existence and uniqueness of a solution u to (2.4) follow from the Lax-Milgram lemma. Note that we assume the source term f is deterministic. This condition can be relaxed to include randomness.

2.2. The Karhunen-Loève expansion. Here, we use the Karhunen-Loève (KL) expansion [21] of the random field $a(x, \omega)$ to separate the deterministic and stochastic components. We assume the mean and the covariance of $a(x, \omega)$ are known respectively as

$$a_0(x) = \int_\Omega a(x, \omega)dP(\omega) \text{ and } C_a(x, x') = \int_\Omega (a(x, \omega) - a_0(x))(a(x', \omega) - a_0(x'))dP(\omega).$$
(2.7)
By the KL expansion, $a(x, \omega)$ can be represented in the form of a series as follows

$$(2.8) \qquad a(x, \omega) = a_0(x) + \sum_{j=1}^{\infty} \sqrt{\lambda_j} \, k_j(x) \, y_j(\omega),$$

where λ_j and $k_j(x)$ are the eigenvalues and orthogonal eigenfunctions of $C_a(x, x')$; i.e.,

$$(2.9) \qquad \int_D C_a(x, x')k_j(x')d(x') = \lambda_j k_j(x).$$

This series converges in the mean-square sense.

By definition, $C_a(x, x')$ is symmetric and positive semidefinite. This implies that there exists a countable sequence of eigenpairs $\{(\lambda_j, k_j)\}$ where the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq \cdots \rightarrow 0$ are nonnegative and the eigenfunctions $\{k_j(x)\}$ are orthogonal in $L^2(D)$. Moreover, $\{y_j\}$ is a set of uncorrelated random variables with mean value zero. If the eigenfunctions are normalized, y_j all have unit variance; i.e., $\langle y_j(\omega) \rangle = 0$, $\langle y_i(\omega) y_j(\omega) \rangle = \delta_{ij}$. For the computation, we approximate $a(x, \omega)$ by a truncation of (2.8),

$$a_M(x, \omega) = a_0(x) + \sum_{j=1}^{M} \sqrt{\lambda_j}\, k_j(x)\, y_j(\omega),$$

where M denotes the number of terms in the truncation.

In this paper, we also assume the $\{y_j\}_{j=1}^{M}$ are independent. The probability density function of y_j is denoted as ρ_j. The joint probability density function of $y = (y_1, \cdots, y_M)$ is $\rho = \rho_1 \times \cdots \times \rho_M$. Let Γ_j denote the image of y_j and $\Gamma = \Gamma_1 \times \cdots \times \Gamma_M$. We treat $a_M \colon D \times \Gamma \to R$ as

$$(2.10) \qquad a_M(x, y) = a_0(x) + \sum_{j=1}^{M} \sqrt{\lambda_j}\, k_j(x)\, y_j(\omega).$$

Now, we arrive at the following deterministic elliptic problem,

$$(2.11) \quad -\nabla(a_M(x, y)\nabla u_M(x, y)) = f(x), \qquad u_M(x, y) \text{ in } H_0^1(D) \times L^2(\Gamma, \rho).$$

2.3. Double orthogonal basis. For $y \in \Gamma$, we use the double orthogonal polynomial function space [3, 16] to approximate $L^2(\Gamma, \rho)$, which decouples the equation in the y-space, yielding a sequence of uncoupled equations. The double orthogonal basis is constructed as follows. For any $r \in \mathbf{N}$, the space of single-variable polynomials of degree at most r is

$$(2.12) \qquad P_r := span\{1, t, t^2, \cdots, t^r\}.$$

For $\mathbf{r} = (r_1, r_2, \cdots, r_M) \in \mathbf{N}^M$, we construct the multi-variable polynomial space

$$(2.13) \qquad P_{\mathbf{r}} := P_{r_1} \otimes P_{r_2} \otimes \cdots \otimes P_{r_M} \in L^2(\Gamma, \rho).$$

For the space $P_{r_j}, j = 1, 2, \cdots, M$, we use the double orthogonal functions, denoted as $\{\psi_{k,j}(t)\}_{k=0}^{r_j}$, as basis instead of $\{1, t, t^2, \cdots, t^{r_j}\}$. We require that $\psi_{k,j}(t), k = 0, \cdots, r_j$ satisfy two orthogonality conditions:

$$(2.14) \quad \begin{cases} \displaystyle\int_{\Gamma_j} \psi_{p,j}(t)\psi_{q,j}(t)\rho_j(t)dt = \delta_{p,q}, & p, q = 0, \cdots, r_j, \\[2mm] \displaystyle\int_{\Gamma_j} t\psi_{p,j}(t)\psi_{q,j}(t)\rho_j(t)dt = C_{p,j}\delta_{p,q}, & p, q = 0, \cdots, r_j, \end{cases}$$

where $\{C_{p,j}\}_{p=0}^{r_j}$ are nonzero constants. Next we construct a basis function of $P_{\mathbf{r}}$ by selecting one polynomial basis function from each $P_{r_j}, j = 1, \cdots, M$, and then multiply these selected M basis functions together. So given a $\mathbf{r} = (r_1, r_2, \cdots, r_m) \in \mathbf{N}^M$, there are total $N_y = \prod_{j=1}^{M}(r_j + 1)$ basis functions for $P_{\mathbf{r}}(y_1, y_2, \cdots, y_M)$.

Let $\mathbf{i} = \{i_1, i_2, \cdots, i_M\}$. If $0 \leq i_j \leq r_j, \forall 1 \leq j \leq M$, we say that $\mathbf{i} \leq \mathbf{r}$. It is obvious that there are N_y multi-indices $\mathbf{i}$, which are less than or equal to $\mathbf{r}$. Each $\mathbf{i}$

corresponds to one basis function of $P_{\mathbf{r}}$. We denote all the basis functions of $P_{\mathbf{r}}$ as the set

$$(2.15) \qquad \left\{ \psi_{\mathbf{i}}(y) \,\middle|\, \psi_{\mathbf{i}}(y) = \prod_{j=1}^{m} \psi_{i_j,j}(y_j), \ \ i_j \in \{0,1,\cdots,r_j\} \right\}_{\mathbf{i} \leq \mathbf{r}}.$$

Finding $\{\psi_{k,j}(y_j)\}_{k=0}^{r_j}$ for spaces P_{r_j}, $j = 1,\cdots,M$, results in an eigenproblem (c. f. section 8.7.2 in [18]). For the probability space of y, generally we do not need high order polynomials. So the computational work for these eigenproblems is negligible compared with the cost required to solve the coupled equations.

2.4. Discretization. We use a Galerkin method with a double orthogonal basis for the discretization of (2.11) in the y-space. Let

$$u_M(x,y) = \sum_{\mathbf{i} \leq \mathbf{r}} u_{\mathbf{i}}(x)\psi_{\mathbf{i}}(y)$$

such that for any $v(x,y) = h(x)\psi_{\mathbf{j}}(y) \in H_0^1(D) \times P_{\mathbf{r}}(\Gamma)$, we have

$$(2.16) \qquad \left\langle \int_D a_M(x,y)\nabla u_M(x,y)\cdot\nabla v(x,y)dx \right\rangle = \left\langle \int_D f(x)v(x,y)dx \right\rangle.$$

Writing in an explicit form,

$$\left\langle \int_D a_M(x,y)\nabla u_M(x,y)\cdot\nabla v(x,y)dx \right\rangle$$
$$= \sum_{j=1}^{M}\sum_{\mathbf{i}\leq\mathbf{r}} \sqrt{\lambda_j}\,(k_j(x)\nabla u_{\mathbf{i}}(x),\ \nabla h(x)) \int_\Gamma y_j\psi_{\mathbf{i}}(y)\psi_{\mathbf{j}}(y)\rho(y)dy.$$

The two orthogonality conditions (2.14) imply that

$$\int_\Gamma y_j\psi_{\mathbf{i}}(y)\psi_{\mathbf{j}}(y)\rho(y)dy = C_{i_j,j}\delta_{\mathbf{i},\,\mathbf{j}}.$$

Consequently,

$$\left\langle \int_D a_M(x,y)\nabla u_M(x,y)\cdot\nabla v(x,y)dx \right\rangle$$
$$= \sum_{j=1}^{M}\sum_{\mathbf{i}\leq\mathbf{r}} \sqrt{\lambda_j}\,(k_j(x)\nabla u_{\mathbf{i}}(x),\ \nabla h(x))\,C_{i_j,j}\,\delta_{\mathbf{i},\mathbf{j}}.$$

On the other hand,

$$\left\langle \int_D f(x)v(x,y)dx \right\rangle = \int_D f(x)h(x)dx\langle\psi_{\mathbf{j}}(y)\rangle.$$

Now, it is easy to see that the variational form of (2.11)

$$\langle B[u_M,v]\rangle = \langle\langle f,v\rangle\rangle,$$

implies that

(2.17) $-\nabla(a_{M,\mathbf{i}}(x)\nabla u_{M,\mathbf{i}}(x)) = f_{\mathbf{i}}(x)$ in $H_0^1(D),$ $with$

(2.18)
$$
\begin{cases}
a_{M,\mathbf{i}}(x) := a_0(x) + \sum_{j=1}^{M} \sqrt{\lambda_j}\, k_j(x) C_{i_j,j}, \\
f_{\mathbf{i}}(x) := f(x) \cdot \prod_{j=1}^{M} \int_{\Gamma_j} \psi_{i_j,j}(t)\rho_j(t)dt.
\end{cases}
$$

Thus, we have decoupled the equation (2.11) into N_y deterministic diffusion problems (2.17) and (2.18) in D. The stochastic solution is

(2.19) $$u_M(x,y) = \sum_{\mathbf{i}\le\mathbf{r}} u_{M,\mathbf{i}}(x)\psi_{\mathbf{i}}(y).$$

We discretize the left-hand side of (2.17) and obtain a sequence of matrices $A_i, i = 1,\cdots,N_y$. It can be shown that these systems are positive definite, if M is large enough. The following result is useful.

LEMMA 2.1. *If C_a is piecewise smooth on $D \times D$ and $y_j, j = 1,\cdots,M$, are all bounded, then for any $\mathbf{i}$,*

$$\frac{\alpha}{2} \le a_{M,\mathbf{i}}(x) \le 2\beta$$

if M is large enough. Constants α and β are defined in (2.6).

Proof. By Proposition 4.3 in [16]: $\alpha/2 \le a_M(x,y) \le 2\beta$ for any x,y. The constant $C_{i_j,j}$ in (2.18) and (2.14) is bounded by $|y_j|$ in (2.10):

$$|C_{i_j,j}| \le \int_{\Gamma_j} |y_j|\psi_{i_j\,j}(y_j)^2 dy_j \le \max|y_j| \int_{\Gamma_j} \psi_{i_j,j}(y_j)^2 dy_j = \max|y_j|.$$

This implies that $a_{M,\mathbf{i}}$ are also bounded in $[\alpha/2, 2\beta]$. $\square$

The statistics of the solution can be found from the approximate solutions. For example, the mean of $u(x,y)$ can be approximated by,

$$\langle u_M \rangle = \sum_{\mathbf{i}\le\mathbf{r}} u_{M,\mathbf{i}}(x) \int_{\Gamma} \psi_{\mathbf{i}}(y)\rho(y)dy = \sum_{\mathbf{i}\le\mathbf{r}} u_{M,\mathbf{i}}(x) \prod_{j=1}^{M} \int_{\Gamma_j} \psi_{i_j,j}(t)\rho_j(t)dt.$$

3. Numerical method. After the discretization of (2.18) we obtain a sequence of independent linear systems of the form (1.1). In this section we propose an additive Schwarz preconditioned recycling Krylov subspace method for solving these systems. The recycling Krylov subspace method, Generalized Conjugate Residual with implicit inner Orthogonalization and Deflated Restarting (GCRO-DR) [24], reduces some work in constructing the Krylov subspace for a new system and the additive Schwarz algorithm [27] provides a preconditioner and a data structure for a parallel implementation. Note that we also reuse the preconditioner to further save the computational time.

3.1. Recycling Krylov subspace method. There are several recycling Krylov subspace methods, see for example [8, 14, 15, 19, 23, 26] and references therein. In our case which has a sequence of linear systems with changing left-hand side matrices and right-hand side vectors, we use GCRO-DR. Here, we only present the main steps of this method. For a detailed description and derivation of the algorithm, we refer to [24].

The idea of GCRO-DR is to retain a Krylov subspace for subsequent restarted GMRES cycles, or for solving other linear systems. For generality, we extend the GMRES framework introduced in [24] to include the flexible version of GMRES (FGMRES). Suppose we have solved the ith system with FGMRES. We retain k vectors

$$(3.1) \qquad \tilde{Y}_k = [\tilde{y}_1, \tilde{y}_2, \cdots, \tilde{y}_k].$$

Let $C_k = Q$, $U_k = \tilde{Y}_k R^{-1}$ where Q and R be from the reduced QR decomposition of $A_{i+1}\tilde{Y}_k$. $U_k, C_k \in R^{n \times k}$ satisfy

$$(3.2) \qquad A_{i+1}U_k = C_k, \quad C_k^H C_k = I_k.$$

Let x_0 and r_0 be the initial guess and initial residual of $A_{i+1}x_{i+1} = b_{i+1}$. We update the solution as $x = x_0 + U_k C_k^H r_0$, and set $r = r_0 - C_k C_k^H r_0$. In our numerical experiment, we find that for some systems, this updated solution x already satisfies the error tolerance so that we do not need any more iterations after the initial step.

However, if the x is still not good enough, we continue to generate a Krylov subspace of dimension $m - k + 1$ with $(I - C_k C_k^H)A_{i+1}$, where m is the maximum number of iterations before restarting. In the case of recycling FGMRES, it produces the Arnoldi relation

$$(3.3) \qquad (I - C_k C_k^H)A_{i+1}Z_{m-k} = V_{m-k+1}\bar{H}_{m-k},$$

where the columns of Z_{m-k} are the preconditioned orthogonal vectors V_{m-k}. In the case of GMRES, Z_{m-k} should be replaced by V_{m-k} everywhere in the algorithm.

Each of the Arnoldi vectors $V_{m-k+1} = [v_1, v_2, \cdots, v_{m-k+1}]$ is orthogonal to the range of C_k. We can rewrite (3.3) as

$$(3.4) \qquad A_{i+1}[U_k \ Z_{m-k}] = [C_k \ V_{m-k+1}] \begin{bmatrix} I_k & B_{m-k} \\ 0 & \bar{H}_{m-k} \end{bmatrix},$$

where $B_{m-k} = C_k^H A_{i+1}Z_{m-k}$. For numerical reasons, we normalize the column vectors of U_k and replace the identity matrix I_k above with a diagonal matrix D_k, such that $U_k D_k$ has unit columns denoted as $\tilde{U}_k$. We define

$$(3.5) \qquad \hat{V}_m = [\tilde{U}_k \ Z_{m-k}], \ \hat{W}_{m+1} = [C_k \ V_{m-k+1}], \ \bar{G}_m = \begin{bmatrix} D_k & B_{m-k} \\ 0 & \bar{H}_{m-k} \end{bmatrix}$$

and write (3.4) more compactly as

$$A_{i+1}\hat{V}_m = \hat{W}_{m+1}\bar{G}_m.$$

Note that $\bar{G}_m = \hat{W}_{m+1}^H A_{i+1}\hat{V}_m$ is upper Hessenberg with diagonal D_k. The columns of $\hat{W}_{m+1}$ are orthogonal, but this is not true for the columns of $\hat{V}_m$.

In each cycle, we need to find y to minimize

$$\begin{aligned} \|r - A_{i+1}\hat{V}_m y\|_2 &= \|r - \hat{W}_{m+1}\bar{G}_m y\|_2 \\ &= \|\hat{W}_{m+1}^H r - \bar{G}_m y\|_2 \\ &= \|e_{k+1}\|r\|_2 - \bar{G}_m y\|_2. \end{aligned}$$

Once we have y, the residual and solution are updated by

$$(3.6) \qquad r := r - A_{i+1}\hat{V}_m y = r - \hat{W}_{m+1}\bar{G}_m y \quad \text{and} \quad x := x + \hat{V}_m y.$$

Next, we solve the generalized eigenvalue problem

$$(3.7) \qquad \bar{G}_m^H \bar{G}_m z_i = \theta_i \bar{G}_m^H \hat{W}_{m+1}\hat{V}_m z_i.$$

The k vectors $\tilde{Y}_k$ are chosen from $\tilde{y}_i = \hat{V}_m z_i, i = 1, 2, \cdots, m$. Theoretically, any k eigenvectors z_i can be used to compute $\tilde{Y}_k$. In our implementation, we choose the eigenvectors corresponding to the eigenvalues of smallest magnitude [24]. For more details in computing the recycling Krylov subspace between cycles and systems, we refer to [24].

3.2. Additive Schwarz preconditioning. In this section, we show that all matrices in the sequence of linear systems are spectrally equivalent, and therefore an optimal preconditioner for anyone of the matrices is also optimal for the whole sequence. Here "optimal" means that the number of iterations is independent of the mesh size and the number of subdomains. Note that "optimal" does not mean the number of iterations is small or the computing time is small, which is of course more important. In the next section, we will introduce a method that partitions the sequence of linear systems into groups and within each group we apply the preconditioning techniques to be discussed briefly below.

We start with a coarse triangulation $\mathcal{T}^H = \{\tau_i^H\}$ and a fine triangulation $\mathcal{T}^h = \{\tau_i\}$ of the domain D. Let $\mathcal{N}(\mathcal{T}^H)$ be the set of nodal points of $\mathcal{T}^H$. Let $h_i =$ diameter of τ_i, and $h = \max_i\{h_i\}$. We assume that the coarse and fine triangulation are shape regular in the sense common to finite element theory. We then partition D into nonoverlapping subdomains $\{D_i\}$, $i = 1, \cdots, N$. Let $H_i =$ diameter of D_i, $H = \max_i\{H_i\}$, We extend each D_i to a larger region $D_i' \subset D$. We assume that the distance between the boundaries ∂D_i and $\partial D_i'$ inside D is bounded from below by a fixed fraction of H_i, and that $\partial D_i'$ does not cut through any elements.

Let V^h and V_0 be the finite element spaces of continuous, piecewise linear functions associated with the triangulations $\mathcal{T}^h$ and $\mathcal{T}^H$ respectively. Let $V_i = \{v \in V^h : v(x) = 0, x \notin D_i'\}, i \geq 1$. Our finite element space V^h can be represented as:

$$V^h = I_H^h V_0 + V_1 + \cdots + V_N,$$

where I_H^h is a coarse to fine mesh interpolation operator [7]. We approximate $H_0^1(D)$ by V^h and define the finite element problem of (2.17) as: Find $u_h \in V^h$ such that

$$(3.8) \qquad \int_D a_{M,\mathbf{i}}(x)\nabla u_h \nabla v_h dx = \int_D f_\mathbf{i} v_h dx, \quad \forall\, v \in V^h.$$

For simplicity, we denote the $a_{M,\mathbf{i}}(x)$ and $f_\mathbf{i}$ for system i, where $i = 1, 2, \cdots, N_y$, as $a_i(x)$ and $f_i(x)$ respectively. $a_0(x)$ is still the mean of $a(x)$. We define the bilinear form, for any $u, v \in H_0^1(D)$,

$$A(u, v) = \int_D a_1(x)\nabla u \nabla v dx,$$

$$B_i(u, v) = \int_D a_i(x)\nabla u \nabla v dx, \quad i = 1, 2, \cdots, N_y.$$

We now define the projection operators, which are the main building blocks of the additive Schwarz methods.

DEFINITION 3.1. P_0: $V^h \longrightarrow V_0$,

$$B_i(P_0 u, v) = B_i(u, v), \quad \forall v \in V_0.$$

For $i = 1, 2, \cdots, N$, P_i: $V^h \longrightarrow V_i$,

$$A(P_i u, v) = B_i(u, v), \quad \forall v \in V_i.$$

Note that for subdomain problems, we define the preconditioners using a bilinear form defined by a system usually different from the one being solved. But for the coarse problem, the preconditioner is defined by using the bilinear form of the same linear system. If we view the additive Schwarz method as a preconditioner, the matrix form of P_i is,

(3.9) $$P_i = R^T K_i^{-1} R K,$$

where K_i is the stiffness matrix from system 1 associated with subdomain D_i' and K is the stiff matrix from system i, $i = 1, 2, \cdots, N_y$, associated with D. R is a restriction matrix. The additive Schwarz operator is defined by

(3.10) $$P = \sum_{i=0}^{N} P_i.$$

Using the standard techniques ([27, 28]), we show easily

THEOREM 3.2. *If the coarse mesh is sufficiently fine, then there exist positive constants c_p and C_p independent of h and H, such that for any $u \in V^h$,*

(3.11) $$c_p A(u, u) \leq A(u, Pu), \quad A(Pu, Pu) \leq C_p A(u, u).$$

According to [5, 6, 7, 10], the convergence rate of the additive Schwarz preconditioned GMRES is bounded by the ratio C_p/c_p, and the theorem implies that the convergence rate is independent of the mesh size h and the number of subdomains N.

3.3. Some analysis of the sequence of linear systems and a grouping algorithm. In this section we provide some analysis of the linear systems. Although mathematically we can prove that all of the matrices are spectrally equivalent, in practice, some of the systems are quite different from the others and the same preconditioner may not work well for all the systems. Similarly, simple recycling of the Krylov subspace from the previous system may not work well for these "bad" systems. Therefore it is very important to identify these bad systems and exclude them from the rest of the systems that can be solved with the recycling of Krylov subspaces and preconditioners.

It is easy to see that the matrices depend on the diffusion coefficients $a_{M,\mathbf{i}}(x), \mathbf{i} \leq \mathbf{r}$. From (2.18), it is obvious that the differences among these coefficients $a_{M,\mathbf{i}}(x)$ are determined by the eigenvalues, the eigenfunctions, and the constants $C_{i_j,j}$. If any of them change drastically, the diffusion coefficient changes drastically.

We note that the eigenfunctions are normalized in (2.18) and the eigenvalues $\lambda_j, j = 1, 2, \cdots, M$, are not increasing and converge to zero as j goes to infinity. Therefore, the changes of the constants $C_{i_j,j}, j = 1, 2, \cdots, M$, play the main role in the perturbation of $a_{M,\mathbf{i}}$ for the terms corresponding to larger values of λ_j; i.e. when

j is small. Based on the eigenvalues and the constants $C_{i_j,j}$, we provide an estimate of when the diffusion coefficients $a_{M,\mathbf{i}}$ have a relatively large change. If the constants $C_{i_j,j}$, corresponding to large eigenvalues, change drastically, the corresponding systems may become very different from the previous systems. In this case the recycled Krylov subspace and the preconditioner should not be used.

Following the index of the diffusion coefficient $a_{M,\mathbf{i}}(x)$, the sequence of systems are currently labeled by the M-dimensional index $\mathbf{i}$ which is not too convenient for the purpose of putting the systems into different groups. We next turn the M-dimensional index $\mathbf{i}$ into a one-dimensional array. Given $\mathbf{r} = \{r_1, r_2, \cdots, r_M\}$, the index set $\{\mathbf{i} = \{i_1, i_2, \cdots, i_M\} \leq \mathbf{r}\}$ is ordered as follows:

Re-ordering Scheme for the Multi-index i:
$$k = 1;$$
$$\text{for } i_1 = 0, 1, \cdots, r_1$$
$$\quad \text{for } i_2 = 0, 1, \cdots, r_2$$
$$\quad \cdots$$
$$\quad\quad \text{for } i_M = 0, 1, \cdots, r_M$$
$$\quad\quad\quad \text{"system } k\text{"} := \text{"system } \mathbf{i}\text{", with } \mathbf{i} = \{i_1, i_2, \cdots, i_M\};$$
$$\quad\quad\quad k = k + 1;$$
$$\quad \text{end}$$

An important characteristic of this ordering is that i_k changes only after all possible ordering of $\{i_{k+1}, \cdots, i_M\}$. Following this one-dimensional re-index of $\mathbf{i}$, the set of basis functions $\psi_{\mathbf{i}}$ has a new index.

As is known that there is a decay of the eigenvalues, consequently, the first few terms in the KL expansion carry a lot more weight in perturbing the diffusion coefficient. Using the eigenvalue information, and the values of $C_{i_j,j}$, we derive a heuristic algorithm that put the matrices into several groups.

We first use the mean value of the diffusion coefficient to determine which eigenvalues should be considered to carry more weight. For this purpose, we choose a parameter

$$(3.12) \qquad\qquad \theta = \frac{1}{3} \max_{x \in \mathcal{N}(\mathcal{T}^H)} \{a_0(x)\},$$

where $\mathcal{N}(\mathcal{T}^H)$ is the set of coarse mesh nodes of the domain D. If an eigenvalue is greater than θ, we consider that the change in the term associated with this eigenvalue causes large change in the diffusion coefficient. Assume, by this criterion, we have decided that the first p terms may cause large perturbation. Then the sequence of linear systems are divided into $(r_1 + 1) \times (r_2 + 1) \times \cdots \times (r_p + 1)$ groups. In each group, there are $(r_{p+1} + 1) \times \cdots \times (r_M + 1)$ systems. The systems in different groups have different constants $C_{i_j,j}, j = 1, 2, \cdots, p$, in the first p terms of the KL expansion. While the systems in the same group have the same first p terms of the KL expansion.

In most situations, the systems in the same group have small differences such that we can use one of the systems to construct a Krylov subspace and a preconditioner that can be reused for solving systems in the whole group. However, in some cases, a subset of systems in the group may be close to being singular, and this subgroup of systems needs to be treated separately. More precisely, for certain systems in a given group of systems, we note that when the minimum value of the diffusion coefficient is close to zero, these systems are more sensitive to small perturbations. The iterative solver may not converge well even when the preconditioner is constructed

from another system in the same group. To single out these bad systems, we choose a cutoff parameter $\delta > 0$, and if

$$(3.13) \qquad \min_{x \in \mathcal{N}(\mathcal{T}^H)} \{a_i(x)\} \leq \delta$$

we consider the system as nearly singular. In our numerical experiments, $\delta = 0.03$. In the heuristic estimate, there is no need to compute the minimum value too accurately, we use a coarse mesh to find the minimums of the diffusion coefficients. If the minimum is less than δ, we consider this system too sensitive to be preconditioned by any other systems that are not close to being singular. We collect these sensitive systems together into a subgroup. Experiences show that systems in the subgroup have similar lower bounds, and therefore we can solve them by recycling the Krylov subspace, recycling the symbolic factorizations of the subdomain matrices, but not the preconditioner, within the subgroup. We note that these sparse matrices have the same nonzero pattern, and by using this fact, certain data structures and communication information can be recycled.

Furthermore, based on the eigenfunctions and the values of $C_{i_j,j}, j = 1, 2, \cdots, M$, we may narrow down the interval in which the sensitive systems belong to without searching through every group of systems.

To help the discussion, we consider an example with the mean and covariance functions

$$(3.14) \qquad a_0(x) = 3 + \sin(\pi x_1) \qquad C_a(x, x') = e^{-|x-x'|^2}, \qquad x \in [0,1]^2.$$

$M = 11$ and $\mathbf{r} = (3, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1)$. This implies that there are 9216 systems of form (2.18). We assume that y_j in the KL expansion is uniformly distributed in $\Gamma_j = [-\sqrt{3}, \sqrt{3}]$.

In Fig. 3.2 we show the decay of the eigenvalues. With the above ordering, for the example (3.14), Fig. 3.3 illustrates how the upper and lower bounds of the diffusion coefficient $a_{M,\mathbf{i}}$ change with $\mathbf{i}$. Note that some of the lower bounds are quite close to zero.

Since λ_1 is dominantly large compared with other eigenvalues, we consider the first term has more weight in perturbing the diffusion coefficient. $r_1 = 3$ implies there are four possible corresponding constants: $C_{i_1,1}, i_1 = 0, 1, 2, 3$. The four constants $C_{i_1,1}, i = 0, 1, 2, 3$ evenly divide all the systems into 4 groups. According to the Re-ordering Scheme for the Multi-index $\mathbf{i}$, the first group includes systems 1 to 2304, in which all the systems' diffusion coefficients share the constant $C_{i_1,1}, i_1 = 0$ in the first term of the KL expansion. Similarly, the second group from 2305 to 4608 corresponds to $C_{i_1,1}, i_1 = 1$, the third group corresponds to $C_{i_1,1}, i_1 = 2$ and the last group corresponds to $C_{i_1,1}, i_1 = 3$. We can see clearly that the bounds of the diffusion coefficient have a dramatic change for systems 2305, 4609, 6913. This matches with the change of $C_{i_1,1}, i_1 = 0, 1, 2, 3$, which are $1.49, 0.59, -1.49, -0.59$ respectively by numerical computation. $C_{i_j,j}$ decreases from 1.49 to 0.59 and then to -1.49. In the end, it increases to -0.59. Correspondingly, the bounds of the last group in the Fig. 3.3 increases too compared with the third group. In particular, since the third constant -1.49 is the smallest value among the first four constants, the minimum values of some diffusion coefficients, which are close to zero, are all in the third group $4609 - 6912$. These systems are the bad cases, for which the iterative solver does not converge well if using a preconditioner not in the same subgroup.

Fig. 3.1 shows the surface plots of the diffusion coefficients of systems 4608 to 4613. We can see clearly that the plot for system 4608 is significantly different from
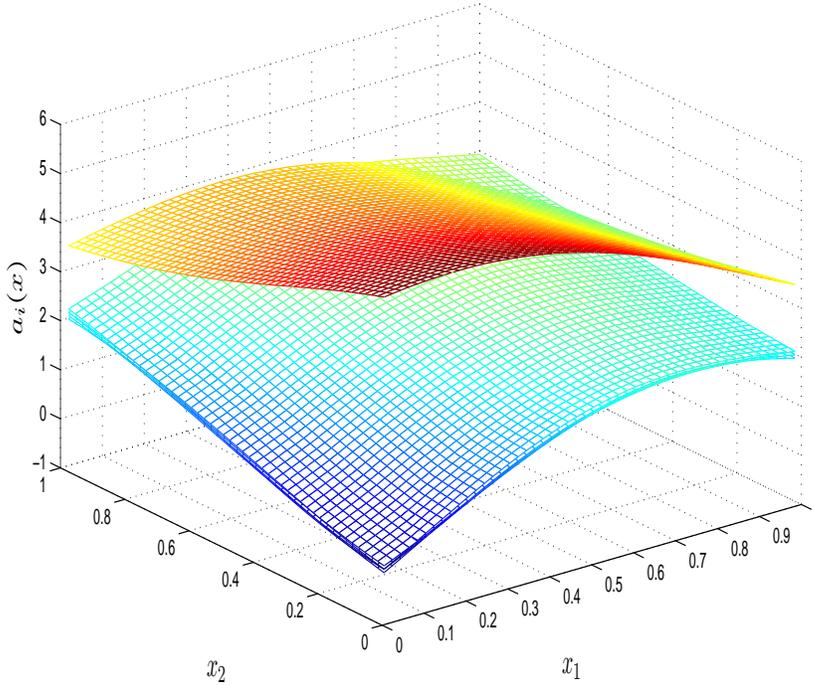
FIG. 3.1. *Surface plots of diffusion coefficients of systems* $4608 - 4613$. *The top surface is for system* 4608, *and the bottom cluster of surfaces is for systems* $4609 - 4613$.

that of the other systems. The plots for systems 4609 to 4613 have very small differences. In particular, the minimums values of the diffusion coefficients of systems 4609 to 4613 are close to zero.

Generally speaking, we would like to recycle both the Krylov subspace and the preconditioner as much as possible. For the preconditioner, there are two ways to recycle, namely (1) recycle the symbolic factorization of the subdomain matrices to save the computational work in the construction of the preconditioners for the next system; and (2) recycle the preconditioner for the next system. Obviously, the second choice saves more work than the first. However, when the matrices have significant differences from each other, the number of iterations can be large if we reuse the old preconditioner, and therefore a new preconditioner is preferred.

We summarize our analysis in term of a grouping algorithm which provides an estimate of when the systems have relatively large change and when the systems are close to be singular.

A Grouping Algorithm:
- Choose M and $\mathbf{r} = \{r_1, r_2, \cdots, r_M\}$. Generate a new index for $\mathbf{i} \leq \mathbf{r}$ according to the Re-ordering Scheme. Compute the double orthogonal basis and obtain the corresponding constants $C_{i_j,j}, j = 1, 2, \cdots, M$. Compute the eigenvalues λ_j and eigenfunctions k_j, $j = 1, 2, \cdots, M$. Choose the parameters $\delta > 0$ and $\theta > 0$ (In our example, we choose $\theta = 1/3 \max_{x \in \mathcal{N}(\mathcal{T}^H)} \{a_0(x)\}$ and $\delta = 0.03$).
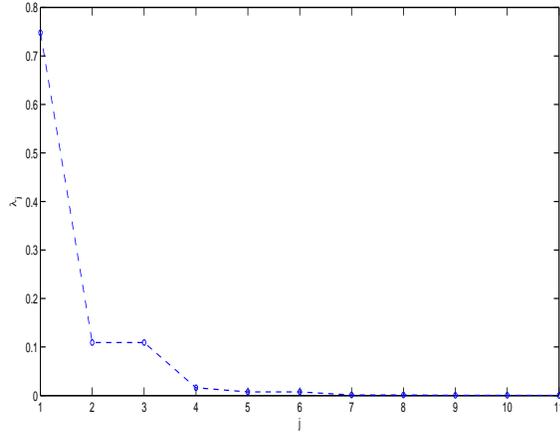
FIG. 3.2. *The decay of eigenvalues when* $C_a(x, x') = e^{-|x-x'|^2}, \quad x \in [0, 1]^2.$

- Compute $p = \max\{i \mid |\lambda_i| > \theta, \ i = 1, 2, \cdots, M\}$. Divide the systems into $(r_1 + 1) \times \cdots \times (r_p + 1)$ groups. In each group, there are $q = (r_{p+1} + 1) \times \cdots \times (r_M + 1)$ systems.
- Compute the approximate minimum value of each diffusion coefficient on the coarse mesh, $\min_{x \in \mathcal{N}(\mathcal{T}^H)}\{a_i(x)\}$. If the minimum is less than or equal to δ, the corresponding system is labeled as nearly singular or a "bad case".
- Put these bad cases in a subgroup. Construct a Krylov subspace and a preconditioner from one system in the subgroup and then recycle the Krylov subspace and recycle the symbolic factorization of the subdomain matrices for other systems in the subgroup.
- For the other systems in each group, construct a Krylov subspace and a preconditioner from one system and then recycle the Krylov subspace and recycle the preconditioner for other systems in this group.

4. Numerical experiments. We test the performance of the additive Schwarz preconditioned recycling FGMRES algorithm for solving the model problem (2.4) on a two-dimensional domain $D = [0, 1]^2$. The source term $f(x) = 1$. The mean and covariance functions are explicitly given as

$$a_0(x) = 3 + \sin(\pi x_1) \qquad C_a(x, x') = e^{-|x-x'|^2}.$$

For the KL expansion of $a(x, \omega)$, we choose $M = 11$ truncation terms as the approximation. This implies the dimension for the y-space is 11. By the selecting algorithm provided in [16], we choose $\mathbf{r} = (3, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1)$. Consequently, we have 9216 systems of form (2.18). At the same time, we assume that $y_j, j = 1, 2, \cdots, 11$ are uniformly distributed in $\Gamma_j = [-\sqrt{3}, \sqrt{3}]$. So the probability density function $\rho_j = 1/(2\sqrt{3})$. Note that this implies that the variance of y_j is unity. We compute the eigenpairs $(\lambda_j, k_j(x))$ using MATLAB. $k_j(x)$ is saved as a step function together with eigenvalues into a file to be used in the main program. The computation of the orthogonal polynomials (2.14) involves the generalized eigenvalue problem $Ax = \lambda Bx$
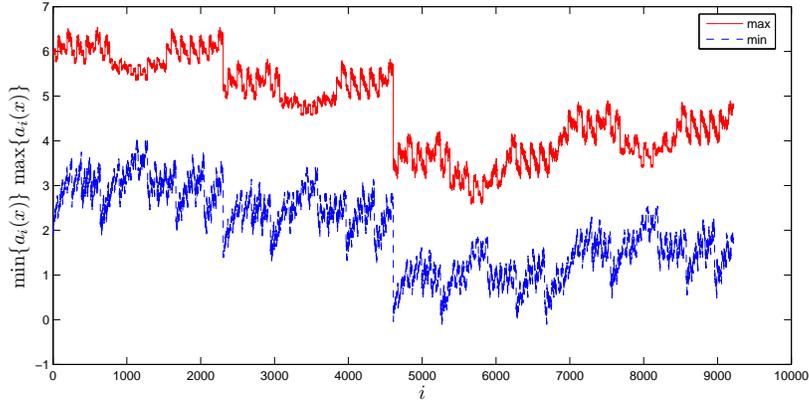
FIG. 3.3. *Top curve is the maximum of the diffusion coefficient* $\max_{x \in \mathcal{N}(\mathcal{T}^H)}\{a_{M,\mathbf{i}}(x)\}$, *and the bottom curve is the minimum of the diffusion coefficient* $\min_{x \in \mathcal{N}(\mathcal{T}^H)}\{a_{M,\mathbf{i}}(x)\}$, $M = 11$, $\mathbf{r} = \{3, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1\}$, $a_0(x) = 3 + \sin(\pi x)$.

with symmetric and positive semidefinite matrices A, B of size $r_j, j = 1, 2, \cdots, 11$. These are solved with MATLAB and saved in a file to be read by the main program.

The main program implements the additive Schwarz preconditioned recycling FGMRES. PETSc [4] is employed for the parallel computation. Both the relative tolerance and absolute tolerance for all the tests are set to 10^{-6}.

4.1. Comparison of different recycling schemes. In this section, we study the following four schemes

Scheme 1. No recycling of Krylov subspace and no recycling of preconditioner

Scheme 2. Recycle Krylov subspace and recycle symbolic factorizations of sub-domain matrices for the entire sequence of systems

Scheme 3. Recycle Krylov subspace and recycle the preconditioner for the entire sequence of systems

Scheme 4. An algorithm based on the Grouping Algorithm introduced in the previous section

Scheme 1 is used to compare with the other schemes. To take advantage of the fact that all the matrices have the same nonzero pattern, we recycle the symbolic factorization of subdomain matrices and call this as scheme 2. In PETSc this ensures that certain data structures and communication information are reused (instead of being regenerated) during successive steps, thereby increases the efficiency of the software. For example, for a parallel preconditioner such as incomplete factorization, matrix colorings and communication patterns are determined only once and then reused throughout the solution process. Scheme 3 uses the same preconditioner for all systems. It saves almost all the preconditioner construction work for all the matrices. However, this causes a steep increase in iteration numbers for some of the systems. We can see in Figs. 4.1, 4.2, 4.4, 4.5, for scheme 2, all the iteration numbers of the following systems are about half of the iteration of the first system. But for scheme 3, there are sudden increases of iteration numbers for some systems. The reason was discussed in the previous section. Scheme 4 combines the advantages of scheme 2 and scheme 3. It adaptively recycle the Krylov subspace and the preconditioner according to the Grouping Algorithm.

TABLE 4.1
Running time for different schemes and preconditioning, in second

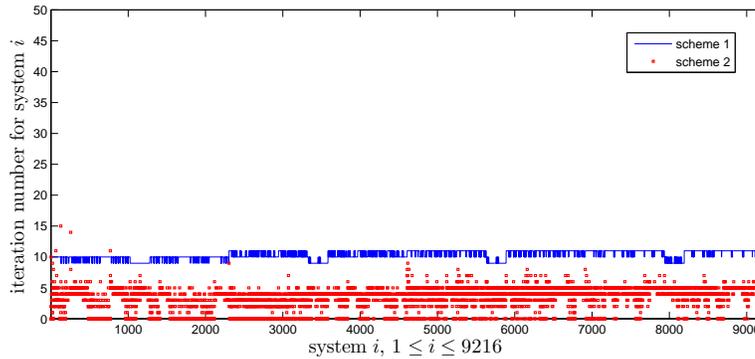| preconditioner	scheme			
	1	2	3	4
one-level ASM	12030	7693	4205	3882
two-level ASM	20980	14130	10740	8476

FIG. 4.1. *One-level ASM, iteration number comparison of scheme* 1 *and* 2

We first study the one-level additive Schwarz method; i.e. without the coarse level preconditioning. The mesh size is 256×256. The number of processors is 16. Figs. 4.1, 4.2 and 4.3 are the comparisons. The running time for both one-level and two-level ASM are in Table 4.1.

From Fig. 4.1, we see that most of the iteration numbers of scheme 2 are less than half of that using scheme 1. Many systems even have zero number of iterations, which means the solution generated from the recycled Krylov subspace is already good enough for the current system. This leads to big savings in the running time. In Fig. 4.2, we use the same preconditioner for all systems. The iteration numbers increase a little for most systems, but still less than that of scheme 1 . In a few cases, the number of iterations even reaches 50, which is the maximum iteration number we set in the implementation. These are the bad cases we discussed in the previous section. Notice that the systems with high iteration numbers in Fig. 4.2 are exactly the systems of which the minima of the coefficients are close to zero. Without surprise, the running time of scheme 3 is much less than that of scheme 2 and only about $1/3$ of scheme 1. In Fig. 4.3, we recycle Krylov subspace and the preconditioner by the Grouping Algorithm. The running time is about the same as scheme 3. The iteration numbers required for solving those few systems are also controlled.

Figs. 4.4, 4.5 and 4.6 are for the comparison of the schemes when we use two-level additive Schwarz preconditioner in the implementation. The coarse mesh is 16×16. It shows the same effects as in the one-level case. In particular, we can see that in scheme 4, there are more systems requiring zero iteration numbers.

4.2. Scalability of additive Schwarz method. In this section, we study the scalability of the additive Schwarz methods. We use recycling scheme 4 for the exper-
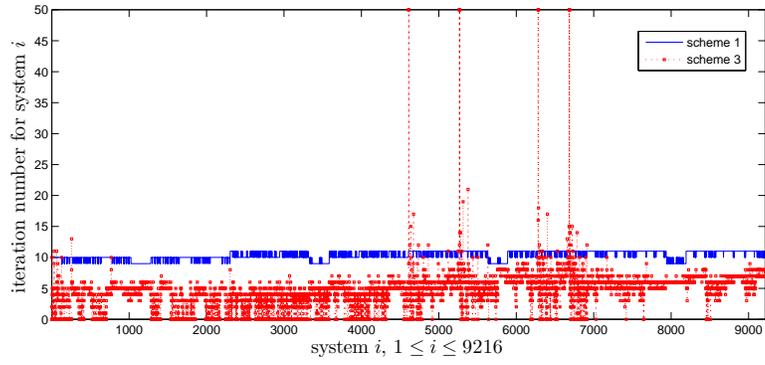
Fig. 4.2. *One-level ASM, iteration number comparison of scheme* 1 *and* 3
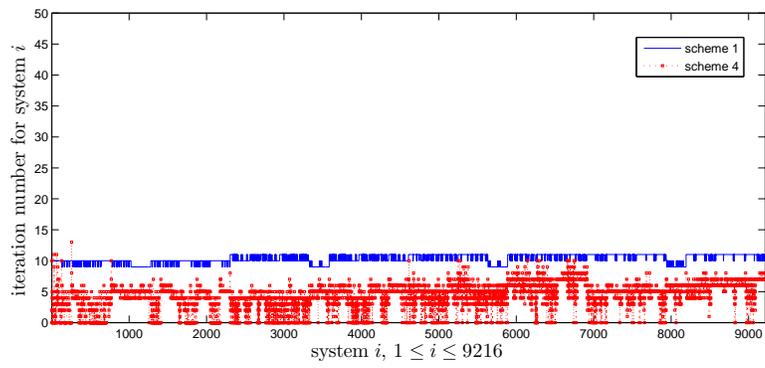

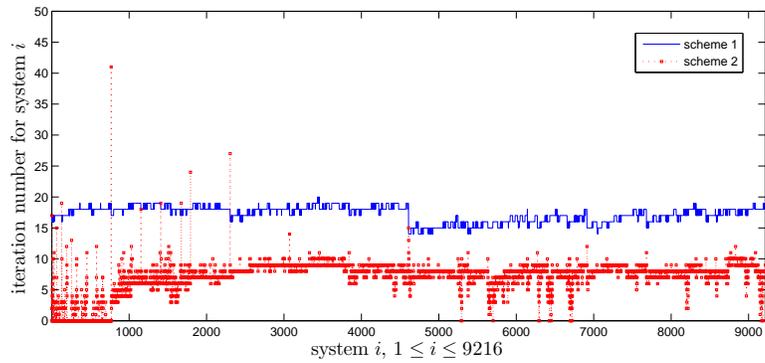Fig. 4.3. *One-level ASM, iteration number comparison of scheme* 1 *and* 4

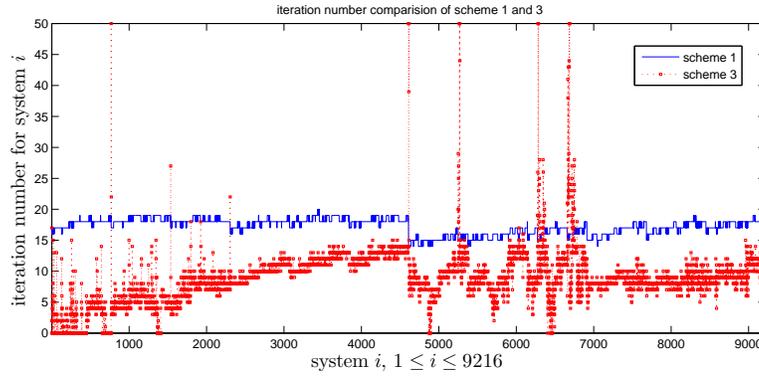Fig. 4.4. *Two-level ASM, iteration number comparison of scheme* 1 *and* 2

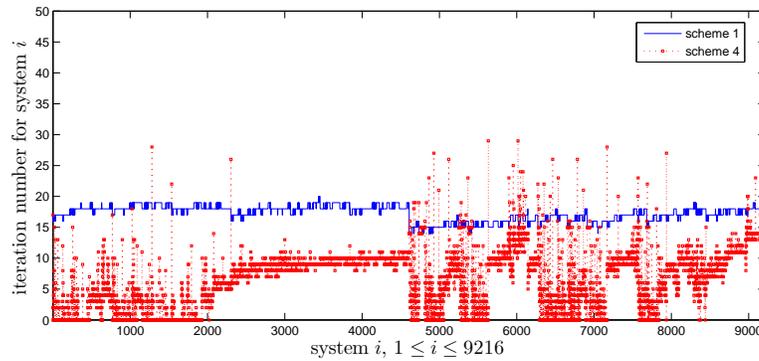FIG. 4.5. *Two-level ASM, iteration number comparison of scheme* 1 *and* 3

FIG. 4.6. *Two-level ASM, iteration number comparison of scheme* 1 *and* 4

iments. Table 4.2 shows the results of the one-level additive Schwarz method. Table 4.3 shows the results of the two-level additive Schwarz method. The iteration numbers in the tables are the average iteration numbers for solving all systems.

From Table 4.2, we can see that, for the one-level ASM, as the mesh size increases, the average iteration numbers actually decrease a little. As the number of processors increases, the average iteration numbers increase except for the 16 processors case. To conclude, with the one-level ASM, the method is scalable with respect to mesh size but not to the number of processors.

From Table 4.3, we can see that the two-level ASM is scalable with respect to both the mesh size and the number of processors. Actually, the average iteration numbers decrease a little as the mesh size increases as in the one-level case. When the number of processors increases, the average iteration numbers decrease too, except for the mesh 136×136. The number of iterations is higher than expected, more research is still needed to fully understand the situation.

TABLE 4.2
Average iteration numbers, one-level additive Schwarz preconditioner

mesh	overlap	number of subdomains			
		1	4	16	64
128×128	4	5.89	5.81	5.83	6.83
256×256	8	3.79	5.52	4.28	7.76
512×512	16	2.08	5.46	4.10	5.78

TABLE 4.3
Average iteration numbers, two-level additive Schwarz preconditioner, coarse mesh 16×16

mesh	overlap	number of subdomains			
		1	4	16	64
136×136	4	7.66	8.53	8.82	8.99
256×256	8	3.93	9.04	6.39	4.14
496×496	16	2.52	2.59	4.59	3.44

5. Conclusion. In this paper, we introduced a parallel domain decomposition preconditioned recycling Krylov subspace method for stochastic elliptic partial differential equations. By recycling the Krylov subspace and the preconditioner, the overall computational cost and time can be saved by as much as 50%. We carefully analyzed the KL expansion and derived an useful algorithm that tells us when to recompute and when to recycle some of the expensive components of the computation. We showed theoretically and experimentally that the multilevel additive Schwarz preconditioned recycling GMRES method is optimal for the entire family of linear systems arising from the discretization of the stochastic elliptic partial differential equation. A fully parallel implementation was provided and tested with up to 64 processors and scalability results were reported in the paper.

6. Acknowledgement. The authors wish to thank Professor Dongbin Xiu for many helpful discussions, and the referees for useful suggestions.

REFERENCES

[1] I. BABUSKA AND P. CHATZIPANTELIDIS, *On solving linear elliptic stochastic partial differential equations*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 4093-4122.

[2] I. BABUSKA, R. TEMPONE, AND G. ZOURARIS, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM J. Numer. Anal., 42 (2004), pp. 800-825.

[3] I. BABUSKA, R. TEMPONE, AND G. ZOURARIS, *Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 1251-1795.

[4] S. BALAY, K. BUSCHELMAN, W. D. GROPP, D. KAUSHIK, M. KNEPLEY, L. C. McINNES, B. F. SMITH, AND H. ZHANG, *PETSc Users Manual*, ANL-95/11 - Revision 2.3.0, Argonne National Laboratory, 2006.

[5] X.-C. CAI, *An additive Schwarz algorithm for nonselfadjoint elliptic equations, Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, T. CHAN, R. GLOWINSKI, J. PERIAUX, AND O. WIDLUND, EDS., SIAM 1990.

[6] X.-C. CAI, *An optimal two-level overlapping domain decomposition method for elliptic problems in two and three dimensions*, SIAM J. Sci. Comput., 14 (1993), pp. 239-247.

[7] X.-C. CAI AND O. B. WILBUND, *Domain decomposition algorithms for indefinite elliptic problems*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 243-258.

[8] A. CHAPMAN AND Y. SAAD, *Deflated and argumented Krylov subspace techniques*, Numer. Linear Algebra Appl., 4 (1997), pp. 43-66.

[9] P. DOSTERT, Y. EFENDIEV, T. Y. HOU, AND W. LUO, *Coarse-gradient Langevin algorithms for dynamic data integration and uncertainty quantification*, J. Comput. Phys., 217 (2006), pp. 123-142.

[10] M. DRYJA AND O. B. WIDLUND, *Some domain decomposition algorithms for elliptic problems*, Iterative Methods for Large Linear Systems, L. HAYES AND D. KINCAID, eds., Academic Press, San Diego, California, 1989.

[11] M. EIERMANN, O. ERNST, AND E. ULLMANN, *Computational aspects of the stochastic finite element method*, Proceedings of ALGORITMY, 2005, pp. 1-10.

[12] H. ELMAN, O. ERNST, D. O'LEARY, AND M. STEWART, *Efficient iterative algorithms for the stochastic finite element method with application to acoustic scattering*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 1037-1055.

[13] H. ELMAN AND D. FURNIVAL, *Solving the stochastic steady-state diffusion problem using multigrid*, Department of Computer Science, University of Maryland, Tech. Report UMCP-CSD:CS-TR-4786, 2006.

[14] C. FARHAT, L. CRIVELLI, AND F. X. ROUX, *Extending substructure based on iterative solvers to multiple load and repeated analyses*, Comput. Methods Appl. Mech. Engrg., 117 (1994), pp. 195-209.

[15] P. F. FISCHER, *Projection techniques for iterative solution of $Ax = b$ with successive right-hand sides*, Comput. Methods Appl. Mech. Engrg., 163 (1998), pp. 193-204.

[16] P. FRAUENFELDER, C. SCHWAB, AND R. A. TODOR, *Finite elements for elliptic problems with stochastic coefficients*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 205-228.

[17] R. G. GHANEM AND P. D. SPANOS, *Stochastic Finite Elements: A Spectral Approach*, Revised Edition, Springer-Verlag, 1991.

[18] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computation*, 3rd Edition, The John Hopkings University Press, 1996.

[19] K. GURUPRASAD, D. E. KEYES, AND J. H. KANE, *GMRES for sequentially multiple nearby sysstems*, Technical Report (1995), Old Dominium University.

[20] H. KLIE, M. F. WHEELER, G. LIU, AND D. ZHANG, *Stochastic subspace projection methods for efficient multiphase flow uncertaintiy assessment*, Tenth European Conference on the Mathematics of Oil Recovery, 2006.

[21] M. LOÈVE, *Probability Theory, Vol. I, II*, Springer, New York, 1978.

[22] H. MATTHIES AND A. KEESE, *Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 1295-1331.

[23] R. B. MORGAN, *A restarted GMRES method augmented with eigenvectors*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 1154-1171.

[24] M. L. PARKS, E. D. STURLER, G. MACKEY, D. JOHNSON, AND S. MAITI, *Recycling Krylov subspaces for sequences of linear systems*, SIAM J. Sci. Comput., 28 (2006), pp. 1651-1674.

[25] L. ROMAN AND M. SARKIS, *Stochastic Galerkin method for elliptic SPDEs: A white noise approach*, Discrete and Continuous Dynamical Systems, 6 (2006), pp. 941-955.

[26] V. SIMONCINI AND E. GALLOPOULOS, *An iterative method for nonsymmetric systems with multiple right-hand sides*, SIAM J. Sci. Comput. 16 (1996), pp. 917-933.

[27] B. SMITH, P. BJØRSTAD, AND W. GROPP, *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*, Cambridge University Press, Cambridge, MA, 1996.

[28] A. TOSELLI AND O. WIDLUND, *Domain Decomposition Methods – Algorithms and Theory*, Springer-Verlag, 2005.

[29] D. XIU AND G. E. KARNIADAKIS, *The Wiener-Askey polynomial chaos for stochastic differential equations*, SIAM J. Sci. Comput., 24 (2002), pp. 619-644.

[30] D. XIU AND G. E. KARNIADAKIS, *Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 4927-4948.

[31] N. WIENER, *The homogeneous chaos*, Amer. J. Math., 60 (1930), pp. 897-936.