Journal of Computational and Applied Mathematics 412 (2022) 114298

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Block triangular preconditioning for stochastic Galerkin method

Dongwu Wang [a], Bin Zheng [b,*], Long Chen [a], Guang Lin [c], Jinchao Xu [d]

[a] Department of Mathematics, University of California, Irvine, CA 92697, USA
[b] College of Mathematics, Beijing University of Technology, Beijing 100124, China
[c] Department of Mathematics, Purdue University, West Lafayette, IN 47906, USA
[d] Department of Mathematics, Pennsylvania State University, University Park, PA 16802, USA

ARTICLE INFO

Article history:
Received 24 December 2020
Received in revised form 15 March 2022

Keywords:
Stochastic Galerkin method
Polynomial chaos
Block triangular preconditioner
Multigrid

ABSTRACT

In this paper, we develop a new block triangular preconditioner for solving partial differential equations with random coefficients. We prove spectral bounds for the preconditioned system. Several numerical examples are provided to demonstrate the efficiency of this preconditioner, especially for stochastic problems with large variance.

1. Introduction

Partial differential equations (PDEs) with random coefficients are widely used as mathematical models to quantify uncertainties in many physical and engineering applications, for example, flows in heterogeneous porous media [1], thermo-fluid processes [2,3], fluid–structure interactions [4], etc.

There has been a tremendous amount of study on the numerical solutions of PDEs with random coefficients, including sampling-based methods (e.g., Monte Carlo method or stochastic collocation method) and intrusive methods (e.g., stochastic Galerkin method). It is well known that the convergence rate of the Monte Carlo method is proportional to $1/\sqrt{N}$ where N is the number of samples. Stochastic collocation method achieves faster convergence by combining a Galerkin discretization in space with collocation in stochastic dimensions [5–7]. To alleviate the curse-of-dimensionality due to the exponential growth of the number of collocation points in a tensor grid, Smolyak sparse grid method has been developed [8]. These non-intrusive methods have the advantage of the ease of implementation as the corresponding deterministic solvers can be applied.

Stochastic Galerkin method [9] is an intrusive type of discretization for solving PDEs with random coefficients. The size of the resulting discrete equations is relatively smaller and scales approximately $1/2^p$ (where p denotes the order of the stochastic discretization) times the number of the sparse-grid stochastic collocation equations [8,10]. In [11,12], experimental comparisons are provided to show that the stochastic Galerkin method is computationally less expensive when efficient solvers are available.

Stochastic Galerkin discretization typically combines a finite element discretization in the spatial dimensions with a Galerkin projection onto orthogonal polynomials [13,14] in stochastic dimensions. Unlike the non-intrusive approach, deterministic solvers cannot be applied directly to solve the coupled system of algebraic equations and fast

* Corresponding author.
E-mail addresses: dongwuw@uci.edu (D. Wang), zhengbin@bjut.edu.cn (B. Zheng), chenlong@math.uci.edu (L. Chen), guanglin@purdue.edu (G. Lin), xu@math.psu.edu (J. Xu).

https://doi.org/10.1016/j.cam.2022.114298
0377-0427/

solvers/preconditioners need to be developed. We refer to [15,16] for comprehensive overviews and comparisons of iterative solvers for linear systems resulting from the stochastic Galerkin discretizations.

There are a number of special structures in these linear system of equations which make it possible to develop fast linear solvers. For example, due to the block sparsity structure, Krylov subspace methods that require only matrix–vector multiplication are well suited for these linear systems. To accelerate their convergence, various preconditioning techniques have been designed, including a block diagonal preconditioner [17–19] and a Kronecker product preconditioner [20], etc. Multigrid methods with the optimal order of computational complexity in the physical space are also investigated theoretically and numerically for stochastic Galerkin discretizations [21–24]. The use of hierarchical polynomial chaos basis functions introduces a hierarchical structure in the stochastic dimensions which can also be beneficial for the design of iterative solvers [15,17,18].

In this work, we explore more on the sparse pattern of the stochastic space by introducing a simplex structure and develop a block triangular preconditioner for the stochastic Galerkin discretization. The geometric multigrid method has been utilized to invert the diagonal blocks. Numerical results show that the GMRES method preconditioned by this preconditioner performs better than MINRes method with traditional block diagonal preconditioner as well as GMRES method with Kronecker product preconditioner. The block triangular preconditioner is more robust and efficient in terms of iteration steps as well as computational time. Theoretical analysis is also given to show the spectral bounds of the preconditioned matrix. In particular, a tighter bound for the preconditioned system is provided using the Gershgorin's theorem for block matrices.

The rest of the paper is organized as follows. In Section 2, we describe the model elliptic problem with random diffusion coefficient and its stochastic Galerkin discretization. In Section 3, we discuss several important structures of the matrix from stochastic Galerkin discretization. In Section 4, we introduce the block triangular preconditioner. Some eigenvalue analysis for the preconditioned GMRES method is described in Section 5. In Section 4.3 we review the Kronecker product preconditioner. Finally, the performance of the proposed block triangular preconditioner is demonstrated in Section 6.

2. Model problem and stochastic Galerkin discretization

Let us consider the following second order elliptic problem

$$\begin{aligned} -\nabla\cdot(a(x,\omega)\nabla u(x,\omega)) &= f(x),\ x\in D\subset\mathbb{R}^d,\\ u(x,\omega) &= 0,\ x\in\partial D, \end{aligned} \tag{1}$$

where the diffusion coefficient is a real-valued random field, i.e., for each $x\in D$, $a(x,\cdot)$ is a random variable defined on a probability space $(\Omega,\mathcal{F},P)$. We assume that a is bounded and uniformly coercive, i.e.

$$\exists\ a_{\min},a_{\max}\in(0,+\infty):\ P(\omega\in\Omega:\ a(x,\omega)\in[a_{\min},a_{\max}],\ \forall x\in\overline{D})=1. \tag{2}$$

Introducing the tensor product Hilbert space $V=L_P^2(\Omega)\otimes H_0^1(D)$ and the corresponding inner product

$$(u,v)_V=\int_\Omega\left(\int_D\nabla u(x,\omega)\cdot\nabla v(x,\omega)\,\mathrm{d}x\right)\mathrm{d}P(\omega),$$

the weak solution $u\in V$ is a random function such that $\forall\ v\in V$:

$$\int_\Omega\left(\int_D a(x,\omega)\nabla u(x,\omega)\cdot\nabla v(x,\omega)\,\mathrm{d}x\right)\mathrm{d}P(\omega)=\int_\Omega\left(\int_D f(x)\,v(x,\omega)\,\mathrm{d}x\right)\mathrm{d}P(\omega). \tag{3}$$

The well-posedness of the above variational problem (3) follows from (2) and the Lax–Milgram lemma.

2.1. Karhunen-Loève expansion

We approximate the input random field $a(x,\omega)$ by the following truncated Karhunen–Loève expansion

$$a(x,\omega)\approx a_m(x,\omega):=\bar{a}(x)+\sum_{k=1}^{m}\sqrt{\lambda_k}b_k(x)\xi_k(\omega), \tag{4}$$

where $\bar{a}(x)$ is the mean value of $a(x,\omega)$, λ_k and $b_k(x)$ are the eigenvalues and eigenfunctions of $C:L^2(D)\to L^2(D)$ defined by $\int_D\mathrm{Cov}_a(x,\cdot)\,u(x)\,\mathrm{d}x$.

Given the continuous covariance function $\mathrm{Cov}_a(x,y)$, the above expansion (4) approximates $a(x,\omega)$ with minimized mean square error [9]. The number of terms in the expansion, m, depends on the eigenvalue decay rate, and in turn, depends on the stochastic regularity, i.e., the smoothness of the covariance function. The expansion coefficients $\xi_k(\omega)$ are pairwise uncorrelated random variables with images $\Gamma_k=\xi_k(\Omega)$, and probability density functions (PDFs) $\rho_k:\Gamma_k\to\mathbb{R}^n$. The joint PDF of the random vector $\xi=(\xi_1,\ldots,\xi_m)$ is denoted by $\rho(\xi)$, and the image $\Gamma=\Pi_{k=1}^m\Gamma_k$.

If $a(x,\omega)$ is a Gaussian random field, ξ_k will also be Gaussian and mutually independent with joint PDF $\rho(\xi)=\Pi_{k=1}^m\rho_k(\xi_k)$. In general, for non-Gaussian random fields, ξ_k may not be independent and their distributions are not known. Several methods have been developed to estimate the distributions of ξ_k and to simulate non-Gaussian processes

using Karhunen–Loève expansion, see [25,26]. A non-Gaussian random field may be approximated by polynomial chaos expansion, see [27,28]. As an example, we consider the exponential covariance function

$$\mathrm{Cov}_a(x, y) = \sigma^2 \exp\left(-|x - y|/L\right), \tag{5}$$

where σ denotes the standard deviation and L is the correlation length.

Remark 1. When replacing the diffusion coefficient $a(x, \omega)$ by the truncated Karhunen–Loève expansion $a_m(x, \omega)$, it is important to verify the uniform coercivity condition (2) so that the problem

$$\begin{aligned} -\nabla \cdot (a_m(x, \omega)\nabla u(x, \omega)) &= f(x), \quad x \in D, \\ u(x, \omega) &= 0, \quad x \in \partial D. \end{aligned} \tag{6}$$

is well-posed. For more discussions, including the estimate of the error between the two solutions of (1) and (6), we refer to [29,30].

One advantage of using Karhunen–Loève expansion is the separation of the stochastic and deterministic variables for the stochastic function $a(x, \omega)$. In addition, from the Doob–Dynkin lemma [31], the solution of (6) can be expressed in terms of ξ, i.e. $u(x, \omega) = u(x, \xi_1(\omega), \ldots, \xi_m(\omega))$. The above problem (6) can be reformulated as the following parametrized deterministic problem:

$$\begin{aligned} -\nabla \cdot (a_m(x, \xi)\nabla u(x, \xi)) &= f(x), \quad x \in D,\ \xi \in \Gamma, \\ u(x, \xi) &= 0, \quad x \in \partial D,\ \xi \in \Gamma. \end{aligned}$$

2.2. Stochastic Galerkin discretization

Since the weak solution $u(x, \xi)$ is defined in a tensor product space V, we consider finite dimensional approximation space also in tensor product form, i.e. $V_{h,p} = X_h \otimes \Xi_p$. When the solution is smooth/analytic in stochastic variables, spectral approximation using global polynomials of total degree $\leq p$ in m variables defined in Γ

$$\Xi_p = \mathrm{span}\{\psi_1(\xi), \ldots, \psi_{N_\xi}(\xi)\} \subset L^2_\rho(\Gamma)$$

are good candidates for approximations in the stochastic space.

For the global polynomials ψ_i, we use the orthogonal polynomials associated with the density function ρ (i.e., generalized Polynomial Chaos (gPC) [14]), e.g., Legendre polynomials for uniform distribution, and Hermite polynomial for Gaussian distribution, etc. The dimension of the space Ξ_p is given by

$$N_\xi = \frac{(m + p)!}{m!p!}.$$

For example, when $m = 6$, $p = 4$, $\dim(\Xi_4) = 210$. Notice that for the same m and p, the stochastic basis of stochastic collocation method has a dimension of $6^4 = 1296$ which is 10 times larger than $\dim(\Xi_4)$.

For the spatial approximation, we choose the standard finite element discretization, i.e., piecewise polynomials with respect to a given mesh $\mathcal{T}_h$ (h is the spatial discretization parameter)

$$X_h = \mathrm{span}\{\phi_1(x), \ldots, \phi_{N_x}(x)\} \subset H^1_0(D),$$

where N_x is the dimension of X_h. Hence, the discrete solution $u_{h,p}$ has the following form

$$u_{h,p}(x, \xi) = \sum_{j=1}^{N_\xi} u_j(x)\psi_j(\xi) = \sum_{j=1}^{N_\xi} \left(\sum_{s=1}^{N_x} U_{j,s}\phi_s(x) \right) \psi_j(\xi). \tag{7}$$

An a priori estimate for the discretization error $\|u - u_{h,p}\|_{L^2_\rho(\Gamma)\otimes H^1_0(D)}$ is given in [7]. Statistical information including mean, variance, etc., can then be obtained from the explicit formula given by Eq. (7), which give good approximations to those of the exact solution.

The stochastic Galerkin finite element method is defined by applying the Galerkin projection in the tensor product space $V_{h,p}$. More precisely, find $u_{h,p} \in V_{h,p}$ such that

$$\mathcal{A}(u_{h,p}, v) = (f, v), \quad \forall\, v \in V_{h,p}$$

where

$$\mathcal{A}(u_{h,p}, v) := \int_\Gamma \rho(\xi) \int_D a_m(x, \xi)\nabla_x u_{h,p}(x, \xi) \cdot \nabla_x v(x, \xi)\mathrm{d}x\mathrm{d}\xi,$$

and

$$(f, v) := \int_\Gamma \rho(\xi) \int_D v(x, \xi) f(x)\mathrm{d}x\mathrm{d}\xi.$$

Since $\mathcal{A}(\cdot,\cdot)$ is symmetric and positive definite, it also introduces an inner product and the associated norm (denote by $(\cdot,\cdot)_a$ and $\|\cdot\|_a$, respectively).

We refer to [32] for a more thorough discussion of the stochastic Galerkin method.

3. Matrix structures and iterative solvers

In this section, we describe some important matrix structures for the stochastic Galerkin matrix $\mathcal{A}$ which are useful in designing efficient and robust preconditioners.

3.1. Tensor product structure

Consider the semi-discretization in the stochastic domain, the corresponding Galerkin projection $u^{(p)}(\cdot,\xi) : \Gamma \to X_h$ satisfies for each stochastic basis polynomial $\psi_i(\xi)$, $i = 1,\ldots,N_\xi$:

$$\int_\Gamma -\nabla\cdot(a_m(x,\xi)\nabla u^{(p)}(x,\xi))\,\psi_i(\xi)\rho(\xi)\,\mathrm{d}\xi = \int_\Gamma f(x)\,\psi_i(\xi)\rho(\xi)\,\mathrm{d}\xi, \tag{8}$$

where the semi-discrete approximation $u^{(p)}(x,\xi) = \sum_{j=1}^{N_\xi} u_j(x)\psi_j(\xi)$. Note Eq. (8) is a system of N_ξ equations with N_ξ unknown functions $\{u_j(x)\}_{j=1}^{N_\xi}$, i.e. the stochastic 'stiffness' matrix is given by

$$\mathcal{A} = \begin{pmatrix} \mathcal{A}_{1,1} & \mathcal{A}_{1,2} & \cdots & \mathcal{A}_{1,N_\xi} \\ \mathcal{A}_{2,1} & \mathcal{A}_{2,2} & \cdots & \mathcal{A}_{2,N_\xi} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{A}_{N_\xi,1} & \mathcal{A}_{N_\xi,2} & \cdots & \mathcal{A}_{N_\xi,N_\xi} \end{pmatrix},$$

where each entry $\mathcal{A}_{i,j}$ contains spatial differentiation and is given by

$$\mathcal{A}_{i,j} = \int_\Gamma -\nabla\cdot(a_m(x,\xi)\nabla u_j(x))\psi_j(\xi)\psi_i(\xi)\rho(\xi)\,\mathrm{d}\xi.$$

Substitute in the truncated Karhunen–Loève expansion (4) for a_m, we get

$$\begin{aligned}\mathcal{A}_{i,j} &= -\nabla\cdot(\bar{a}(x)\nabla u_j(x))\int_\Gamma \psi_j(\xi)\psi_i(\xi)\rho(\xi)\,\mathrm{d}\xi \\ &- \sum_{k=1}^{m}\sqrt{\lambda_k}\nabla\cdot(b_k(x)\nabla u_j(x))\int_\Gamma \xi_k\psi_j(\xi)\psi_i(\xi)\rho(\xi)\,\mathrm{d}\xi.\end{aligned}$$

Next, we apply Galerkin projection to discretize spatial differential operators. Let

$$u_j(x) = \sum_{s=1}^{N_x} U_{j,s}\phi_s(x),$$

where $\{\phi_s(x)\}_{s=1}^{N_x}$ is the set of finite element basis functions. Multiply the spatial derivative terms in each $\mathcal{A}_{i,j}$ by the test function $\phi_r(x)$ (for $r = 1,\ldots,N_x$) and integration by parts gives

$$\sum_{s=1}^{N_x} U_{j,s}\int_D \bar{a}(x)\nabla\phi_s(x)\cdot\nabla\phi_r(x)\,\mathrm{d}x + \sum_{k=1}^{m}\sqrt{\lambda_k}\sum_{s=1}^{N_x}U_{j,s}\int_D b_k(x)\nabla\phi_s(x)\cdot\nabla\phi_r(x)\,\mathrm{d}x. \tag{9}$$

From (9), we define spatial stiffness matrices K_0 and K_k (for $k = 1,\ldots,m$) as

$$\begin{aligned}K_0(r,s) &= \int_D \bar{a}(x)\nabla\phi_s(x)\cdot\nabla\phi_r(x)\,\mathrm{d}x, \\ K_k(r,s) &= \sqrt{\lambda_k}\int_D b_k(x)\nabla\phi_s(x)\cdot\nabla\phi_r(x)\,\mathrm{d}x,\ \ r,s = 1,\ldots,N_x.\end{aligned} \tag{10}$$

Similarly, we define the stochastic matrices G_0 and $\{G_k\}_{k=1}^m$ as

$$\begin{aligned}G_0(i,j) &= \int_\Gamma \psi_j(\xi)\psi_i(\xi)\rho(\xi)\,\mathrm{d}\xi, \\ G_k(i,j) &= \int_\Gamma \xi_k\psi_j(\xi)\psi_i(\xi)\rho(\xi)\,\mathrm{d}\xi,\ \ i,j = 1,\ldots,N_\xi.\end{aligned} \tag{11}$$

Finally, the stochastic Galerkin matrix A can be described as the sum of Kronecker products:

$$A = G_0 \otimes K_0 + \sum_{k=1}^{m} G_k \otimes K_k. \tag{12}$$

In practice, one does not assemble the stochastic Galerkin matrix A explicitly. Instead, using the tensor product structure (12), only $m+1$ spatial stiffness matrices K_k and $m+1$ stochastic matrices G_k need to be stored.

3.2. Simplex structure

We explore the sparse pattern of the stochastic matrices G_k. Consider a lattice with vertices defined as

$$\mathcal{V}_{p,m}^{\mathfrak{L}} = \left\{ \boldsymbol{i} = (i_1, i_2, \ldots, i_m) \in \mathbb{Z}^m \mid 0 \leqslant i_k \leqslant p,\ k = 1, 2, \ldots, m \right\},$$

where m is the number of truncated terms in the Karhunen–Loève expansion and p is the maximum degree for each univariate orthonormal polynomial. There exists a bijection between the lattice vertices $\mathcal{V}_{p,m}^{\mathfrak{L}}$ and the tensor product polynomial space $\mathfrak{T}_{p,m}$

$$\mathfrak{T}_{p,m} := \left\{ \prod_{k=1}^{m} \psi_k(\xi_k) \mid \max_{1 \leq k \leq m} \deg(\psi_k) \leqslant p \right\} \tag{13}$$

if we consider each i_k in the multi-index $\boldsymbol{i}$ as the degree of univariate polynomial of $\psi_{i_k}(\xi_k)$. Alternatively, define a simplex with vertices $\mathcal{V}_{p,m}^{\mathfrak{S}}$ as a subset of $\mathcal{V}_{p,m}^{\mathfrak{L}}$:

$$\mathcal{V}_{p,m}^{\mathfrak{S}} = \left\{ \boldsymbol{i} = (i_1, i_2, \ldots, i_m) \in \mathbb{Z}^m \mid \sum_{k=1}^{m} i_k \leqslant p,\ i_k \geqslant 0,\ k = 1, 2, \ldots, m \right\}.$$

There also exists a bijection between the simplex vertices $\mathcal{V}_{p,m}^{\mathfrak{S}}$ and the complete polynomial space $\mathfrak{C}_{p,m}$

$$\mathfrak{C}_{p,m} := \left\{ \prod_{k=1}^{m} \psi_k(\xi_k) \mid \sum_{k=1}^{m} \deg(\psi_k) \leqslant p \right\}. \tag{14}$$

In view of lattice, stochastic collocation method is corresponding to the lattice vertices $\mathcal{V}_{p,m}^{\mathfrak{L}}$ with tensor product polynomial space $\mathfrak{T}_{p,m}$, while stochastic Galerkin method is corresponding to the simplex vertices $\mathcal{V}_{p,m}^{\mathfrak{S}}$ with complete polynomial space $\mathfrak{C}_{p,m}$.

For any symmetric $N_\xi \times N_\xi$ matrix G, we can define a bijection between G and an undirected graph $\mathcal{G}_{p,m} = \left(\mathcal{V}_{p,m}^{\mathfrak{S}}, \mathcal{E}_{p,m}^{\mathfrak{S}}\right)$ on the simplex vertices $\mathcal{V}_{p,m}^{\mathfrak{S}}$ and their edges $\mathcal{E}_{p,m}^{\mathfrak{S}} = \left\{(\boldsymbol{i}, \boldsymbol{j}) \ :\ \boldsymbol{i},\ \boldsymbol{j} \in \mathcal{V}_{p,m}^{\mathfrak{S}}\right\}$ by setting the weight in the graph $w_{\boldsymbol{ij}} = G(\hat{i}, \hat{j})$, where $\hat{i},\ \hat{j}$ are the global index corresponding to the two multi-indices $\boldsymbol{i},\ \boldsymbol{j}$.

Recall the three-term recurrence formula for orthonormal polynomials with the initial terms $\psi_{-1}(\xi) = 0$ and $\psi_0(\xi) = 1$:

$$\psi_{k+1}(\xi) = (a_k \xi + b_k)\psi_k(\xi) - c_k \psi_{k-1}(\xi), \quad \forall k \geqslant 1, \tag{15}$$

where $a_k,\ c_k$ are non-zero and $c_k a_k a_{k-1} > 0$ for all $k \geqslant 1$. Now (11) turns into

$$G_k(\hat{i}, \hat{j}) = \begin{cases} \dfrac{c_{i_k+1}}{a_{i_k+1}}, & \text{if } i_k = j_k - 1 \text{ and } i_l = j_l,\ l \in \{1, 2, \ldots, m\} \setminus \{k\}, \\ \dfrac{c_{i_k}}{a_{i_k}}, & \text{if } i_k = j_k + 1 \text{ and } i_l = j_l,\ l \in \{1, 2, \ldots, m\} \setminus \{k\}, \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

The result above indicates that the stochastic matrices G_k are sparse. The $(\hat{i}, \hat{j})$-th entry of G_k is not zero only when their corresponding multi-index points on $\mathcal{V}_{p,m}$ are adjacent in the kth coordinate, i.e. $|i_k - j_k| = 1$ and $i_l = j_l,\ l \in \{1, 2, \ldots, m\} \setminus \{k\}$.

In the sum $G = \sum_{k=1}^{m} G_k$, $G(\hat{i}, \hat{j})$ is non-zero only when $\|\boldsymbol{i} - \boldsymbol{j}\|_1 = 1$, where $\|\cdot\|_1$ is the l_1-norm of vectors. Therefore, we can define a simplex graph $\mathcal{G}_{p,m} = \left(\mathcal{V}_{p,m}^{\mathfrak{S}}, \mathcal{E}_{p,m}^{\mathfrak{S}}\right)$ on the simplex vertices $\mathcal{V}_{p,m}^{\mathfrak{S}}$ and the edges

$$\mathcal{E}_{p,m}^{\mathfrak{S}} = \left\{(\boldsymbol{i}, \boldsymbol{j}) \mid \boldsymbol{i},\ \boldsymbol{j} \in \mathcal{V}_{p,m}^{\mathfrak{S}},\ \|\boldsymbol{i} - \boldsymbol{j}\|_1 = 1\right\}.$$

Fig. 1 is an example for the simplex graph $\mathcal{G}_{p,m}$ with normalized Hermite polynomials when $p = 5$ and $m = 2$.

Due to the simplex structure, the vertices of the simplex graph $\mathcal{G}_{p,m}$ are well-separated into two groups: odd total degree and even total degree. Inside each group, the vertices are disjointed. More precisely, define

$$\mathcal{I}_{\text{odd}} = \left\{ \boldsymbol{i} \in \mathcal{V}_{p,m}^{\mathfrak{S}} \mid \sum_{k=1}^{m} i_k \text{ is odd} \right\},$$

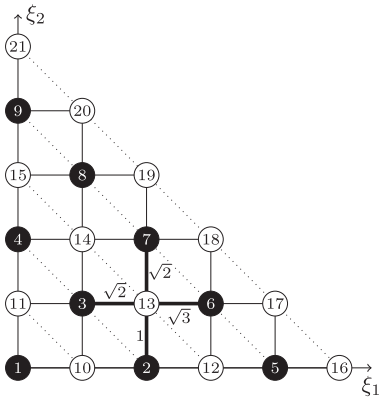

Fig. 1. The simplex lattice and a corresponding graph of stochastic matrix G with parameters $p = 5$ and $m = 2$. The black points are those with even total degrees while the white points are those with odd total degrees. The black points are only adjacent (connected) to white points and vice versa. For example, point #13 (corresponding to $\psi_2(\xi_1)\psi_1(\xi_2)$ with total degree three) is connected to point #2, #3, #6, #7, all with an even total degree. Moreover, the entries of $G(13, \cdot)$ are zero except $G(13, 7) = G_1(13, 7) = \sqrt{2}$, $G(13, 6) = G_1(13, 6) = \sqrt{3}$, $G(13, 2) = G_2(13, 2) = 1$ and $G(13, 3) = G_2(13, 3) = \sqrt{2}$.

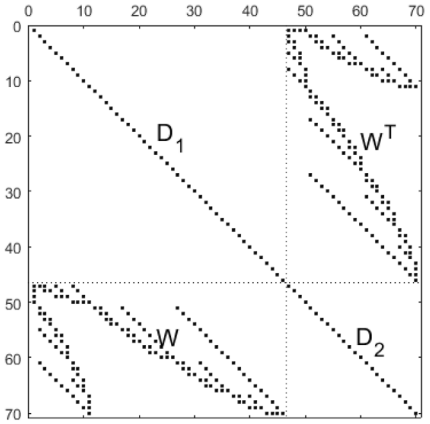

Fig. 2. Block sparsity structure of the Galerkin matrix A under the red–black ordering of their total degree ($m = 4$, $p = 4$).

$$\mathcal{I}_{\text{even}} = \left\{ \boldsymbol{i} \in \mathcal{V}^{\mathfrak{S}}_{p,m} \mid \sum_{k=1}^{m} i_k \text{ is even} \right\}.$$

Then $G = \sum_{k=1}^{m} G_k$ admits a sparse block matrix structure (after suitable reordering)

$$G = \begin{pmatrix} O & B^{\mathsf{T}} \\ B & O \end{pmatrix}, \text{ with } B = G(\mathcal{I}_{\text{odd}}, \mathcal{I}_{\text{even}}).$$

Fig. 2 is an example for the simplex graph $\mathcal{G}_{p,m}$ with normalized Hermite polynomials when $p = 4$ and $m = 4$. We will use this property to design a block Gauss–Seidel preconditioner in the next section.

4. Block preconditioners

Preconditioning techniques are often necessary in order to accelerate the convergence of the Krylov subspace methods when solving ill-conditioned linear system of equations. In particular, for the stochastic Galerkin matrix A with block sparse structure, it is natural to consider block preconditioners.

4.1. Block-diagonal preconditioner

It is known that the block-diagonal preconditioner (also known as the mean-based preconditioner) defined by

$$B_D := G_0 \otimes K_0, \tag{17}$$

works very well with the conjugate gradient (CG) method when the variance of the diffusion coefficient $a(x,\omega)$ is small [17,18]. By the uniform ellipticity assumption, the mean stiffness matrix K_0 is SPD. The multigrid method can be used to invert each diagonal block approximately.

For the block-diagonal preconditioned systems, the spectral bounds are derived in [19]. We first recall results in [19]. We begin with the estimate on the generalized eigenvalues of K_k relative to K_0. Recall that μ is mean of ω and σ is its variance.

Lemma 1 (*Lemma 3.4 in [19]*). *Let $\beta_k = \sqrt{\lambda_k}\|b_k\|_\infty$ for $k = 1, 2, \ldots, m$, then*

$$\rho(K_0^{-1}K_k) \le \frac{\sigma}{\mu}\beta_k.$$

We then recall their bound obtained for the block-diagonal preconditioner $B_D = G_0 \otimes K_0$.

Lemma 2 (*Theorem 3.8 in [19]*).

$$1 - \tau \leqslant \lambda_{\min}(B_D^{-1}A) \leqslant \lambda_{\max}(B_D^{-1}A) \leqslant 1 + \tau,$$

where

$$\tau = \frac{\sigma}{\mu} C_{p+1}^{\max} \sum_{k=1}^{m} \beta_k,$$

$$C_{p+1}^{\max} = \begin{cases} H_{p+1}^{\max} = \sqrt{p-1} + \sqrt{p}, & \text{for Hermite polynomials,} \\ L_{p+1}^{\max} = \sqrt{3}, & \text{for Legendre polynomials.} \end{cases}$$

Define $\mathcal{N}(\boldsymbol{i}) = \left\{\boldsymbol{j} \in \mathcal{V}_{p,m} \mid \|\boldsymbol{i} - \boldsymbol{j}\|_1 = 1\right\}$ as the neighbors of vertex $\boldsymbol{i}$, and $\mathcal{N}(\boldsymbol{i}, k) = \{\boldsymbol{j} \in \mathcal{N}(\boldsymbol{i}) \mid |i_k - j_k| = 1\}$ as the neighbors of vertex $\boldsymbol{i}$ in k-direction. Further define $k(\boldsymbol{i}, \boldsymbol{j}) = \left\{k \in \{1, 2, \ldots, L\} \mid (\boldsymbol{i}, \boldsymbol{j}) \in \mathcal{E}_{p,m} \text{ and } |i_k - j_k| = 1\right\}$ as the different direction between indices $\boldsymbol{i}$ and $\boldsymbol{j}$. Based on the simplex structure of the stochastic matrix, the result in Lemma 2 can be improved to the following result.

Theorem 1.

$$1 - \hat{\tau} \leqslant \lambda_{\min}(B_D^{-1}A) \leqslant \lambda_{\max}(B_D^{-1}A) \leqslant 1 + \hat{\tau}, \tag{18}$$

where

$$\hat{\tau} = \frac{\sigma}{\mu} \max_{\boldsymbol{i}, |\boldsymbol{i}| \leqslant p} \sum_{k=1}^{m} \sum_{\boldsymbol{j} \in \mathcal{N}(\boldsymbol{i}, k)} G(\boldsymbol{i}, \boldsymbol{j})\beta_k, \quad G = \sum_{k=1}^{m} G_k. \tag{19}$$

We claim that $\hat{\tau}$ in Theorem 1 is indeed an improvement of τ in Lemma 2, i.e. $\hat{\tau} \leqslant \tau$. In each direction $k \in \{1, 2, \ldots, m\}$, for any $\boldsymbol{i}$ given, there are at most two vertices $\boldsymbol{j}_1, \boldsymbol{j}_2$ in $\mathcal{N}(\boldsymbol{i}, k)$, cf. Fig. 1. Therefore

$$\sum_{\boldsymbol{j} \in \mathcal{N}(\boldsymbol{i}, k)} G(\boldsymbol{i}, \boldsymbol{j})\beta_k \le C_{p+1}^{\max}\beta_k,$$

and the claim holds by simply summing up all the directions. The above equality holds only when $i_k = p - 1$ and all other $i_l = 0$, $l \neq k$. This implies that $\hat{\tau}$ in Theorem 1 is a much tighter bound compared to τ in Lemma 2, and $\tau = \hat{\tau}$ only when $m = 1$.

In order to prove (18), we introduce the Gershgorin's Theorem for block matrices [33–35]. For completeness, a proof is included here.

Theorem 2 (*Gershgorin's Theorem for Block Matrices [33–35]*). *Consider $\bar{A} = (\bar{A}_{ij}) \in \mathbb{R}^{dn \times dn}$ where $\bar{A}_{ij} \in \mathbb{R}^{d \times d}$. Denote by $\sigma(\cdot)$ the spectrum of a matrix. Define*

$$\mathcal{G}_i = \sigma(\bar{A}_{ii}) \bigcup \left\{ \lambda \notin \sigma(\bar{A}_{ii}) : \left\|(\bar{A}_{ii} - \lambda I)^{-1}\right\|^{-1} \leqslant \sum_{j \neq i} \|\bar{A}_{ij}\| \right\}, \tag{20}$$

for $i = 1, 2, \ldots, n$. Then

$$\sigma(\bar{A}) \subseteq \bigcup_{i=1}^{n} \mathcal{G}_i.$$

Proof. Suppose that $\lambda \notin \bigcup_{i=1}^{n} \sigma(\bar{A}_{ii})$ then we have

$$\bar{A} - \lambda I = \begin{pmatrix} \bar{A}_{11} - \lambda I & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \bar{A}_{nn} - \lambda I \end{pmatrix} (I + M(\lambda)), \tag{21}$$

where

$$M(\lambda) = \begin{pmatrix} 0 & (\bar{A}_{11} - \lambda I)^{-1}\bar{A}_{12} & \cdots & (\bar{A}_{11} - \lambda I)^{-1}\bar{A}_{1n} \\ (\bar{A}_{22} - \lambda I)^{-1}\bar{A}_{21} & 0 & \cdots & (\bar{A}_{22} - \lambda I)^{-1}\bar{A}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ (\bar{A}_{nn} - \lambda I)^{-1}\bar{A}_{n1} & \cdots & \cdots & 0 \end{pmatrix}.$$

Assume $\|(\bar{A}_{ii} - \lambda I)^{-1}\|^{-1} > \sum_{j \neq i} \|\bar{A}_{ij}\|$ for $i = 1, 2, \ldots, n$. Then $\|M(\lambda)\| < 1$ and hence $I + M(\lambda)$ is not singular. It implies that $\bar{A} - \lambda I$ is not singular which leads to a contradiction. Therefore λ must be in the given region. □

Now we give a proof of Theorem 1.

Proof. Since $G_0 = I$ and K_0 is positive definite, we can define $B_D^{\frac{1}{2}} = I \otimes K_0^{\frac{1}{2}}$ and then

$$\sigma(B_D^{-1}A) = \sigma(B_D^{-\frac{1}{2}} A B_D^{-\frac{1}{2}}).$$

It suffices to show the following bound for the symmetric matrix $B_D^{-\frac{1}{2}} A B_D^{-\frac{1}{2}}$

$$1 - \hat{\tau} \leqslant \lambda_{\min}(B_D^{-\frac{1}{2}} A B_D^{-\frac{1}{2}}) \leqslant \lambda_{\max}(B_D^{-\frac{1}{2}} A B_D^{-\frac{1}{2}}) \leqslant 1 + \hat{\tau}.$$

Let

$$\bar{A} = B_D^{-\frac{1}{2}} A B_D^{-\frac{1}{2}} = I + \sum_{k=1}^{m} G_k \otimes \left(K_0^{-\frac{1}{2}} K_k K_0^{-\frac{1}{2}} \right).$$

The diagonal blocks $\bar{A}_{ii} = I$. For a scalar index, we shall use the boldface letter to denote its corresponding multi-index, and the structure is easier to explore in the multi-index system, cf. Fig. 1. Using these notation, together with $G = \sum_{k=1}^{m} G_k$, we have for $i \neq j$,

$$\bar{A}_{ij} = \sum_{k=1}^{m} (G_k)_{ij} \left(K_0^{-\frac{1}{2}} K_k K_0^{-\frac{1}{2}} \right) = G(\boldsymbol{i}, \boldsymbol{j}) \left(K_0^{-\frac{1}{2}} K_{k(\boldsymbol{i},\boldsymbol{j})} K_0^{-\frac{1}{2}} \right).$$

Therefore, (20) can be written as

$$\begin{aligned} \mathcal{G}_i &= \{1\} \bigcup \left\{ \lambda \neq 1 : |1 - \lambda| \leqslant \sum_{j \neq i} \left\| G(\boldsymbol{i}, \boldsymbol{j}) \left(K_0^{-\frac{1}{2}} K_{k(\boldsymbol{i},\boldsymbol{j})} K_0^{-\frac{1}{2}} \right) \right\| \right\} \\ &= \left\{ \lambda : |1 - \lambda| \leqslant \sum_{\boldsymbol{j} \in \mathcal{N}(\boldsymbol{i})} G(\boldsymbol{i}, \boldsymbol{j}) \left\| K_0^{-\frac{1}{2}} K_{k(\boldsymbol{i},\boldsymbol{j})} K_0^{-\frac{1}{2}} \right\| \right\} \\ &= \left\{ \lambda : |1 - \lambda| \leqslant \sum_{k=1}^{m} \sum_{\boldsymbol{j} \in \mathcal{N}(\boldsymbol{i},k)} G(\boldsymbol{i}, \boldsymbol{j}) \left\| K_0^{-\frac{1}{2}} K_k K_0^{-\frac{1}{2}} \right\| \right\} \\ &\subseteq \left\{ \lambda : |1 - \lambda| \leqslant \frac{\sigma}{\mu} \sum_{k=1}^{m} \sum_{\boldsymbol{j} \in \mathcal{N}(\boldsymbol{i},k)} G(\boldsymbol{i}, \boldsymbol{j}) \beta_k \right\} \end{aligned} \tag{22}$$

Since all $\mathcal{G}_i$ are concentric at 1, we can conclude that

$$\sigma(B_D^{-\frac{1}{2}} A B_D^{-\frac{1}{2}}) \subseteq \bigcup_{i=1}^{n} \mathcal{G}_i \subseteq D\left(1, \hat{\tau}\right),$$

where $D\left(1, \hat{\tau}\right)$ is the disk centered at 1 with radius $\hat{\tau}$. And so is $\sigma(B_D^{-1}A)$. □

Since B_D is symmetric positive definite, we can write a symmetric version of the preconditioned system as

$$B_D^{-\frac{1}{2}} A B_D^{-\frac{1}{2}} = \sum_{k=0}^{m} G_k \otimes (K_0^{-\frac{1}{2}} K_k K_0^{-\frac{1}{2}}) = I + \sum_{k=1}^{m} G_k \otimes (K_0^{-\frac{1}{2}} K_k K_0^{-\frac{1}{2}}) = \begin{bmatrix} I & B^\intercal \\ B & I \end{bmatrix}$$

where $B = \sum_{k=1}^{m} G_k(\mathcal{I}_2, \mathcal{I}_1) \otimes (K_0^{-\frac{1}{2}} K_k K_0^{-\frac{1}{2}})$ is the off-diagonal of the preconditioned system $B_D^{-\frac{1}{2}} A B_D^{-\frac{1}{2}}$ corresponding to the even–odd ordering described before based on the total degrees in the stochastic space. Here $\mathcal{I}_1$ is the one in $\{\mathcal{I}_{odd}, \mathcal{I}_{even}\}$ with larger cardinality, $\mathcal{I}_2$ is the one with smaller cardinality.

With the help of Theorem 2, we can prove that the off-diagonal blocks of the preconditioned system satisfy

$$\left\| \begin{bmatrix} 0 & B^\intercal \\ B & 0 \end{bmatrix} \right\| = \hat{\tau}, \tag{23}$$

where $\hat{\tau}$ is given in (19). We will use this fact in the spectral analysis of the block triangular preconditioner in Section 5.

4.2. Block-triangular preconditioner

We can use the lower block triangular part of the matrix A as a preconditioner. This can be viewed as applying one step of the block Gauss–Seidel method with zero initial guess.

Consider a splitting of the stochastic approximation space Ξ_p according to whether their total polynomial degrees are even or odd,

$$\Xi_p = \Xi_1 \oplus \Xi_2$$

and the corresponding splitting of the global approximation space

$$V_{h,p} = (X_h \otimes \Xi_1) \oplus (X_h \otimes \Xi_2),$$

which results in a 2×2 block structure.

Given $u^{(0)} = 0$, for $k = 1, 2, \ldots$, the block Gauss–Seidel iterate $u^{(k+1)}$ is given by the following two steps:

- Find $u^{(k+1/2)} \in u^{(k)} + X_h \otimes \Xi_1$ such that

$$\mathcal{A}(u^{(k+1/2)}, v) = \int_\Gamma \rho(\xi) \int_D v(x, \xi) f(x)\, \mathrm{d}x\, \mathrm{d}\xi, \quad \forall v \in X_h \otimes \Xi_1.$$

- Find $u^{(k+1)} \in u^{(k+1/2)} + X_h \otimes \Xi_2$ such that

$$\mathcal{A}(u^{(k+1)}, v) = \int_\Gamma \rho(\xi) \int_D v(x, \xi) f(x)\, \mathrm{d}x\, \mathrm{d}\xi, \quad \forall v \in X_h \otimes \Xi_1.$$

In matrix notation, the above block Gauss–Seidel method can be described by the following matrix splitting

$$A = \begin{bmatrix} D_1 & 0 \\ W & D_2 \end{bmatrix} - \begin{bmatrix} 0 & -W^\intercal \\ 0 & 0 \end{bmatrix},$$

where

$$D_1 = I_1 \otimes K_0, \quad D_2 = I_2 \otimes K_0,$$

$$W = \left(\sum_{k=1}^{m} G_k \otimes K_k \right)(\mathcal{I}_2, \mathcal{I}_1) = \sum_{k=1}^{m} G_k(\mathcal{I}_2, \mathcal{I}_1) \otimes K_k.$$

We define the block triangular preconditioner

$$B_T := \begin{bmatrix} D_1 & 0 \\ W & D_2 \end{bmatrix}. \tag{24}$$

The corresponding preconditioner system

$$\begin{bmatrix} D_1 & 0 \\ W & D_2 \end{bmatrix} \begin{bmatrix} \tilde{U}_1 \\ \tilde{U}_2 \end{bmatrix} = \begin{bmatrix} \tilde{F}_1 \\ \tilde{F}_2 \end{bmatrix}$$

may be solved inexactly by the standard multigrid V-cycle.

Remark 2. The block triangular preconditioner B_T may also be motivated by considering the block LU factorization

$$A = \begin{bmatrix} D_1 & 0 \\ W & S \end{bmatrix} \begin{bmatrix} I & D_1^{-1} W^\intercal \\ 0 & I \end{bmatrix}, \quad S = D_2 - W D_1^{-1} W^\intercal. \tag{25}$$

It is known that with the "ideal" block triangular preconditioner

$$\tilde{B}_T = \begin{bmatrix} D_1 & 0 \\ W & S \end{bmatrix}$$

the GMRES method converges in at most two iterations . However, $\tilde{B}_T$ is impractical because the Schur complement S is computationally expensive to invert. Replacing S by D in $\tilde{B}_T$ results in the block triangular preconditioner B_T.

Since B_T is nonsymmetric, we will use it with the GMRES method [36] or GPCG method [37]. To apply the standard PCG method, we may consider the symmetric block Gauss–Seidel method as the preconditioner, i.e.

$$B_S := \begin{bmatrix} D_1 & 0 \\ W & D_2 \end{bmatrix} \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}^{-1} \begin{bmatrix} D_1 & W^\top \\ 0 & D_2 \end{bmatrix}.$$

It is clear that B_S is SPD and the standard PCG method is guaranteed to converge. However, it might not be efficient as one more matrix–vector products are computed compared with the block-triangular preconditioner while the iteration steps do not decrease much.

4.3. A Kronecker product preconditioner

Another way to develop a good preconditioner of the $N_xN_\xi \times N_xN_\xi$ matrix A is to approximate it by a single tensor product $L \otimes R$ with certain matrix L of size $N_\xi \times N_\xi$ and R of size $N_x \times N_x$. Van Loan and Pitsianis [38] presented a general framework (for any $N_xN_\xi \times N_xN_\xi$ matrix A) to find such L and R by minimizing the Frobenius norm $\|A - L \otimes R\|_F$. However, the additional computational cost introduced in solving L and R might be dominant compare to solving the original stochastic Galerkin system with the previous preconditioners, and therefore is not worth it in practice. A simpler approximation with fixed L or R is brought up in Van Loan and Pitsianis [38] and then specified to the stochastic Galerkin system by Ullmann [20]. We first recall a lemma in [38] for fixed L or R.

Lemma 3 (*[38], Theorem 3*)**.** *If $R \in \mathbb{R}^{N_x \times N_x}$ is fixed, then the matrix $L \in \mathbb{R}^{N_\xi \times N_\xi}$ defined by*

$$L_{st} = \frac{\mathrm{tr}(A_{st}^\top R)}{\mathrm{tr}(R^\top R)}, \quad s, t = 1, 2, \ldots, N_\xi, \tag{26}$$

minimizes $\|A - L \otimes R\|_F$, where A_{st} is the s, t-th $N_x \times N_x$ block of A. Similarly, if $L \in \mathbb{R}^{N_\xi \times N_\xi}$ is fixed, then the matrix $R \in \mathbb{R}^{N_x \times N_x}$ defined by

$$R_{ij} = \frac{\mathrm{tr}(\bar{A}_{ij}^\top L)}{\mathrm{tr}(L^\top L)}, \quad i, j = 1, 2, \ldots, N_x, \tag{27}$$

minimizes $\|A - L \otimes R\|_F$, where $\bar{A}_{ij}$ is the i, j-th $N_\xi \times N_\xi$ block of A.

In the stochastic Galerkin system, if $L = I$ is fixed, R is simply equal to K_0 according to (27), which results in the mean-based preconditioner $B_D = I \otimes K_0$. If $R = K_0$ is fixed, then based on (26),

$$\begin{aligned} L_{st} &= \frac{\mathrm{tr}(A_{st}^\top K_0)}{\mathrm{tr}(K_0^\top K_0)}, \\ &= \frac{1}{\mathrm{tr}(K_0^\top K_0)} \left[\mathrm{tr}(K_0^\top K_0)\delta_{st} - \sum_{k=1}^{m} (G_k)_{st} \mathrm{tr}(K_k^\top K_0) \right] \end{aligned} \tag{28}$$

which implies that

$$L = I - \sum_{k=1}^{m} \frac{\mathrm{tr}(K_k^\top K_0)}{\mathrm{tr}(K_0^\top K_0)} G_k. \tag{29}$$

The corresponding preconditioner is the Kronecker product preconditioner

$$B_K = L \otimes K_0 \tag{30}$$

presented in [20]. Instead of using only the mean information for the mean-based preconditioner, the Kronecker product preconditioner utilizes all the information in the stochastic Galerkin matrices.

5. Spectral analysis

We give spectral bounds for the matrix preconditioned by the block triangular preconditioner (24). For saddle point problems, the spectral properties of block triangular preconditioner have been studied in [39–41].

According to whether the total degree on the simplex points is odd or even, we can reorder the matrix A and B_T as the following:

$$A = \begin{bmatrix} D_1 & W^\intercal \\ W & D_2 \end{bmatrix}, \quad B_T = \begin{bmatrix} D_1 & 0 \\ W & D_2 \end{bmatrix},$$

where recall that

$$D_1 = I_1 \otimes K_0, \quad D_2 = I_2 \otimes K_0,$$

$$W = \left(\sum_{k=1}^{m} G_k \otimes K_k \right)(\mathcal{I}_2, \mathcal{I}_1) = \sum_{k=1}^{m} G_k(\mathcal{I}_2, \mathcal{I}_1) \otimes K_k,$$

and $\mathcal{I}_1$ is the one in $\{\mathcal{I}_{odd}, \mathcal{I}_{even}\}$ with larger cardinality, $\mathcal{I}_2$ is the one with smaller cardinality. By direct computation,

$$B_T^{-1} = \begin{bmatrix} D_1^{-1} & 0 \\ -D_2^{-1} W D_1^{-1} & D_2^{-1} \end{bmatrix}.$$

The following lemmas are useful. Here $\sigma(A) = \{\lambda \mid \det(A - \lambda I) = 0\}$ is the spectrum of A and $\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda|$ is the spectral radius.

Lemma 4 ([40]). *The eigenvalues of AB_T^{-1} are positive numbers, and the spectrum satisfies*

$$\sigma(AB_T^{-1}) \subset \{1\} \cup \sigma(S, D_1)$$

where $S = D_1 - W^\intercal D_2^{-1} W$ is the Schur complement of D_2 in A, and $\sigma(S, D_1)$ contains the eigenvalues μ corresponding to the generalized eigenvalue problem

$$Sz = \mu D_1 z. \tag{31}$$

Proof. Let $\mu \in \sigma(AB_T^{-1})$ and $(v_1, v_2)^\intercal$ is the corresponding eigenvector. Then we have

$$AB_T^{-1} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} I - W^\intercal D_2^{-1} W D_1^{-1} & W^\intercal D_2^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \mu \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

It can be written as

$$\begin{cases} (I - W^\intercal D_2^{-1} W D_1^{-1}) v_1 + (W^\intercal D_2^{-1}) v_2 = \mu v_1 \\ 0 = (1 - \mu) v_2 \end{cases}$$

If $\mu = 1$, then $(v_1, v_2)^\intercal$ satisfies $W^\intercal D_2^{-1} v_2 = W^\intercal D_2^{-1} W D_1^{-1} v_1$. If $\mu \neq 1$, from the first equation we get $(I - W^\intercal D_2^{-1} W D_1^{-1}) v_1 = \mu v_1$ which implies that $\mu \in \sigma(S, D_1)$. □

Remark 3. The eigenvalue 1 has multiplicity $N_x \cdot (|\mathcal{I}_1| - |\mathcal{I}_2|)$. Having eigenvalue 1 with multiplicity $N_x \cdot (|\mathcal{I}_1| - |\mathcal{I}_2|)$ is advantageous when applying the Krylov subspace iterative methods as one iteration is enough to take care of these well centered eigenvalues.

Lemma 5 ([42]). *For a rectangular matrix B,*

$$\rho(B^\intercal B) = \left\| \begin{bmatrix} 0 & B^\intercal \\ B & 0 \end{bmatrix} \right\|^2.$$

Next, we give the main result of this section.

Theorem 3. *The spectrum of the preconditioned Galerkin matrix AB_T^{-1} satisfies*

$$\sigma(AB_T^{-1}) \subset (1 - \hat{\tau}^2, 1],$$

where $\hat{\tau}$ is defined in (19).

Proof. By Theorem 1 , it suffices to show that $\sigma(S, D_1) \subset (1 - \hat{\tau}^2, 1]$. For the upper bound, let μ be an eigenvalue of the generalized eigenvalue problem (31), and z be the corresponding eigenvector,

$$\begin{aligned} & (D_1 - W^\intercal D_2^{-1} W) z = \mu D_1 z \\ \Rightarrow\ & (1 - \mu) D_1 z = W^\intercal D_2^{-1} W z \end{aligned}$$

$$\Rightarrow\ (1-\mu)z^{\intercal}D_1 z = (Wz)^{\intercal}D_2^{-1}(Wz)$$
$$\Rightarrow\ 1-\mu = \frac{(Wz)^{\intercal}D_2^{-1}(Wz)}{z^{\intercal}D_1 z}, \tag{32}$$

by the symmetric positive definiteness of D_1 and D_2^{-1}, we conclude that

$$1-\mu \geq 0 \ \ \Rightarrow \ \ \mu \leq 1.$$

For the lower bound, we introduce the change of variables $\tilde{z} := D_1^{1/2} z$ in the Rayleigh quotient (32) and arrive at

$$\begin{aligned}\mu &= 1 - \frac{\tilde{z}^{\intercal}B^{\intercal}B\tilde{z}}{\tilde{z}^{\intercal}\tilde{z}}\\ &\geq 1-\rho(B^{\intercal}B)\\ &= 1-\left\|\begin{bmatrix} 0 & B^{\intercal}\\ B & 0\end{bmatrix}\right\|^2\\ &= 1-\hat{\tau}^2,\end{aligned}$$

where $B = \sum_{k=1}^{m} G_k(\mathcal{I}_2,\mathcal{I}_1)\otimes(K_0^{-\frac{1}{2}}K_k K_0^{-\frac{1}{2}}) = D_2^{-1/2}WD_1^{-1/2}$ introduced in Section 4.1 is the off-diagonal of the preconditioned system $B_D^{-\frac{1}{2}}AB_D^{-\frac{1}{2}}$. The third equation holds by Lemma 5 and the last equation holds by (23).

Therefore, we conclude that $\mu \in (1-\hat{\tau}^2, 1]$. □

We emphasize the bound for the triangular block preconditioner is better than the one for the block diagonal preconditioner since $\hat{\tau}$ is always between 0 and 1 and thus $0 < 1-\hat{\tau} < 1-\hat{\tau}^2 < 1$.

6. Numerical results

In this section, we evaluate the performance of the block triangular preconditioner B_T and compare with the block diagonal preconditioner B_D and Kronecker product preconditioner B_K. We use linear or quadratic finite element space as X_h and assemble the stiffness matrices K_k, $k = 0, 1, 2, \ldots, m$ using the software package *i*FEM [43]. We apply the MINRES method with tolerance $= 10^{-8}$ for B_D and apply GMRES method with same tolerance and restart $= 20$ for B_T and B_K. Diagonal matrices are inverted by the geometric multigrid method with one V-cycle $V(2,2)$, i.e., two pre-smoothing and two post-smoothing steps in each V-cycle. For the Kronecker product preconditioner $B_K = L\otimes K_0$, the matrix L is inverted by MATLAB's backslash operator. All the time costs (in seconds) are for solver time only, set up time not included. All computations are performed using MATLAB 2017b on a Lenove-Y40 with Intel Core i7-4510U CPU and 8 GB of RAM.

For each example, we present extremal eigenvalues for the un-preconditioned system A as well as the preconditioned system $P^{-1}A$ with various preconditioners; see Tables 1, 4, 8, 11, and 15. The extremal eigenvalues for $B_T^{-1}A$ lie closer around one compared to the other two preconditioned systems $B_D^{-1}A$ and $B_K^{-1}A$. We also present iteration steps and time cost with various h, p, and m. All preconditioners are robust to h and iteration steps are positively correlated to p and m. Under the same h, p, m values, the block triangular preconditioner B_T outperforms the block diagonal preconditionerB_D and B_K with CPU time ratio 0.5 and 0.65, respectively.

6.1. A 2-D homogeneous boundary problem

First, we look at a 2-D benchmark problem taken from [19,44]. Let $D = (-0.5, 0.5)\times(-0.5, 0.5)$. The 2-D covariance function is given by (5) with $L = 1$. Let $\bar{a}(x) = 1$, and $f(x_1, x_2) = 2(0.5 - x_1^2 - x_2^2)$.

6.1.1. Gaussian Random variables with Hermite polynomials (P1-element)

Table 1 is the extremal eigenvalues for the un-preconditioned system A as well as the preconditioned systems. Tables 2 and 3 show the results of iteration steps and time cost with various h, p, and m under Hermite polynomials and piecewise linear (P1) finite element space as X_h. All preconditioners are robust to h and iteration steps are positively correlated to p and m. Under the same h, p, m values, the triangular preconditioner B_T performs better than the Kronecker product B_K and the mean-based preconditioner B_D. When p, m are large, GMRES with preconditioner B_T converges almost twice faster than MINRES with preconditioner B_D.

6.1.2. Uniform random variables with Legendre polynomials (P1-element)

The extremal eigenvalues for the un-preconditioned and preconditioned systems are summarized in Table 4. Tables 5 and 6 show the results of iteration steps and time cost with various h, p, and m under Legendre polynomials and piecewise linear (P1) finite element space as X_h. All preconditioners are robust to h and iteration steps are positively correlated to p and m. Under the same h, p, m values, the triangular preconditioner B_T performs better than the Kronecker product B_K and the mean-based preconditioner B_D. When p, m are large, the computational cost of GMRES with preconditioner B_T is around 0.65 of that of MINRES using B_K and 0.5 of that with preconditioner B_D.

Table 1
Extremal Eigenvalues of A and $P^{-1}A$ with Different Preconditioners P ($\mu = 1,\ \sigma = 0.3,\ h = 1/8$, Gaussian Random Variables with Hermite Polynomials, P1-element).

m	p	A		$B_D^{-1}A$		$B_K^{-1}A$		$B_T^{-1}A$	
		$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$
4	2	0.1847	11.1479	0.5294	1.4706	0.6944	1.3593	0.7785	1.2215
	3	0.1424	12.3212	0.3658	1.6342	0.5669	1.5461	0.5978	1.4022
	4	0.1047	13.3309	0.2239	1.7761	0.4256	1.7797	0.3976	1.6024
6	2	0.1841	11.2045	0.5235	1.4765	0.6561	1.4019	0.7730	1.2270
	3	0.1416	12.3930	0.3579	1.6421	0.5144	1.6097	0.5877	1.4123
	4	0.1036	13.4153	0.2143	1.7857	0.3629	1.8687	0.3826	1.6174

Table 2
Iteration Steps with Different Preconditioners ($\mu = 1,\ \sigma = 0.3$, Gaussian Random Variables with Hermite Polynomials, P1-element).

m	h	B_D			B_K			B_T		
		$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$
4	1/16	13	17	24	10	12	15	6	8	11
	1/32	13	18	25	10	12	15	6	8	11
	1/64	13	18	25	10	12	15	6	8	11
	1/128	13	18	25	10	12	15	6	8	11
6	1/16	11	17	23	10	13	17	6	8	10
	1/32	12	18	24	10	13	17	6	8	10
	1/64	13	18	25	10	13	17	6	8	10
	1/128	13	18	26	10	13	17	6	8	10

Table 3
Time Costs with Different Preconditioners ($\mu = 1, \sigma = 0.3$, Gaussian Random Variables with Hermite Polynomials, P1-element).

m	h	B_D			B_K			B_T		
		$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$
4	1/16	0.1	0.2	1.2	0.2	0.2	0.8	0.1	0.1	0.4
	1/32	0.4	1.1	4.4	0.3	0.8	2.8	0.3	0.7	1.5
	1/64	1.3	6.3	19.3	1.2	4.7	13.5	0.9	2.2	7.9
	1/128	8.8	29.4	80.4	7.9	23.5	57.5	4.8	14.5	43.1
6	1/16	0.1	1.0	2.8	0.2	0.8	2.6	0.1	0.5	1.7
	1/32	0.7	3.8	16.6	0.5	3.1	10.6	0.4	1.9	6.8
	1/64	2.6	16.6	59.5	2.3	14.7	48.3	1.4	7.8	28.7
	1/128	16.7	69.8	248.4	15.9	63.6	195.6	9.4	36.8	118.2

Table 4
Extremal Eigenvalues of A and $P^{-1}A$ with Different Preconditioners P ($\mu = 1,\ \sigma = 0.7,\ h = 1/8$, Uniform Random Variables with Legendre Polynomials, P1-element).

m	p	A		$B_D^{-1}A$		$B_K^{-1}A$		$B_T^{-1}A$	
		$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$
4	2	0.1744	11.6305	0.4301	1.5699	0.6022	1.4954	0.6752	1.3248
	3	0.1559	12.3074	0.2931	1.7069	0.4543	1.6995	0.5003	1.4997
	4	0.1436	12.7364	0.1986	1.8014	0.3285	1.8676	0.3577	1.6423
6	2	0.1730	11.7466	0.4026	1.5974	0.5418	1.5636	0.6431	1.3569
	3	0.1535	12.4745	0.2440	1.7560	0.3626	1.8119	0.4285	1.5715
	4	0.1392	12.9380	0.1273	1.8727	0.2016	2.0298	0.2384	1.7616

Table 7 shows the iteration steps and time costs for different preconditioners with various variances σ. We can find that the iteration steps increase as σ increases for all preconditioners. Again the block triangular preconditioner B_T outperforms the block diagonal preconditionerB_D and B_K with CPU time ratio 0.5 and 0.65, respectively.

6.1.3. GaussIan random variables with Hermite polynomials (P2-element)

The extremal eigenvalues for the un-preconditioned and preconditioned systems are summarized in Table 8.

Tables 9 and 10 show the results of iteration steps and time cost with various h, p, and m under Hermite polynomials and quadratic (P2) finite element space as X_h. All preconditioners are robust to h and iteration steps are positively correlated to p and m. Under the same h, p, m values, the triangular preconditioner B_T performs better than the Kronecker product B_K and the mean-based preconditioner B_D. Again the block triangular preconditioner B_T outperforms the block diagonal preconditionerB_D and B_K with CPU time ratio 0.5 and 0.65, respectively.

Table 5
Iteration Steps with Different Preconditioners ($\mu = 1$, $\sigma = 0.7$, Uniform Random Variables with Legendre Polynomials, P1-element).

m	h	B_D			B_K			B_T		
		$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$
4	1/16	16	22	29	12	15	19	7	10	12
	1/32	16	22	30	12	15	19	7	10	12
	1/64	16	23	30	12	15	19	7	10	12
	1/128	16	23	30	12	15	18	7	10	12
6	1/16	17	25	37	13	17	26	8	11	16
	1/32	17	25	39	13	17	26	8	11	16
	1/64	17	25	40	13	17	26	8	11	16
	1/128	17	26	40	13	17	26	8	11	16

Table 6
Time Costs with Different Preconditioners ($\mu = 1, \sigma = 0.7$, Uniform Random Variables with Legendre Polynomials, P1-element).

m	h	B_D			B_K			B_T		
		$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$
4	1/16	0.3	0.8	1.4	0.3	0.5	1.0	0.3	0.2	0.5
	1/32	0.5	1.4	5.5	0.5	1.0	3.6	0.4	0.7	1.6
	1/64	1.8	7.8	23.4	1.4	5.7	17.1	0.9	2.7	8.9
	1/128	10.8	37.4	98.0	9.5	29.5	68.7	5.5	18.1	46.1
6	1/16	0.2	1.5	4.7	0.2	1.2	3.8	0.1	0.7	2.5
	1/32	0.8	5.1	26.3	0.7	4.0	19.2	0.5	2.6	12.4
	1/64	4.2	29.0	112.0	3.5	22.6	87.2	2.4	13.1	54.6
	1/128	28.5	126.2	480.7	24.5	98.5	351.2	15.0	66.2	191.9

Table 7
Iteration Steps and Time Costs with Different Preconditioners and various σ ($\mu = 1$, $p = 4$, $m = 6$, $h = 1/64$, Uniform Random Variables with Legendre Polynomials, P1-element).

σ	B_D		B_K		B_T	
	Iter	Time	Iter	Time	Iter	Time
0.1	7	21.0	6	20.6	4	14.0
0.3	12	35.1	9	31.4	6	21.2
0.5	19	48.3	13	38.5	8	25.1
0.7	40	95.5	26	74.1	16	47.2

Table 8
Extremal Eigenvalues of A and $P^{-1}A$ with Different Preconditioners P ($\mu = 1$, $\sigma = 0.3$, $h = 1/8$, Gaussian Random Variables with Hermite Polynomials, P2-element).

m	p	A		$B_D^{-1}A$		$B_K^{-1}A$		$B_T^{-1}A$	
		$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$
4	2	0.0465	15.3835	0.5240	1.4760	0.6843	1.3747	0.7734	1.2266
	3	0.0358	17.0475	0.3584	1.6416	0.5506	1.5740	0.5884	1.4116
	4	0.0263	18.4825	0.2148	1.7852	0.4036	1.8249	0.3834	1.6166
6	2	0.0464	15.4837	0.5159	1.4841	0.6463	1.4211	0.7657	1.2343
	3	0.0356	17.1762	0.3476	1.6524	0.4990	1.6415	0.5744	1.4256
	4	0.0260	18.6336	0.2016	1.7984	0.3397	1.9161	0.3625	1.6375

6.1.4. Uniform random variables with Legendre polynomials (P2-element)

Table 11 provides the extremal eigenvalues for the un-preconditioned and preconditioned systems with different preconditioners.

Tables 12 and 13 show the results of iteration steps and time cost with various h, p, and m under Legendre polynomials and quadratic (P2) finite element space as X_h. All preconditioners are robust to h and iteration steps are positively correlated to p and m. Under the same h, p, m values, the triangular preconditioner B_T performs better than the Kronecker product B_K and the mean-based preconditioner B_D. Again the block triangular preconditioner B_T outperforms the block diagonal preconditionerB_D and B_K with CPU time ratio 0.5 and 0.65, respectively.

Table 14 shows the iteration steps and time costs for different preconditioners with various variances σ. We can find that the iteration steps increase as σ increases for all preconditioners. The conclusion is still true: the block triangular preconditioner B_T outperforms the block diagonal preconditionerB_D and B_K with CPU time ratio 0.5 and 0.65, respectively.

Table 9
Iteration Steps with Different Preconditioners ($\mu = 1$, $\sigma = 0.3$, Gaussian Random Variables with Hermite Polynomials, P2-element).

m	h	B_D			B_K			B_T		
		$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$
4	1/8	13	17	24	10	12	15	6	8	11
	1/16	13	18	25	10	12	15	6	8	11
	1/32	13	18	25	10	12	15	6	8	11
	1/64	13	18	25	10	12	15	6	8	11
6	1/8	11	17	23	10	13	17	6	8	10
	1/16	12	18	25	10	13	17	6	8	10
	1/32	13	18	25	10	13	17	6	8	10
	1/64	13	18	26	10	13	17	6	8	10

Table 10
Time Costs with Different Preconditioners ($\mu = 1, \sigma = 0.3$, Gaussian Random Variables with Hermite Polynomials, P2-element).

m	h	B_D			B_K			B_T		
		$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$
4	1/8	0.2	0.3	2.0	0.1	0.3	1.3	0.0	0.1	0.7
	1/16	0.5	1.2	5.8	0.4	1.1	3.9	0.4	0.8	2.8
	1/32	2.6	9.6	28.0	2.0	6.4	22.4	1.3	3.3	13.7
	1/64	13.4	42.7	118.0	11.8	33.2	83.8	7.1	21.2	61.4
6	1/8	0.2	1.5	4.0	0.2	1.2	3.5	0.1	0.8	2.5
	1/16	0.8	5.1	44.0	0.8	4.4	29.4	0.6	2.8	22.4
	1/32	5.1	20.7	72.8	3.8	17.2	58.8	2.8	10.0	35.5
	1/64	21.1	88.7	310.5	19.2	75.8	238.5	11.9	46.8	149.2

Table 11
Extremal Eigenvalues of A and $P^{-1}A$ with Different Preconditioners P ($\mu = 1$, $\sigma = 0.7$, $h = 1/8$, Uniform Random Variables with Legendre Polynomials, P2-element).

m	p	A		$B_D^{-1}A$		$B_K^{-1}A$		$B_T^{-1}A$	
		$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$
4	2	0.0439	16.1900	0.4209	1.5791	0.5866	1.5166	0.6646	1.3354
	3	0.0393	17.3396	0.2798	1.7202	0.4319	1.7259	0.4813	1.5187
	4	0.0361	18.0996	0.1818	1.8182	0.3004	1.8955	0.3306	1.6694
6	2	0.0436	16.3732	0.3903	1.6097	0.5273	1.5892	0.6283	1.3717
	3	0.0386	17.6129	0.2268	1.7732	0.3393	1.8424	0.4022	1.5978
	4	0.0350	18.4923	0.1061	1.8939	0.1697	2.0597	0.2009	1.7991

Table 12
Iteration Steps with Different Preconditioners ($\mu = 1$, $\sigma = 0.7$, Uniform Random Variables with Legendre Polynomials, P2-element).

m	h	B_D			B_K			B_T		
		$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$
4	1/8	16	22	29	12	15	19	7	10	12
	1/16	16	22	30	12	15	19	7	10	12
	1/32	16	23	30	12	15	19	7	10	12
	1/64	16	23	30	12	15	18	7	10	12
6	1/8	17	25	37	13	17	26	8	12	16
	1/16	17	25	39	13	17	26	8	11	16
	1/32	17	25	40	13	17	26	8	11	16
	1/64	17	25	40	13	17	26	8	11	16

6.2. A 3-D homogeneous boundary problem

Now we look at a 3-D homogeneous boundary problem. Let $D = (-0.5, 0.5)^3$. The 3-D covariance function is given by (5) with $L = 1$. The mean value of the random coefficient $\bar{a}(x) = 1$, and the right hand side $f(x_1, x_2, x_3) = 2(0.5 - x_1^2 - x_2^2 - x_3^2)$.

We use a uniform random variable in the truncated Karhunen–Loève expansion and corresponding Legendre polynomials as the gPC basis. Now we compare the performance of the mean-based block preconditioner B_D, the Kronecker product preconditioner B_K and the block triangular preconditioner B_T. Table 15 provides the extremal eigenvalues for the un-preconditioned and preconditioned systems with these preconditioners.

Table 13
Time Costs with Different Preconditioners ($\mu = 1, \sigma = 0.7$, Uniform Random Variables with Legendre Polynomials, P2-element).

m	h	B_D			B_K			B_T		
		$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$	$p=2$	$p=3$	$p=4$
4	1/8	0.3	0.4	2.9	0.2	0.4	1.5	0.1	0.3	0.7
	1/16	0.9	2.4	6.6	0.6	1.3	5.9	0.5	1.1	2.2
	1/32	2.6	10.4	29.7	2.0	7.9	21.4	1.5	3.6	11.7
	1/64	13.4	53.9	133.8	11.9	43.4	87.1	7.0	25.3	59.4
6	1/8	0.3	2.2	6.2	0.2	1.5	4.6	0.3	1.0	3.2
	1/16	1.1	6.3	28.2	1.0	5.0	21.3	0.7	3.1	12.9
	1/32	4.9	36.9	145.3	4.8	25.9	106.7	3.4	18.4	71.1
	1/64	39.1	163.5	648.7	32.3	125.6	480.8	20.0	83.0	301.9

Table 14
Iteration Steps and Time Costs with Different Preconditioners and various σ ($\mu = 1$, $p = 4$, $m = 6$, $h = 1/32$, Uniform Random Variables with Legendre Polynomials, P2-element).

σ	B_D		B_K		B_T	
	Iter	Time	Iter	Time	Iter	Time
0.1	7	24.7	6	24.1	4	15.4
0.3	12	55.1	9	39.2	6	31.7
0.5	19	59.7	13	45.8	8	34.5
0.7	40	120.9	26	90.8	16	58.6

Table 15
Extremal Eigenvalues of A and $P^{-1}A$ with Different Preconditioners P (Legendre polynomials with $h = 1/8$, $\mu = 1$, $m = 10$, three dimensional P1-element).

p	A		$B_D^{-1}A$		$B_K^{-1}A$		$B_T^{-1}A$	
	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$	$\lambda_{\min}$	$\lambda_{\max}$
2	0.0352	2.1185	0.4378	1.5622	0.5105	1.7807	0.6840	1.3160
3	0.0219	2.2674	0.2696	1.7304	0.3398	2.1698	0.4666	1.5334
4	0.0114	2.3699	0.1374	1.8626	0.1825	2.5423	0.2558	1.7442

Table 16
Iteration steps and time costs with different m and p for both preconditioners (Legendre polynomials with $h = 1/16$, $\mu = 1$, $\sigma = 0.7$, three dimensional P1-element).

p	B_D		B_K		B_T	
	Iter	Time	Iter	Time	Iter	Time
2	17	28.6	12	22.0	7	14.1
3	25	181.1	16	128.3	10	80.6
4	39	980.5	24	712.9	14	452.2

From Table 16 we can see that similar to the 2-D example, the block triangular preconditioner B_T outperforms the Kronecker product preconditioner B_K and the mean-based preconditioner B_D. Our conclusion on the comparison of preconditioners: the block triangular preconditioner B_T outperforms the block diagonal preconditionerB_D and B_K with CPU time ratio 0.5 and 0.65, respectively, is further confirmed by the 3-D example.

7. Conclusions

In this paper, we study block preconditioners for the coupled linear systems resulting from the stochastic Galerkin discretizations of elliptic equations with random coefficients. The proposed block triangular preconditioner utilizes the block sparsity and hierarchical structure of the stochastic Galerkin matrix. The preconditioner is solved inexactly by geometric multigrid V-cycle and applied to Krylov subspace methods such as GMRES. Numerical results demonstrate that the block triangular preconditioner is computationally more efficient than the block diagonal preconditioner, especially when the random coefficient has large variance. We also give theoretical bounds for the spectrum of the preconditioned system.

Acknowledgments

Long Chen would like to acknowledge the support by NSF DMS-1418934, DMS-1913080, and DMS-2012465. Guang Lin gratefully acknowledges the support from the National Science Foundation, United States (DMS-1555072, DMS-1736364, CMMI-1634832, and CMMI-1560834), Brookhaven National Laboratory, United States Subcontract 382247, ARO/MURI

grant W911NF-15-1-0562, and U.S. Department of Energy (DOE) Office of Science Advanced Scientific Computing Research Program DE-SC-0021142. Bin Zheng would like to acknowledge the support by Beijing Institute for Scientific and Engineering Computing (BISEC) and a Laboratory Directed Research and Development (LDRD) Program, United States from Pacific Northwest National Laboratory.

References

[1] X. Ma, N. Zabaras, A stochastic mixed finite element heterogeneous multiscale method for flow in porous media, J. Comput. Phys. 230 (12) (2011) 4696–4722.
[2] O.P.L. Maître, O.M. Knio, H.N. Najm, R.G. Ghanem, A stochastic projection method for fluid flow: I. Basic formulation, J. Comput. Phys. 173 (2) (2001) 481–511.
[3] O.P.L. Maître, M.T. Reagan, H.N. Najm, R.G. Ghanem, O.M. Knio, A stochastic projection method for fluid flow: II. Random process, J. Comput. Phys. 181 (1) (2002) 9–44.
[4] D. Xiu, D. Lucor, C.-H. Su, G. Karniadakis, Stochastic modeling of flow-structure interactions using generalized polynomial chaos, ASME J. Fluids Eng. 124 (1) (2001) 51–59.
[5] M.A. Tatang, W. Pan, R.G. Prinn, G.J. McRae, An efficient method for parametric uncertainty analysis of numerical geophysical models, J. Geophys. Res.: Atmos. 102 (D18) (1997) 21925–21932.
[6] L. Mathelin, M.Y. Hussaini, T.A. Zang, Stochastic approaches to uncertainty quantification in CFD simulations, Numer. Algorithms 38 (1) (2005) 209–236.
[7] I. Babuška, F. Nobile, R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, SIAM J. Numer. Anal. 45 (3) (2007) 1005–1034.
[8] D. Xiu, J. Hesthaven, High-order collocation methods for differential equations with random inputs, SIAM J. Sci. Comput. 27 (3) (2005) 1118–1139.
[9] R.G. Ghanem, P.D. Spanos, Stochastic Finite Element: A Spectral Approach, Springer, 1991.
[10] D. Xiu, J. Shen, Efficient stochastic Galerkin methods for random diffusion equations, J. Comput. Phys. 228 (2) (2009) 266–281.
[11] H.C. Elman, C.W. Miller, E.T. Phipps, R.S. Tuminaro, Assessment of collocation and Galerkin approaches to linear diffusion equations with random data, Int. J. Uncertain. Quantif. 1 (1) (2011) 19–33.
[12] J. Bäck, F. Nobile, L. Tamellini, R. Tempone, Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: A numerical comparison, in: J.S. Hesthaven, E.M. Rønquist (Eds.), Spectral and High Order Methods for Partial Differential Equations, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 43–62.
[13] N. Wiener, The homogeneous chaos, Amer. J. Math. 60 (4) (1938) 897–936.
[14] D. Xiu, G. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2) (2002) 619–644.
[15] E. Rosseel, S. Vandewalle, Iterative solvers for the stochastic finite element method, SIAM J. Sci. Comput. 32 (1) (2010) 372–397.
[16] R. Tipireddy, E.T. Phipps, R.G. Ghanem, A comparison of solution methods for stochastic partial differential equations, in: CSRI Summer Proceedings, 2010, pp. 79–90.
[17] R.G. Ghanem, R.M. Kruger, Numerical solution of spectral stochastic finite element systems, Comput. Methods Appl. Mech. Engrg. 129 (3) (1996) 289–303.
[18] M. Pellissetti, R. Ghanem, Iterative solution of systems of linear equations arising in the context of stochastic finite elements, Adv. Eng. Softw. 31 (8) (2000) 607–616.
[19] C.E. Powell, H.C. Elman, Block-diagonal preconditioning for spectral stochastic finite-element systems, IMA J. Numer. Anal. 29 (2) (2009) 350–375.
[20] E. Ullmann, A kronecker product preconditioner for stochastic Galerkin finite element discretizations, SIAM J. Sci. Comput. 32 (2) (2010) 923–946.
[21] O.P.L. Maître, O. Knio, B. Debusschere, H. Najm, R. Ghanem, A multigrid solver for two-dimensional stochastic diffusion equations, Comput. Methods Appl. Mech. Engrg. 192 (41) (2003) 4723–4744.
[22] B. Seynaeve, E. Rosseel, B. Nicolaï, S. Vandewalle, Fourier mode analysis of multigrid methods for partial differential equations with random coefficients, J. Comput. Phys. 224 (1) (2007) 132–149.
[23] H.C. Elman, D. Furnival, Solving the stochastic steady-state diffusion problem using multigrid, IMA J. Numer. Anal. 27 (4) (2007) 675–688.
[24] E. Rosseel, T. Boonen, S. Vandewalle, Algebraic multigrid for stationary and time-dependent partial differential equations with stochastic coefficients, Numer. Linear Algebra Appl. 15 (2–3) (2008) 141–163.
[25] K. Phoon, H. Huang, S. Quek, Simulation of strongly non-Gaussian processes using Karhunen–Loeve expansion, Probab. Eng. Mech. 20 (2) (2005) 188–198.
[26] X. Wan, G. Karniadakis, Solving elliptic problems with non-Gaussian spatially-dependent random coefficients, Comput. Methods Appl. Mech. Engrg. 198 (21) (2009) 1985–1995, Advances in Simulation-Based Engineering Sciences – Honoring J. Tinsley Oden.
[27] R. Ghanem, Stochastic finite elements with multiple random non-Gaussian properties, J. Eng. Mech. 125 (1) (1999) 26–40.
[28] B. Puig, F. Poirion, C. Soize, Non-Gaussian simulation using Hermite polynomial expansion: convergences and algorithms, Probab. Eng. Mech. 17 (3) (2002) 253–264.
[29] P. Frauenfelder, C. Schwab, R.A. Todor, Finite elements for elliptic problems with stochastic coefficients, Comput. Methods Appl. Mech. Engrg. 194 (2) (2005) 205–228, Selected papers from the 11th Conference on The Mathematics of Finite Elements and Applications.
[30] I. Babuška, P. Chatzipantelidis, On solving elliptic stochastic partial differential equations, Comput. Methods Appl. Mech. Engrg. 191 (37) (2002) 4093–4122.
[31] B. Øksendal, Stochastic Differential Equations: An Introduction with Applications, 5th ed., Springer, Berlin, 1998.
[32] I. Babuska, R. Tempone, G. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, SIAM J. Numer. Anal. 42 (2) (2004) 800–825.
[33] D.G. Feingold, R.S. Varga, Block diagonally dominant matrices and generalizations of the gerschgorin circle theorem, Pacific J. Math. 12 (1962) 1241–1250.
[34] H. Salas, Gershgorin's theorem for matrices of operators, Linear Algebra Appl. 291 (1) (1999) 15–36.
[35] C. Tretter, Spectral Theory of Block Operator Matrices and Applications, Imperial College Press, 2008.
[36] Y. Saad, M. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Stat. Comput. 7 (3) (1986) 856–869.
[37] R. Blaheta, GPCG–generalized preconditioned CG method and its use with non-linear and non-symmetric displacement decomposition preconditioners, Numer. Linear Algebra Appl. 9 (6–7) (2002) 527–550.
[38] C.F. Van Loan, N. Pitsianis, Approximation with kronecker products, in: Linear Algebra for Large Scale and Real-Time Applications, Springer Netherlands, Dordrecht, 1993, pp. 293–314.
[39] A. Klawonn, Block-triangular preconditioners for saddle point problems with a penalty term, SIAM J. Sci. Comput. 19 (1) (1998) 172–184.

[40] V. Simoncini, Block triangular preconditioners for symmetric saddle-point problems, Appl. Numer. Math. 49 (1) (2004) 63–80, Numerical Algorithms, Parallelism and Applications.
[41] O. Axelsson, R. Blaheta, Preconditioning of matrices partitioned in 2×2 block form: eigenvalue estimates and Schwarz DD for mixed FEM, Numer. Linear Algebra Appl. 17 (5) (2010) 787–810.
[42] V. Paulsen, Completely Bounded Maps and Operator Algebras, Cambridge Studies in Advanced Mathematics (78), Cambridge University Press, 2003.
[43] L. Chen, iFEM: an Integrated Finite Element Methods Package in MATLAB, Technical report, University of California at Irvine, 2009.
[44] M.K. Deb, I. Babuška, J. Oden, Solution of stochastic partial differential equations using Galerkin finite element techniques, Comput. Methods Appl. Mech. Engrg. 190 (48) (2001) 6359–6372, URL http://www.sciencedirect.com/science/article/pii/S0045782501002377.