INDUSTRIAL AND SYSTEMS ENGINEERING

ISE

Stochastic Galerkin Method And Hierarchical Preconditioning For PDE-Constrained Optimization

ZHENDONG LI[1], AKWUM ONWUNTA[1], AND BEDŘICH SOUSEDÍK[2]

[1]Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA, 18015 USA

[2]Department of Mathematics and Statistics, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250 USA

ISE Technical Report 25T-024

LEHIGH UNIVERSITY

STOCHASTIC GALERKIN METHOD AND HIERARCHICAL PRECONDITIONING FOR PDE-CONSTRAINED OPTIMIZATION *

ZHENDONG LI†, AKWUM ONWUNTA†, AND BEDŘICH SOUSEDÍK‡

Abstract. We develop efficient hierarchical preconditioners for optimal control problems governed by partial differential equations with uncertain coefficients. Adopting a discretize-then-optimize framework that integrates finite element discretization, stochastic Galerkin approximation, and advanced time-discretization schemes, the approach addresses the challenge of large-scale, ill-conditioned linear systems arising in uncertainty quantification. By exploiting the sparsity inherent in generalized polynomial chaos expansions, we derive hierarchical preconditioners based on truncated stochastic expansion that strike an effective balance between computational cost and preconditioning quality. Numerical experiments demonstrate that the proposed preconditioners significantly accelerate the convergence of iterative solvers compared to existing methods, providing robust and efficient solvers for both steady-state and time-dependent optimal control applications under uncertainty.

1. Introduction. Optimal control problems governed by partial differential equations (PDEs) arise in numerous applications, including fluid mechanics, structural optimization, and inverse problems. These problems have been extensively studied over the past decades. For a theoretical overview and computational methods related to deterministic problems, we refer readers to, e.g., [9, 27]. In many practical applications, the PDE coefficients are uncertain. Such uncertainties originate from various sources, including measurement errors, model approximations, and environmental variations, and they are modeled as random variables or stochastic processes. Recently, there has been an increased interest in optimal control problems governed by PDEs with random coefficients, see e.g., [17]. These stochastic problems are inherently more complex than their deterministic counterparts, thus necessitating specialized numerical methods.

Two alternative strategies are used for the optimal control problems: optimize-then-discretize and discretize-then-optimize. The optimize-then-discretize approach involves deriving continuous optimality conditions and then discretizing them. Conversely, the discretize-then-optimize approach discretizes the objective and the PDE first and then solves the resulting discrete optimization problem. The discretize-then-optimize is widely used in practice, because it allows the employment of efficient numerical methods such as finite element or finite difference methods, see e.g., [1, 3, 4, 5, 9]. This approach is also used in this study.

In the formulation of numerical methods, discretizations in the stochastic space, spatial domain, and potentially time domain are required. For stochastic-space discretization, several methods exist, including Monte Carlo method, stochastic collocation method and stochastic Galerkin method. The stochastic collocation method discretizes random variables using a set of collocation points, solving the resulting deterministic PDEs at these points [8, 13, 12, 26]. The stochastic Galerkin method expands the solution in terms of orthogonal polynomials and solves the resulting coupled system of equations [10, 14, 15, 16, 21, 30].

Monte Carlo method is simple and therefore popular; however, it typically demands a large number of samples for acceptable accuracy, making it computationally

*Dedicated to the memory of Professor Howard C. Elman, our dear colleague and mentor. We also want to thank Lehigh's High Performance Computing systems for providing computational resources.

†Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA 18015 (zhendong.li@lehigh.edu, ako221@lehigh.edu).

‡Department of Mathematics and Statistics, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250 (sousedik@umbc.edu).

intensive for high-dimensional problems. The stochastic collocation method is in general more efficient, but it may lose efficiency in high-dimensional random spaces, and it becomes challenging to implement in practice. Here, we use the stochastic Galerkin method, which systematically captures uncertainties by employing orthonormal polynomial expansions, preserving optimal convergence properties while improving computational efficiency for large-scale stochastic PDE-constrained optimization problems.

In practice, it is common to combine temporal discretization (e.g., backward Euler scheme), stochastic expansions (e.g., generalized polynomial chaos expansions), and spatial discretization (e.g., finite element method). Specifically, let $t_0, t_1, \ldots, t_{N_t}$ denote a partition of the time interval of interest, $\phi_i(\mathbf{x})_{i=1}^{N_h}$ represent a spatial discretization basis and $\psi_k(\xi)_{k=1}^{N_\xi}$ represent an orthonormal polynomial basis in the random space. Then the state $y(t, \mathbf{x}, \xi)$ can be approximated by

$$y(t, \mathbf{x}, \xi) \approx \sum_{n=0}^{N_t} \sum_{i=1}^{N_h} \sum_{k=1}^{N_\xi} y_{n,i,k}\, \Theta_n(t)\, \phi_i(\mathbf{x})\, \psi_k(\xi),$$

where $\Theta_n(t)$ is a temporal basis, and $y_{n,i,k}$ are expansion coefficients (i.e., the degrees of freedom of the numerical solution) in the basis $\{\Theta_n(t)\, \phi_i(\mathbf{x})\, \psi_k(\xi)\}$, which are determined by a suitable numerical method.

The discretize-then-optimize approach leads to large-scale linear systems obtained via finite element (or possibly finite difference) discretizations, which are then typically solved using Krylov subspace methods, e.g., by the generalized minimal residual method (GMRES). These linear systems are often ill-conditioned, which causes a slow convergence of the iterative method. To address this issue, we construct preconditioners that improve the convergence of iterative solvers. We note that the development of efficient solvers is a general challenge in optimal control, including for deterministic problems solved via duality-based approaches [28].

In this paper, we introduce a hierarchical preconditioning framework specifically tailored for stochastic PDE-constrained optimal control problems. Although the core concepts are inspired by preconditioners for PDE problems [6, 24, 25, 31], an application to the Karush-Kuhn-Tucker (KKT) systems arising from optimization problems and an extension to time-dependent problems are nontrivial and constitutes a primary contribution of this work. We provide a systematic derivation of the preconditioner for both steady-state and all-at-once formulation of time-dependent problems. The method is supported by a rigorous spectral analysis, proving that the proposed preconditioner is spectrally equivalent to the ideal but computationally prohibitive exact Schur complement. The performance is then evaluated using a set of numerical experiments.

The paper is organized as follows. In Section 2, we address the steady-state stochastic optimal control problem. We introduce the problem formulation and discretization into a large-scale KKT system. Then, we construct the hierarchical Gauss-Seidel preconditioner for PDE-constrained optimal control problems (hGSoc). In Section 3, we extend the framework to the time-dependent case. We present an all-at-once discretization and we develop a corresponding parallel-in-time preconditioner. In Section 4, we provide a spectral analysis of the proposed preconditioners. In Section 5, we demonstrate the efficiency of the proposed methods by a series of numerical experiments. Finally, in Section 6 we conclude and summarize our work.

2. Steady-state problem. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a complete probability space, where Ω is the sample space, $\mathcal{F}$ is the σ-algebra generated by Ω, and $\mathcal{P}$ is the proba-

bility measure. We assume that the uncertainty in our model arises from a vector of independent random variables $\xi = \left(\xi_1, \ldots, \xi_{m_\xi}\right)^T$ defined as $\xi : \Omega \to \Phi \subset \mathbb{R}^{m_\xi}$. Let $\mathcal{B}(\Phi)$ be the Borel σ-algebra on Φ induced by ξ, and let μ denote the corresponding probability measure. The expectation of the product of measurable functions u and v depending on ξ defines a Hilbert space $L^2(\Phi) \coloneqq L^2(\Phi, \mathcal{B}(\Phi), \mu)$ equipped with the inner product

$$\langle u, v \rangle = \mathbb{E}[uv] = \int_\Phi u(\xi)\, v(\xi)\; d\mu(\xi),$$

where $\mathbb{E}$ denotes the mathematical expectation.

We consider a stochastic PDE coefficient $\Bbbk(x, \rho)$, where x is the coordinate in the physical space, and ρ represents random inputs. It is assumed that $\Bbbk(x, \rho)$ is bounded away from zero and infinity, i.e., $0 < \Bbbk_{\min} \le \Bbbk(x, \rho) \le \Bbbk_{\max} < \infty$ for some constants $\Bbbk_{\min}$, $\Bbbk_{\max}$, and it can be represented by a truncated Karhunen-Loève(KL) expansion

$$\Bbbk(x, \rho) \approx \kappa_0(x) + \sum\nolimits_{i=1}^{m_\xi} \sqrt{\theta_i}\kappa_i(x)\xi_i(\rho), \tag{2.1}$$

where $\kappa_0(x)$ is the mean function, and $(\theta_i, \kappa_i(x))$ is the i-th eigenpair (eigenvalue and eigenfunctions, respectively) of the covariance function

$$C_\Bbbk(x, y) = \mathbb{E}\left[\left(\Bbbk(x, \rho) - \kappa_0(x)\right)\left(\Bbbk(y, \rho) - \kappa_0(y)\right)\right], \tag{2.2}$$

defined in physical domain $\mathcal{D} \subset \mathbb{R}^n$ with piecewise continuous boundary $\partial\mathcal{D}$, such that

$$\int_\mathcal{D} C_\Bbbk(x, y)\kappa_i(y)dy = \theta_i\kappa_i(x). \tag{2.3}$$

The system of eigenfunctions $\kappa_i(x)$ is orthonormal in the space $L^2(\mathcal{D})$, and the eigenvalues θ_i are arranged in a nonincreasing order. We next introduce the random variables $\xi_i(\rho)$. They are centered, normalized, and uncorrelated. In that they satisfy

$$\mathbb{E}[\xi_i] = 0, \quad \mathbb{E}[\xi_i\xi_j] = \delta_{ij}, \tag{2.4}$$

where δ_{ij} is the Kronecker delta. Specifically, they are defined as

$$\xi_i(\rho) = \frac{1}{\sqrt{\theta_i}} \int_\mathcal{D} \left(\Bbbk(x, \rho) - \kappa_0(x)\right) \kappa_i(x)\, dx.$$

We will assume that the variables ξ_i are independent, identically distributed (i.i.d.).

2.1. Problem formulation. We consider the steady-state optimal control problem given by

$$\min_{y,u} J(y, u) := \frac{1}{2}\int_\Phi \int_\mathcal{D} |y - y_d|^2 dx\, d\mu(\xi) + \frac{\beta}{2}\int_\Phi \int_\mathcal{D} |u|^2\, dx\, d\mu(\xi) + \frac{\gamma}{2}\int_\mathcal{D} |\sigma(y)|^2\, dx, \tag{2.5}$$

subject to

$$\begin{cases} -\nabla \cdot \left(\Bbbk(x, \xi)\, \nabla y(x, \xi)\right) = u(x, \xi), & \text{in } \mathcal{D} \times \Phi, \\ y(x, \xi) = g(x), & \text{on } \partial\mathcal{D} \times \Phi, \end{cases} \tag{2.6}$$

where y is the state, y_d is the target state, and u is the (distributed) control. The parameter γ penalizes the variance $\sigma^2(y)$ of the state y, which is defined as

$$\sigma(y) = \sqrt{\mathbb{E}\left[(y - \mathbb{E}[y])^2\right]} = \left[\int_\Phi |y - \mathbb{E}[y]|^2 d\mu(\xi)\right]^{1/2}. \tag{2.7}$$

Observe that both the state y and control u are stochastic. In view of the Doob-Dynkin lemma (see, e.g., [2]), both y and u admit the same parametric dependence on ξ. In computations, we work with a finite dimensional subspace $\mathcal{T}_p \subset L^2(\Phi, \mathcal{B}(\Phi), \mu)$, spanned by a set of generalized polynomial chaos (gPC) functions $\{\psi_\ell(\xi)\}_{\ell=1}^{N_\xi}$ with

$$\langle \psi_\ell, \psi_k \rangle = \mathbb{E}[\psi_\ell \psi_k] = \int_\Phi \psi_\ell(\xi)\psi_k(\xi) d\mu = \delta_{\ell k}, \tag{2.8}$$

$\psi_1(\xi) = 1$, and $\mathbb{E}[\psi_k(\xi)] = 0$ for $k > 1$. Considering polynomials of total degree p and the number of random variables m_ξ, the dimension of $\mathcal{T}_p$ is $N_\xi = \binom{m_\xi + p}{p}$. After the application of the stochastic Galerkin finite element discretization to (2.5)–(2.6), both the state y and control u are expanded as

$$v = \sum_{k=1}^{N_\xi} \sum_{i=1}^{N_h} v_{ik}\phi_i(x)\psi_k(\xi), \qquad v = y \text{ or } u. \tag{2.9}$$

Using the KLE as given by (2.1) in the finite element discretization of the PDE constraint (2.6) yields the stiffness matrices A_ℓ associated with the expansion coefficients,

$$A_\ell = [(A_\ell)_{ab}], \quad (A_\ell)_{ab} = \int_{\mathcal{D}} \sqrt{\theta_\ell}\kappa_\ell(x)\nabla\phi_a(x) \cdot \nabla\phi_b(x)\, dx, \quad a, b = 1, 2, \ldots, N_h. \tag{2.10}$$

We note that all matrices A_ℓ share the same sparsity pattern. The global stiffness matrix is given by $\mathcal{A} = \sum_{\ell=1}^{n_A} H_\ell \otimes A_\ell$, where the matrices H_ℓ are defined below.

In implementation, we use the *matricized* format, which utilizes the isomorphism between $\mathbb{R}^{N_h N_\xi}$ and $\mathbb{R}^{N_h \times N_\xi}$, defined via the operators vec and mat. Specifically,

$$\bar{V} = \text{mat}(\bar{v}) = \left[v_1, v_2, \ldots, v_{N_\xi}\right] \in \mathbb{R}^{N_h \times N_\xi}, \tag{2.11}$$

where the column k contains the coefficients associated with the basis function ψ_k, and $\bar{v} = \text{vec}\left(\bar{V}\right) \in \mathbb{R}^{N_h N_\xi}$. We will use lowercase letters for the *vectorized* representation and uppercase letters for the *matricized* counterpart so, e.g., $\bar{R} = \text{mat}(\bar{r})$, etc.

We will also use the notation

$$H_\ell = [h_{\ell,jk}], \quad h_{\ell,jk} \equiv \mathbb{E}[\psi_\ell \psi_j \psi_k], \quad \ell = 1, \ldots, n_A, \; j, k = 1, \ldots, N_\xi,$$

where we note that all matrices H_ℓ are symmetric, and

$$H^\sigma = \text{diag}\,(0, h_{1,jj}), \quad j = 2, \ldots, N_\xi,$$

which is obtained from H_1 by setting $h_{1,11} = 0$. We note that in our settings $H_1 = I_{N_\xi}$, i.e., it is an identity matrix of size N_ξ. Finally, we denote

$$\mathcal{M} = H_1 \otimes M, \qquad \mathcal{M}_\sigma = H^\sigma \otimes M, \tag{2.12}$$

where M is the mass matrix. Using the stochastic Galerkin framework to discretize problem (2.5)–(2.6), we get

$$\min_{\mathbf{y},\mathbf{u}} J(\mathbf{y}, \mathbf{u}) = \frac{1}{2}(\mathbf{y} - \mathbf{y}_d)^T \mathcal{M}(\mathbf{y} - \mathbf{y}_d) + \frac{\gamma}{2}\mathbf{y}^T \mathcal{M}_\sigma \mathbf{y} + \frac{\beta}{2}\mathbf{u}^T \mathcal{M}\mathbf{u},$$

subject to

$$-\mathcal{A}\mathbf{y} + \mathcal{M}\mathbf{u} = \mathbf{g}. \tag{2.13}$$

The Lagrangian corresponding to this problem is

$$\mathcal{L}(\mathbf{y}, \mathbf{u}, \lambda) = \frac{1}{2}(\mathbf{y} - \mathbf{y}_d)^T \mathcal{M}(\mathbf{y} - \mathbf{y}_d) + \frac{\gamma}{2}\mathbf{y}^T \mathcal{M}_\sigma \mathbf{y} + \frac{\beta}{2}\mathbf{u}^T \mathcal{M}\mathbf{u} + \lambda^T(-\mathcal{A}\mathbf{y} + \mathcal{M}\mathbf{u} - \mathbf{g}). \tag{2.14}$$

Applying the first-order optimality conditions to (2.14), using $H^\gamma = H_1 + \gamma H^\sigma$, and

$$\mathcal{M}_\gamma = H^\gamma \otimes M, \qquad \mathcal{A} = \sum\nolimits_{\ell=1}^{n_A} H_\ell \otimes A_\ell,$$

we get the KKT system of equations, which can be written in the matrix form as

$$\begin{bmatrix} \mathcal{M}_\gamma & 0 & -\mathcal{A}^T \\ 0 & \beta\mathcal{M} & \mathcal{M}^T \\ -\mathcal{A} & \mathcal{M} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathcal{M}\mathbf{y}_d \\ 0 \\ \mathbf{g} \end{bmatrix}. \tag{2.15}$$

The stochastic Galerkin matrix in (2.15) is symmetric, indefinite, and in general very large. It is ill-conditioned, and therefore, a good preconditioner is required to solve the system efficiently. Next, we introduce a preconditioner to tackle this problem.

2.2. Schur complement based preconditioner. A block-diagonal preconditioner for (2.15) is given by (see, e.g. Benner et al. [5]):

$$\mathcal{P} := \begin{bmatrix} \mathcal{M}_\gamma & 0 & 0 \\ 0 & \beta\mathcal{M} & 0 \\ 0 & 0 & \mathcal{S}_{\text{exact}} \end{bmatrix}, \tag{2.16}$$

where $\mathcal{S}_{\text{exact}}$ is the exact Schur complement

$$\mathcal{S}_{\text{exact}} = \mathcal{A}\mathcal{M}_\gamma^{-1}\mathcal{A}^T + \frac{1}{\beta}\mathcal{M}. \tag{2.17}$$

The first two blocks corresponding to (scaling of) the mass matrix are block diagonal, and the inverses are approximated by Chebyshev semi-iteration listed as Algorithm 5. However, forming and applying the inverse of $\mathcal{S}_{\text{exact}}$ is computationally prohibitive. The primary difficulty stems from its additive structure. Therefore, the key to an efficient solution lies in designing an approximation $\mathcal{S}$ that is spectrally close to $\mathcal{S}_{\text{exact}}$ and easy to invert. Following the approach in [5], we employ the approximation

$$\mathcal{S} = \mathcal{Z}\mathcal{M}_\gamma^{-1}\mathcal{Z}^T, \qquad \mathcal{Z} = \mathcal{A} + \sqrt{\frac{1+\gamma}{\beta}}\mathcal{M} = \sum_{\ell=1}^{n_A} H_\ell \otimes \tilde{A}_\ell, \tag{2.18}$$

where

$$\tilde{A}_1 = A_1 + \sqrt{\frac{1+\gamma}{\beta}}M, \qquad \tilde{A}_\ell = A_\ell, \quad \ell = 2, \ldots, n_A.$$

Since $\mathcal{Z}$ is symmetric, we have $\mathcal{S}^{-1} = \mathcal{Z}^{-1}\mathcal{M}_\gamma\mathcal{Z}^{-1}$. In [5], the authors studied a mean-based preconditioner derived from (2.18) by dropping all the terms of $\mathcal{Z}$ except the first term; that is, $\mathcal{Z} \approx H_1 \otimes \tilde{A}_1$. However, this mean-based preconditioner is

only effective when the variability in the model is relatively small, but struggles as the standard deviation increases.

In this study, we overcome this shortcoming by using hierarchical preconditioning introduced in [24, 25] in the context of forward problems. More specifically, unlike [24, 25], we extend this strategy to the more challenging setting – stochastic optimal control problem, and rigorously study the spectral analysis of the new preconditioner. To this end, we formulate the preconditioners in the matricized format. First, we rewrite the linear system (2.15) using this format. Using (2.11) and the identity

$$(V \otimes W)\,\mathrm{vec}(X) = \mathrm{vec}(WXV^T), \tag{2.19}$$

the system (2.15), can be equivalently written, cf. also [5, eq. (53)], as

$$\begin{aligned} M\bar{Y}H^\gamma - \sum_{\ell=1}^{n_A} A^T\bar{\Lambda}H_\ell &= M\bar{Y}_d H_1, \\ \beta M\bar{U}H_1 + M^T\bar{\Lambda}H_1 &= 0, \\ -\sum_{\ell=1}^{n_A} A^T\bar{Y}H_\ell + M\bar{U}H_1 &= \bar{G}. \end{aligned} \tag{2.20}$$

Finally, we note that a matrix-vector product with $\mathcal{Z}$ from (2.18) can be written as

$$\mathcal{Z}\bar{v} = \mathrm{vec}\left(\textstyle\sum_{\ell=1}^{n_A} \tilde{A}_\ell \bar{V} H_\ell\right). \tag{2.21}$$

2.3. Block-diagonal hierarchical Gauss-Seidel preconditioner. We first recall the preconditioner for the forward PDE problem from [24]. However, in this work, we present it in the matricized format as Algorithm 1. Moreover, we formulate the corresponding preconditioner for the optimal control problem as Algorithm 2–3. To set the notation, we will denote by $\bar{V}_{(i:n)}$ a submatrix of $\bar{V}$ containing columns $i, i+1, \ldots, n$, and, in particular, $\bar{V} = \bar{V}_{(1:n_\xi)}$. There are two components of the preconditioner. The first component consists of block-diagonal solves with blocks of varying sizes. The second component is used in the setup of the right-hand sides for the solves, and consists of matrix-vector products by certain subblocks of the stochastic Galerkin matrix by vectors of corresponding sizes. To this end, we will write $\left[h_{t,(\ell)(k)}\right]$, with (ℓ) and (k) denoting a set of (consecutive) rows and columns of matrix H_t so that, in particular, $H_t = \left[h_{t,(1:n_\xi)(1:n_\xi)}\right]$. Let us also denote $\bar{v}_{(\ell)} = \mathrm{vec}(\bar{V}_{(\ell)})$. Then, the matrix-vector products can be written, cf. (2.19) and noting the symmetry of H_t, as

$$\bar{v}_{(\ell)} = \textstyle\sum_{t\in\mathcal{I}_t} (\left[h_{t,(\ell)(k)}\right] \otimes \tilde{A}_t)\bar{u}_{(k)} \quad \Leftrightarrow \quad \bar{V}_{(\ell)} = \sum_{t\in\mathcal{I}_t} \tilde{A}_t \bar{U}_{(k)} \left[h_{t,(k)(\ell)}\right], \tag{2.22}$$

where $\mathcal{I}_t$ is an index set $\mathcal{I}_t \subseteq \{1, \ldots, n_\xi\}$ indicating that the matrix-vector products may be truncated. Possible strategies for truncation are discussed in [24]. In this study, we use $\mathcal{I}_t = \{1, \ldots, n_t\}$ with $n_t = \binom{m_\xi + p_t}{p_t}$ for some $p_t \le p$. In particular, we set $t = \{0, 1, 2\}$ and with $\mathcal{I}_t = \emptyset$ both preconditioners in Algorithms 1 and 2–3 reduce to mean-based variants. We also note that, since the initial guess is zero in Algorithm 1, the multiplications by $\mathcal{F}_1$ and $\mathcal{F}_{d+1}$ vanish from (2.23)–(2.24).

Next, we apply the hGS strategy to form a preconditioner for the KKT system. The preconditioner is formulated as Algorithm 2–3. It adapts the core idea of Algorithm 1 to handle the coupled variables corresponding to the state $\bar{V}^Y$, control $\bar{V}^U$,

Algorithm 1 [24, Algorithm 3] Hierarchical Gauss-Seidel preconditioner (hGS)

The preconditioner $\mathcal{Z}_{hGS} : \bar{R} \longmapsto \bar{V}$ is defined as follows.

1: Set the initial solution $\bar{V}$ to zero and update in the following steps:
2: Solve

$$\tilde{A}_1 \bar{V}_{(1)} = \bar{R}_{(1)} - \mathcal{F}_1, \qquad \text{where } \mathcal{F}_1 = \sum_{t \in \mathcal{I}_t} \tilde{A}_t \bar{V}_{(2:N_\xi)} \left[h_{t,(2:N_\xi)(1)} \right]. \tag{2.23}$$

3: **for** $d = 1, \ldots, p-1$ **do**
4: Set $\ell = (n_\ell + 1 : n_u)$, where $n_\ell = \binom{m_\xi + d - 1}{d-1}$ and $n_u = \binom{m_\xi + d}{d}$.
5: Solve

$$\tilde{A}_1 \bar{V}_{(\ell)} = \bar{R}_{(\ell)} - \mathcal{E}_{d+1} - \mathcal{F}_{d+1}, \tag{2.24}$$

where

$$\mathcal{E}_{d+1} = \sum_{t \in \mathcal{I}_t} \tilde{A}_t \bar{V}_{(1:n_\ell)} \left[h_{t,(1:n_\ell)(\ell)} \right], \qquad \mathcal{F}_{d+1} = \sum_{t \in \mathcal{I}_t} \tilde{A}_t \bar{V}_{(n_u+1:N_\xi)} \left[h_{t,(n_u+1:N_\xi)(\ell)} \right].$$

6: **end for**
7: Set $\ell = (n_u + 1 : n_\xi)$.
8: Solve

$$\tilde{A}_1 \bar{V}_{(\ell)} = \bar{R}_{(\ell)} - \mathcal{E}_{p+1}, \qquad \text{where } \mathcal{E}_{p+1} = \sum_{t \in \mathcal{I}_t} \tilde{A}_t \bar{V}_{(1:n_u)} \left[h_{t,(1:n_u)(\ell)} \right],$$

9: **for** $d = p-1, \ldots, 1$ **do**
10: Set $\ell = (n_\ell + 1 : n_u)$, where $n_\ell = \binom{m_\xi + d - 1}{d-1}$ and $n_u = \binom{m_\xi + d}{d}$.
11: Solve (2.24).
12: **end for**
13: Solve (2.23).

and adjoint $\bar{V}^\Lambda$ simultaneously at each hierarchical level. A key computational step of Algorithm 2–3 entails a solve with an approximation $\tilde{\mathcal{P}}$ of the deterministic KKT system. Since this system is relatively small and constant across all hierarchical levels, it can be handled efficiently, for instance, by computing a direct factorization of $\tilde{\mathcal{P}}$ once and reusing it for all subsequent solves. The overall performance of the preconditioner is thus determined by the cost of these deterministic solves and the number of truncated off-diagonal matrix-vector products.

3. Time-dependent problem.

3.1. Problem formulation. The time-dependent optimal control problem is given by

$$\begin{aligned} \min_{y,u} \mathcal{J}(y,u) \quad = \quad & \frac{1}{2} \int_0^T \int_\Phi \int_\mathcal{D} |y - y_d|^2 dx \, d\mu(\xi) \, dt + \frac{\beta}{2} \int_0^T \int_\Phi \int_\mathcal{D} |u|^2 dx \, d\mu(\xi) \, dt \\ & + \frac{\gamma}{2} \int_0^T \int_\Phi \int_\mathcal{D} |\sigma(y)|^2 dx \, d\mu(\xi) \, dt, \end{aligned} \tag{3.1}$$

Algorithm 2 hGS preconditioner for the optimal control problem (hGSoc)

The preconditioner $\mathcal{P}_{hGSoc} : \left(\bar{R}^Y, \bar{R}^U, \bar{R}^\Lambda\right) \longmapsto \left(\bar{V}^Y, \bar{V}^U, \bar{V}^\Lambda\right)$ is defined as follows.

1: Set the initial solution $\left(\bar{V}^Y, \bar{V}^U, \bar{V}^\Lambda\right)$ to zero and update in the following steps:

2: Solve

$$\tilde{P} \begin{bmatrix} \bar{V}^Y_{(1)} \\ \bar{V}^U_{(1)} \\ \bar{V}^\Lambda_{(1)} \end{bmatrix} = \begin{bmatrix} \bar{R}^Y_{(1)} - \mathcal{C}_1 + \mathcal{D}_1 \\ \bar{R}^U_{(1)} - \mathcal{E}_1 - \mathcal{F}_1 \\ \bar{R}^\Lambda_{(1)} + \mathcal{G}_1 - \mathcal{H}_1 \end{bmatrix}, \tag{2.25}$$

where

$$\mathcal{C}_1 = M\bar{V}^Y_{(2:n_\xi)}\left[h^\alpha_{(2:n_\xi)(1)}\right], \qquad \mathcal{D}_1 = \sum_{t\in\mathcal{I}_t} A_t \bar{V}^\Lambda_{(2:n_\xi)}\left[h_{t,(2:n_\xi)(1)}\right],$$
$$\mathcal{E}_1 = \beta M\bar{V}^U_{(2:n_\xi)}\left[h_{1,(2:n_\xi)(1)}\right], \qquad \mathcal{F}_1 = M\bar{V}^\Lambda_{(2:n_\xi)}\left[h_{1,(2:n_\xi)(1)}\right],$$
$$\mathcal{G}_1 = \sum_{t\in\mathcal{I}_t} A_t \bar{V}^Y_{(2:n_\xi)}\left[h_{t,(2:n_\xi)(1)}\right], \qquad \mathcal{H}_1 = M\bar{V}^U_{(2:n_\xi)}\left[h_{1,(2:n_\xi)(1)}\right].$$

3: **for** $d = 1, \ldots, p-1$ **do**

4: Set $\ell = (n_\ell + 1 : n_u)$, where $n_\ell = \binom{m_\xi+d-1}{d-1}$ and $n_u = \binom{m_\xi+d}{d}$.

5: Solve

$$\tilde{P} \begin{bmatrix} \bar{V}^Y_{(\ell)} \\ \bar{V}^U_{(\ell)} \\ \bar{V}^\Lambda_{(\ell)} \end{bmatrix} = \begin{bmatrix} \bar{R}^Y_{(\ell)} - \mathcal{C}_{d+1} + \mathcal{D}_{d+1} \\ \bar{R}^U_{(\ell)} - \mathcal{E}_{d+1} - \mathcal{F}_{d+1} \\ \bar{R}^\Lambda_{(\ell)} + \mathcal{G}_{d+1} - \mathcal{H}_{d+1} \end{bmatrix}, \tag{2.26}$$

where

$$\mathcal{C}_{d+1} = M\left(\bar{V}^Y_{(1:n_\ell)}\left[h^\alpha_{(1:n_\ell)(\ell)}\right] + \bar{V}^Y_{(n_u+1:n_\xi)}\left[h^\alpha_{(n_u+1:n_\xi)(\ell)}\right]\right),$$
$$\mathcal{D}_{d+1} = \sum_{t\in\mathcal{I}_t} A_t\left(\bar{V}^\Lambda_{(1:n_\ell)}\left[h_{t,(1:n_\ell)(\ell)}\right] + \bar{V}^\Lambda_{(n_u+1:n_\xi)}\left[h_{t,(n_u+1:n_\xi)(\ell)}\right]\right),$$
$$\mathcal{E}_{d+1} = \beta M\left(\bar{V}^U_{(1:n_\ell)}\left[h_{1,(1:n_\ell)(\ell)}\right] + \bar{V}^U_{(n_u+1:n_\xi)}\left[h_{1,(n_u+1:n_\xi)(\ell)}\right]\right),$$
$$\mathcal{F}_{d+1} = M\left(\bar{V}^\Lambda_{(1:n_\ell)}\left[h_{1,(1:n_\ell)(\ell)}\right] + \bar{V}^\Lambda_{(n_u+1:n_\xi)}\left[h_{1,(n_u+1:n_\xi)(\ell)}\right]\right),$$
$$\mathcal{G}_{d+1} = \sum_{t\in\mathcal{I}_t} A_t\left(\bar{V}^Y_{(1:n_\ell)}\left[h_{t,(1:n_\ell)(\ell)}\right] + \bar{V}^Y_{(n_u+1:n_\xi)}\left[h_{t,(n_u+1:n_\xi)(\ell)}\right]\right),$$
$$\mathcal{H}_{d+1} = M\left(\bar{V}^U_{(1:n_\ell)}\left[h_{1,(1:n_\ell)(\ell)}\right] + \bar{V}^U_{(n_u+1:n_\xi)}\left[h_{1,(n_u+1:n_\xi)(\ell)}\right]\right).$$

6: **end for**

subject to

$$\begin{cases} \dfrac{\partial y(t,\mathbf{x},\xi)}{\partial t} - \nabla\cdot(\Bbbk(\mathbf{x},\xi)\nabla y(t,\mathbf{x},\xi)) = u(t,\mathbf{x},\xi) \text{ in } (0,T]\times\mathcal{D}\times\Phi, \\ y(t,\mathbf{x},\xi) = g \text{ on } (0,T]\times\partial\mathcal{D}\times\Phi, \\ y(0,\mathbf{x},\xi) = y_0 \text{ in } \mathcal{D}\times\Phi. \end{cases} \tag{3.2}$$

Algorithm 3 hGS preconditioner for the optimal control problem (hGSoc), cont'd

7: Set $\ell = (n_u + 1 : n_\xi)$.
8: Solve

$$\tilde{P} \begin{bmatrix} \bar{V}^Y_{(\ell)} \\ \bar{V}^U_{(\ell)} \\ \bar{V}^\Lambda_{(\ell)} \end{bmatrix} = \begin{bmatrix} \bar{R}^Y_{(\ell)} - \mathcal{C}_{p+1} + \mathcal{D}_{p+1} \\ \bar{R}^U_{(\ell)} - \mathcal{E}_{p+1} - \mathcal{F}_{p+1} \\ \bar{R}^\Lambda_{(\ell)} + \mathcal{G}_{p+1} - \mathcal{H}_{p+1} \end{bmatrix},$$

where

$$\mathcal{C}_{p+1} = M\bar{V}^Y_{(1:n_u)} \left[h^\alpha_{(1:n_u)(\ell)}\right], \qquad \mathcal{D}_{p+1} = \sum_{t\in\mathcal{I}_t} A_t \bar{V}^\Lambda_{(1:n_u)} \left[h_{t,(1:n_u)(\ell)}\right],$$
$$\mathcal{E}_{p+1} = \beta M\bar{V}^U_{(1:n_u)} \left[h_{1,(1:n_u)(\ell)}\right], \qquad \mathcal{F}_{p+1} = M\bar{V}^\Lambda_{(1:n_u)} \left[h_{1,(1:n_u)(\ell)}\right],$$
$$\mathcal{G}_{p+1} = \sum_{t\in\mathcal{I}_t} A_t \bar{V}^Y_{(1:n_u)} \left[h_{t,(1:n_u)(\ell)}\right], \qquad \mathcal{H}_{p+1} = M\bar{V}^U_{(1:n_u)} \left[h_{1,(1:n_u)(\ell)}\right].$$

9: **for** $d = p-1, \ldots, 1$ **do**
10: Set $\ell = (n_\ell + 1 : n_u)$, where $n_\ell = \binom{m_\xi+d-1}{d-1}$ and $n_u = \binom{m_\xi+d}{d}$.
11: Solve (2.26).
12: **end for**
13: Solve (2.25).

After the application of the stochastic Galerkin finite element discretization to (3.1), and using the trapezoidal rule for the time discretization, where $N_t = T/\tau$ is the number of time steps over the interval $[0, T]$ with time-step size τ, we obtain

$$\min_{\mathbf{y},\mathbf{u}} \mathcal{J}(\mathbf{y}, \mathbf{u}) = \frac{\tau}{2}(\mathbf{y} - \mathbf{y}_d)^T (D \otimes \mathcal{M}_\gamma)(\mathbf{y} - \mathbf{y}_d) + \frac{\tau\beta}{2}\mathbf{u}^T (D \otimes \mathcal{M})\mathbf{u}, \tag{3.3}$$

where $\mathcal{M}$ and $\mathcal{M}_\gamma$ are defined in (2.12), D in (3.7) below, $\mathbf{y}$, $\mathbf{y}_d$, and $\mathbf{u}$ are vectors corresponding to the state, desired state, and control, respectively, that contain concatenated vectors $\mathbf{y}_i, \mathbf{y_d}_i, \mathbf{u}_i \in \mathbb{R}^{N_h N_\xi \times 1}, i = 1, \ldots, N_t$, due to the time-stepping,

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{N_t} \end{bmatrix}, \quad \mathbf{y}_d = \begin{bmatrix} \mathbf{y_d}_1 \\ \vdots \\ \mathbf{y_d}_{N_t} \end{bmatrix}, \quad \text{and } \mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_{N_t} \end{bmatrix}.$$

After the application of the stochastic Galerkin finite element discretization to (3.2), and using the implicit Euler method for the time discretization, we obtain

$$M\mathbf{y}_k + \tau A\mathbf{y}_k = M\mathbf{y}_{k-1} + \tau M\mathbf{u}_k. \tag{3.4}$$

Combining all time steps of (3.4) in all-at-once discretization ([18, 20]), we can write

$$\mathcal{A}_t\mathbf{y} - \tau\mathcal{N}\mathbf{u} = \left[\mathcal{M}\mathrm{y}_0\,, \quad 0\,, \quad \ldots, 0\right]^T.$$

where

$$\mathcal{A}_t = \begin{bmatrix} \mathcal{L} & & & \\ -\mathcal{M} & \mathcal{L} & & \\ & \ddots & \ddots & \\ & & -\mathcal{M} & \mathcal{L} \end{bmatrix}, \qquad \mathcal{N} = \begin{bmatrix} \mathcal{M} & & & \\ & \mathcal{M} & & \\ & & \ddots & \\ & & & \mathcal{M} \end{bmatrix},$$

and

$$\mathcal{L} = H_1 \otimes (M + \tau A_1) + \tau \sum_{\ell=2}^{n_A} H_\ell \otimes A_\ell. \tag{3.5}$$

The matrices $\mathcal{A}_t$ and $\mathcal{N}$ can be constructed using Kronecker product as

$$\mathcal{A}_t = (I_{N_t} \otimes \mathcal{L}) - (C \otimes \mathcal{M}), \qquad \mathcal{N} = I_{N_t} \otimes H_1 \otimes M, \tag{3.6}$$

where the matrix C and matrix D, used in (3.3) and also below, are defined as

$$C = \begin{bmatrix} 0 & & & \\ -1 & 0 & & \\ & \ddots & \ddots & \\ & & -1 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} \frac{1}{2} & & & \\ & 1 & & \\ & & \ddots & \\ & & & \frac{1}{2} \end{bmatrix}. \tag{3.7}$$

Forming the Lagrangean and applying the first-order optimality conditions, we get

$$\begin{bmatrix} \tau D \otimes \mathcal{M}_\gamma & 0 & -\mathcal{A}_t^T \\ 0 & \beta\tau D \otimes \mathcal{M} & \tau \mathcal{N}^T \\ -\mathcal{A}_t & \tau \mathcal{N} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \lambda \end{bmatrix} = \begin{bmatrix} \tau (D \otimes \mathcal{M}) \cdot (\mathbf{1}_{N_t} \otimes \mathbf{y}_d) \\ \mathbf{0} \\ \mathbf{d} \end{bmatrix}, \tag{3.8}$$

where $\mathbf{d} = \begin{bmatrix} \mathcal{M}\mathrm{y}_0 + \mathbf{g} & \mathbf{g} & \cdots & \mathbf{g} \end{bmatrix}$, and $\mathbf{1}_{N_t} \in \mathbb{R}^{N_t \times 1}$ is the column all-ones vector.

3.2. PINT-based block-diagonal hierarchical Gauss-Seidel preconditioner. In analogy to (2.16), we propose a preconditioner for (3.8) as

$$\mathcal{P}_{\text{hGSoc-PINT}} \approx \begin{bmatrix} \tau D \otimes \mathcal{M}_\gamma & & \\ & \tau\beta D \otimes \mathcal{M} & \\ & & \bar{\mathcal{S}} \end{bmatrix}, \tag{3.9}$$

where $\bar{\mathcal{S}}$ is a computationally efficient approximation of the exact Schur complement

$$\bar{\mathcal{S}}_{\text{exact}} = \frac{1}{\tau} \mathcal{A}_t (D \otimes \mathcal{M}_\gamma)^{-1} \mathcal{A}_t^T + \frac{\beta}{\tau} \mathcal{N} (D \otimes \mathcal{M})^{-1} \mathcal{N}^T. \tag{3.10}$$

As before, the first two blocks correspond to (scaling of) the mass matrix are approximated by Chebyshev semi-iteration from Algorithm 5. Since the iteration entails matrix-vector multiplications, we note that using (2.19) we have

$$(\tau D \otimes H_1 \otimes M)\, \mathbf{v_1} = \tau \begin{bmatrix} \frac{1}{2} \operatorname{vec}(M V_{11} H_\gamma) \\ \operatorname{vec}(M V_{12} H_\gamma) \\ \vdots \\ \frac{1}{2} \operatorname{vec}(M V_{1N_t} H_\gamma) \end{bmatrix}, \tag{3.11}$$

$$(\beta\tau D \otimes H_1 \otimes M)\, \mathbf{v_2} = \beta\tau \begin{bmatrix} \frac{1}{2} \operatorname{vec}(M V_{21} H_1) \\ \operatorname{vec}(M V_{22} H_1) \\ \vdots \\ \frac{1}{2} \operatorname{vec}(M V_{2N_t} H_1) \end{bmatrix}, \tag{3.12}$$

where $\mathbf{v_1}$ and $\mathbf{v_2}$ are the vectors obtained by concatenating $\mathbf{v_{1i}}$ and $\mathbf{v_{2i}}$, $i = 1, \ldots, N_t$, respectively, which correspond to the time steps. The matrices V_{1i}, $V_{2i} \in \mathbb{R}^{N_h \times N_\xi}$ are

then the *matricized* counterparts of $\mathbf{v_{1i}}$ and $\mathbf{v_{2i}}$, respectively. Next, since inverting $\bar{\mathcal{S}}_{\text{exact}}$ is computationally prohibitive, we propose an approximation as

$$\bar{\mathcal{S}} = \frac{1}{\tau} \underbrace{\left(\mathcal{A}_t + \tau\sqrt{\frac{1+\gamma}{\beta}}\mathcal{N}\right)}_{=:\bar{\mathcal{Z}}} (D \otimes \mathcal{M}_\gamma)^{-1} \left(\mathcal{A}_t + \tau\sqrt{\frac{1+\gamma}{\beta}}\mathcal{N}\right)^T, \tag{3.13}$$

where using (3.5) and (3.6), we can rewrite $\bar{\mathcal{Z}}$ as

$$\bar{\mathcal{Z}} = I_{N_t} \otimes \left(H_1 \otimes (M + \tau A_1) + \tau \sum_{\ell=2}^{n_A} H_\ell \otimes A_\ell\right) + \tau\sqrt{\frac{1+\gamma}{\beta}} I_{N_t} \otimes H_1 \otimes M - C \otimes H_1 \otimes M. \tag{3.14}$$

By dropping the last term $C \otimes H_1 \otimes M$, we further approximate $\bar{\mathcal{Z}}$ by

$$\tilde{\mathcal{Z}} = I_{N_t} \otimes \left\{H_1 \otimes \left[(1 + \tau\sqrt{\frac{1+\gamma}{\beta}})M + \tau A_1\right] + \tau \sum_{\ell=2}^{n_A} H_\ell \otimes A_\ell\right\}, \tag{3.15}$$

which is symmetric. The idea is to use $\bar{\mathcal{S}} \approx \tilde{\mathcal{Z}} (D \otimes \mathcal{M}_\gamma)^{-1} \tilde{\mathcal{Z}}$, and in particular the solves with $\tilde{\mathcal{Z}}$ are approximated by Algorithm 1, similarly to the steady-state case. We remark that by dropping all terms with $\ell > 2$ from $\tilde{\mathcal{Z}}$ in (3.15), that is considering

$$\tilde{\mathcal{Z}}_0 = I_{N_t} \otimes H_1 \otimes \left[(1 + \tau\sqrt{\frac{1+\gamma}{\beta}})M + \tau A_1\right],$$

we recover the mean-based preconditioner [5]. Since the application of Algorithm 1 entails matrix-vector multiplications, using (2.19) we formulate $\tilde{\mathcal{Z}}\mathbf{v_3}$ as

$$\begin{aligned} \tilde{\mathcal{Z}}\mathbf{v_3} &= \left(I_{N_t} \otimes \left\{H_1 \otimes \left[(1 + \tau\sqrt{\frac{1+\gamma}{\beta}})M + \tau A_1\right] + \tau \sum_{\ell=2}^{N} H_\ell \otimes A_\ell\right\}\right) \mathbf{v_3} \\ &= \begin{bmatrix} \sum_{\ell=1}^{n_A} \text{vec}\left(\hat{A}_\ell V_{31} H_\ell\right) \\ \sum_{\ell=1}^{n_A} \text{vec}\left(\hat{A}_\ell V_{32} H_\ell\right) \\ \vdots \\ \sum_{\ell=1}^{n_A} \text{vec}\left(\hat{A}_\ell V_{3N_t} H_\ell\right) \end{bmatrix}, \end{aligned} \tag{3.16}$$

where

$$\hat{A}_\ell = \begin{cases} \left(1 + \tau\sqrt{\frac{1+\gamma}{\beta}}\right) M + \tau A_1, & \ell = 1, \\ \tau A_\ell, & \ell = 2, \ldots, n_A, \end{cases}$$

and $V_{3i} \in \mathbb{R}^{N_h \times N_\xi}$ is the *matricized* form of the i-th block of $\mathbf{v_3}$, with $i = 1, \ldots, N_t$.

The practical implementation of the preconditioner for the time-dependent system leverages the inherent structure of the *all-at-once* formulation. As defined in (3.15), the core operator of the Schur complement preconditioner, $\tilde{\mathcal{Z}}$, is block-diagonal with respect to the time steps. This structure extends to the entire KKT system, which then makes the preconditioning easily parallelizable. Specifically, an application of the preconditioner $\mathcal{P}_{\text{hGSoc-PINT}}$ entails an application of the steady-state optimal control

preconditioner $\mathcal{P}_{\text{hGSoc}}$ from Algorithm 2–3 to all time steps simultaneously, and so it represents *parallel-in-time* (PINT) approach. It is summarized as Algorithm 4.

Algorithm 4 Parallel-in-time hGSoc preconditioner (hGSoc-PINT)

The preconditioner $\mathcal{P}_{\text{hGSoc-PINT}} : \bar{\mathbf{R}} \longmapsto \bar{\mathbf{V}}$ is defined as:

1: **for** $k = 1, \ldots, N_t$ **do**
2: Extract $(\bar{R}_k^Y, \bar{R}_k^U, \bar{R}_k^\Lambda)$. (the subvector k of $\bar{\mathbf{R}}$)
3: Calculate $(\bar{V}_k^Y, \bar{V}_k^U, \bar{V}_k^\Lambda) = \mathcal{P}_{hGSoc}(\bar{R}_k^Y, \bar{R}_k^U, \bar{R}_k^\Lambda)$ (apply Algorithm 2–3)
4: **end for**
5: Concatenate $\{(\bar{V}_k^Y, \bar{V}_k^U, \bar{V}_k^\Lambda)\}_{k=1}^{N_t}$ into $\bar{\mathbf{V}}$.

4. Spectral analysis (time-dependent case). Since the steady-state optimal control problem can be viewed as a special case of the time-dependent formulation (with $N_t = 1$), we focus on analyzing the time-dependent setting; the steady-state results then follow as a direct consequence. The all-at-once discretization, presented in Section 3.1, couples all time steps simultaneously, yielding a significantly larger KKT system than its steady-state counterpart. Our goal is to prove that the proposed parallel-in-time preconditioner, based on the hGSoc-PINT in Algorithm 4, is spectrally equivalent to the ideal (but computationally inexpensive) preconditioner.

We begin by recalling the notion of spectral equivalence, which serves as the foundation for our analysis.

DEFINITION 4.1 (Spectral Equivalence). *Two matrices A and B are said to be spectrally equivalent, denoted $A \sim B$, if there exist positive constants $a \leq b$, such that*

$$a\mathbf{v}^T B\mathbf{v} \leq \mathbf{v}^T A\mathbf{v} \leq b\mathbf{v}^T B\mathbf{v}$$

holds for all non-zero vectors $\mathbf{v}$. Equivalently, all eigenvalues of the preconditioned matrix $B^{-1}A$ are contained within the fixed interval, which means $\lambda(B^{-1}A) \subset [a, b]$.

Our proof proceeds by establishing a chain of spectral equivalences among the following operators:

$$\bar{\mathcal{S}}_{\text{exact}} \sim \bar{\mathcal{S}} \sim \tilde{\mathcal{S}} \sim \tilde{\mathcal{S}}_r \sim \tilde{\mathcal{S}}_{\text{hGS-PINT}}.$$

Here, $\bar{\mathcal{S}}_{\text{exact}}$ denotes the exact Schur complement (3.10), $\bar{\mathcal{S}}$ is an approximation of $\bar{\mathcal{S}}_{\text{exact}}$ defined in (3.13), and $\tilde{\mathcal{S}}$ is a block-diagonal approximation obtained by replacing $\bar{\mathcal{Z}}$ in (3.13) with $\tilde{\mathcal{Z}}$ from (3.15), which eliminates time-coupling terms and enables parallel-in-time computation. The operator $\tilde{\mathcal{S}}_r$ represents the truncated hierarchical preconditioner

$$\tilde{\mathcal{S}}_r \;=\; \tilde{\mathcal{Z}}_r \,(D \otimes \mathcal{M}_\gamma)^{-1}\, \tilde{\mathcal{Z}}_r^\top,\; \tilde{\mathcal{Z}}_r \;=\; I_{N_t} \otimes \left\{ H_1 \otimes \left[(1 + \tau\sqrt{\tfrac{1+\gamma}{\beta}})M + \tau A_1 \right] + \tau \sum_{\ell=2}^{r} H_\ell \otimes A_\ell \right\},$$

with $r = 1, \ldots, n_A$. When $r = 1$, $\tilde{\mathcal{S}}_r$ reduces to the mean-based preconditioner employed in [5], and when $r = n_A$, it recovers the full operator $\tilde{\mathcal{S}}$. Finally, $\tilde{\mathcal{S}}_{\text{hGS-PINT}} = \tilde{\mathcal{Z}}_{\text{hGS-PINT}}(D \otimes \mathcal{M}_\gamma)^{-1}\tilde{\mathcal{Z}}_{\text{hGS-PINT}}^T$ represents the computationally feasible approximation of $\tilde{\mathcal{S}}_r$ in which the linear systems $\tilde{\mathcal{Z}}_r\mathbf{x} = \mathbf{b}$ are solved approximately via the hierarchical Gauss-Seidel method (Algorithm 1), as implemented in the parallel-in-time framework of Algorithm 4.

To establish the spectral equivalences in this chain, we require several technical results. We begin with two auxiliary lemmas concerning matrix perturbations and congruence transformations, which will serve as building blocks for the main theorems.

LEMMA 4.2. *Let A and A_r be symmetric positive definite matrices satisfying $(1-\varepsilon_1)A \preceq A_r \preceq (1+\varepsilon_2)A$ in the Loewner order for some $0 < \varepsilon_1, \varepsilon_2 < 1$. Let B be a symmetric positive semidefinite matrix. Then, the eigenvalues of the preconditioned matrix $(A+B)^{-1}(A_r+B)$ are bounded as follows:*

$$1 - \varepsilon_1 \leq \lambda\big((A+B)^{-1}(A_r+B)\big) \leq 1 + \varepsilon_2.$$

Proof. The eigenvalues λ are given by the generalized eigenvalue problem $(A_r + B)\mathbf{v} = \lambda(A+B)\mathbf{v}$. By the Courant-Fischer theorem, these eigenvalues are bounded by the range of the corresponding Rayleigh quotient. For any non-zero vector $\mathbf{v}$, let $a \coloneqq \mathbf{v}^T A\mathbf{v} > 0$ and $b \coloneqq \mathbf{v}^T B\mathbf{v} \geq 0$. The assumption $(1-\varepsilon_1)A \preceq A_r \preceq (1+\varepsilon_2)A$ implies that

$$(1-\varepsilon_1)a \leq \mathbf{v}^T A_r \mathbf{v} \leq (1+\varepsilon_2)a.$$

We can now bound the Rayleigh quotient:

$$\lambda\left((A+B)^{-1}(A_r+B)\right) = \frac{\mathbf{v}^T(A_r+B)\mathbf{v}}{\mathbf{v}^T(A+B)\mathbf{v}} = \frac{\mathbf{v}^T A_r\mathbf{v} + b}{a+b}.$$

Applying the bounds for $\mathbf{v}^T A_r \mathbf{v}$, we get

$$\frac{(1-\varepsilon_1)a + b}{a+b} \leq \lambda \leq \frac{(1+\varepsilon_2)a + b}{a+b}.$$

Simplifying the lower and upper bounds yields

$$1 - \varepsilon_1 \frac{a}{a+b} \leq \lambda \leq 1 + \varepsilon_2 \frac{a}{a+b}.$$

Let $c \coloneqq \frac{a}{a+b}$. Since $a > 0$, $b \geq 0$, it follows that $0 < c \leq 1$. Thus, for any vector $\mathbf{v} \neq 0$, the corresponding value of the Rayleigh quotient lies in the interval $[1-\varepsilon_1 c, 1+\varepsilon_2 c]$. Because $0 < c \leq 1$, this interval is always contained within the larger, fixed interval $[1-\varepsilon_1, 1+\varepsilon_2]$. Therefore, all eigenvalues are bounded by $1-\varepsilon_1$ and $1+\varepsilon_2$. ∎

LEMMA 4.3 (Eigenvalues under Congruence Transformation). *Let C and D be symmetric positive definite matrices, and let Q be a nonsingular matrix. The eigenvalues of the pair (C, D) are identical to those of the transformed pair $(Q^T CQ, Q^T DQ)$.*

Proof. Let $(\lambda, \mathbf{x})$ be an eigenpair satisfying the generalized eigenvalue problem $C\mathbf{x} = \lambda D\mathbf{x}$, with eigenvector $\mathbf{x} \neq \mathbf{0}$. We perform a change of variables by setting $\mathbf{x} = Q\mathbf{y}$. Since Q is nonsingular, $\mathbf{x} \neq \mathbf{0}$ implies that the transformed vector $\mathbf{y} \neq \mathbf{0}$. Substituting $\mathbf{x} = Q\mathbf{y}$ into the original problem gives:

$$C(Q\mathbf{y}) = \lambda D(Q\mathbf{y}).$$

Multiplying from the left by Q^T, we obtain:

$$(Q^T CQ)\mathbf{y} = \lambda(Q^T DQ)\mathbf{y}.$$

This final expression is the generalized eigenvalue problem for the pair $(Q^T CQ, Q^T DQ)$, which is satisfied by the same eigenvalue λ with the transformed eigenvector $\mathbf{y}$. Therefore, the sets of eigenvalues for both pairs are identical. ∎

With these auxiliary results in place, we now proceed to establish the spectral equivalences in the chain $\bar{\mathcal{S}}_{\text{exact}} \sim \bar{\mathcal{S}} \sim \tilde{\mathcal{S}} \sim \tilde{\mathcal{S}}_r \sim \tilde{\mathcal{S}}_{\text{hGS-PINT}}$. We begin by proving the

first equivalence, $\bar{\mathcal{S}} \sim \bar{\mathcal{S}}_{\text{exact}}$, which shows that our factorized approximation (3.13) is spectrally close to the exact Schur complement.

THEOREM 4.4 (Theorems 4, 6 in [5]). *Let $\bar{\mathcal{S}}_{exact}$ be the exact Schur complement and $\bar{\mathcal{S}}$ be its approximation as defined in (3.13) for the steady-state case or (3.9) for the time-dependent case. For any $\alpha \geq 0$ satisfying the condition below, the eigenvalues of $\bar{\mathcal{S}}^{-1}\bar{\mathcal{S}}_{exact}$ are bounded as*

$$\lambda(\bar{\mathcal{S}}^{-1}\bar{\mathcal{S}}_{exact}) \subseteq \left[\frac{1}{2(1+\alpha)}, 1\right). \tag{4.1}$$

The condition on α satisfies

$$\alpha < \left(\frac{\sqrt{\kappa(\mathcal{A}_t)}+1}{\sqrt{\kappa(\mathcal{A}_t)}-1}\right)^2 - 1. \tag{4.2}$$

Next, we show the relationship between $\bar{\mathcal{S}}$ and the matrix $\tilde{\mathcal{S}} = \tilde{\mathcal{Z}}(D \otimes \mathcal{M}_\gamma)^{-1}\tilde{\mathcal{Z}}^T$ defined in (3.15). To do this, we first introduce a standard lemma concerning singular value perturbation.

LEMMA 4.5. *For any matrices A and B of the same dimensions,*

$$\sigma_{\min}(A+B) \geq \sigma_{\min}(B) - \|A\|_2.$$

Here, $\sigma_{\min}(\cdot)$ denotes the smallest singular value of its matrix argument.

Proof. From the triangle inequality, for any vector x, with $\|x\|_2 = 1$, we have

$$\|(A+B)x\|_2 = \|Bx - (-A)x\|_2 \geq \|Bx\|_2 - \|Ax\|_2.$$

Taking the minimum over all unit vectors x on both sides of the inequality, we get

$$\min_{\|x\|_2=1} \|(A+B)x\|_2 \geq \min_{\|x\|_2=1} \left(\|Bx\|_2 - \|Ax\|_2\right).$$

Using the property that $\min(f-g) \geq \min(f) - \max(g)$, we obtain

$$\min_{\|x\|_2=1} \left(\|Bx\|_2 - \|Ax\|_2\right) \geq \min_{\|x\|_2=1} \|Bx\|_2 - \max_{\|x\|_2=1} \|Ax\|_2.$$

By the definitions of the minimum singular value and the operator norm, the above expression is equivalent to $\sigma_{\min}(A+B) \geq \sigma_{\min}(B) - \|A\|_2$. ∎

LEMMA 4.6. *Let $W = (D \otimes \mathcal{M}_\gamma)^{-1}$, and assume that there exists a constant $\mu > 1$ such that*

$$\tau\sqrt{\frac{1+\gamma}{\beta}}\sigma_{\min}(\mathcal{N}W^{\frac{1}{2}}) \geq \mu\|(\mathcal{A}_t + C \otimes \mathcal{M})W^{\frac{1}{2}}\|_2. \tag{4.3}$$

Then the minimum eigenvalue of $\tilde{\mathcal{S}}$ has the following lower bound

$$\lambda_{\min}(\tilde{\mathcal{S}}) \geq \frac{\tau}{\beta}\left(1 - \frac{1}{\mu}\right)^2 \sigma^2_{\min}(M^{\frac{1}{2}}).$$

Proof. From the definition of $\tilde{\mathcal{S}}$ and the properties of the minimum eigenvalue,

$$\lambda_{\min}(\tilde{\mathcal{S}}) = \lambda_{\min}(\tilde{\mathcal{Z}}W\tilde{\mathcal{Z}}^T) = \sigma^2_{\min}(\tilde{\mathcal{Z}}W^{\frac{1}{2}}).$$

Since $\tilde{\mathcal{Z}}W^{\frac{1}{2}} = \tau\sqrt{\frac{1+\gamma}{\beta}}\mathcal{N}W^{\frac{1}{2}} + (\mathcal{A}_t + C \otimes \mathcal{M})W^{\frac{1}{2}}$, applying Lemma 4.5 yields

$$\sigma_{\min}(\tilde{\mathcal{Z}}W^{\frac{1}{2}}) \geq \tau\sqrt{\frac{1+\gamma}{\beta}}\sigma_{\min}(\mathcal{N}W^{\frac{1}{2}}) - \|(\mathcal{A}_t + C \otimes \mathcal{M})W^{\frac{1}{2}}\|_2.$$

We assume there exists a constant $\mu > 1$ such that

$$\tau\sqrt{\frac{1+\gamma}{\beta}}\sigma_{\min}(\mathcal{N}W^{\frac{1}{2}}) \geq \mu\|(\mathcal{A}_t + C \otimes \mathcal{M})W^{\frac{1}{2}}\|_2.$$

Using the assumption (4.3), we obtain[1]

$$\sigma_{\min}(\tilde{\mathcal{Z}}W^{\frac{1}{2}}) \geq \left(1 - \frac{1}{\mu}\right)\tau\sqrt{\frac{1+\gamma}{\beta}}\sigma_{\min}(\mathcal{N}W^{\frac{1}{2}}).$$

Consequently, the lower bound for $\lambda_{\min}(\tilde{\mathcal{S}})$ is:

$$\lambda_{\min}(\tilde{\mathcal{S}}) \geq \left(1 - \frac{1}{\mu}\right)^2 \left(\tau\sqrt{\frac{1+\gamma}{\beta}}\right)^2 \sigma^2_{\min}(\mathcal{N}W^{\frac{1}{2}}).$$

Next, observe that

$$\begin{aligned}
\sigma^2_{\min}(\mathcal{N}W^{\frac{1}{2}}) &= \sigma^2_{\min}\left((I_{N_t} \otimes \mathcal{M})(D^{-\frac{1}{2}} \otimes \mathcal{M}_\gamma^{-\frac{1}{2}})\right) \\
&= \sigma^2_{\min}\left((I_{N_t}D^{-\frac{1}{2}}) \otimes (\mathcal{M}\mathcal{M}_\gamma^{-\frac{1}{2}})\right) \\
&= \sigma^2_{\min}(D^{-\frac{1}{2}}) \cdot \sigma^2_{\min}(\mathcal{M}\mathcal{M}_\gamma^{-\frac{1}{2}}).
\end{aligned}$$

It is easy to verify that $\sigma_{\min}(D^{-\frac{1}{2}}) = 1$ and

$$\begin{aligned}
\sigma^2_{\min}(\mathcal{M}\mathcal{M}_\gamma^{-\frac{1}{2}}) &= \sigma^2_{\min}((H_1 \otimes M)(H_\gamma^{-\frac{1}{2}} \otimes M^{-\frac{1}{2}})) \\
&= \sigma^2_{\min}((H_1H_\gamma^{-\frac{1}{2}}) \otimes (MM^{-\frac{1}{2}})) \\
&= \frac{1}{1+\gamma}\sigma^2_{\min}(M^{\frac{1}{2}}).
\end{aligned}$$

Thus, we have

$$\sigma^2_{\min}(\mathcal{N}W^{\frac{1}{2}}) = \frac{1}{1+\gamma}\sigma^2_{\min}(M^{\frac{1}{2}}).$$

Substituting this into the expression for the lower bound of $\lambda_{\min}(\tilde{\mathcal{S}})$, we finally get

$$\lambda_{\min}(\tilde{\mathcal{S}}) \geq \left(1 - \frac{1}{\mu}\right)^2 \frac{\tau^2(1+\gamma)}{\beta}\left(\frac{1}{1+\gamma}\sigma^2_{\min}(M^{\frac{1}{2}})\right) = \frac{\tau}{\beta}\left(1 - \frac{1}{\mu}\right)^2 \sigma^2_{\min}(M^{\frac{1}{2}}). \tag{4.4}$$

□

We can now state the following result.

THEOREM 4.7. *Assume that the conditions of Lemma 4.6 hold. Then, the the eigenvalues $\lambda(\bar{\mathcal{S}}^{-1}\tilde{\mathcal{S}})$ satisfy*

$$(1-\theta)^2 \leq \lambda(\bar{\mathcal{S}}^{-1}\tilde{\mathcal{S}}) \leq (1+\theta)^2,$$

[1] In our experience, this condition is often satisfied numerically when $\tau \gg \sqrt{\beta}$.

where θ is the perturbation parameter, with

$$\theta \coloneqq \sup_{\mathbf{v}\neq 0} \frac{\|W^{\frac{1}{2}}(C\otimes\mathcal{M})^T\mathbf{v}\|_2}{\|W^{\frac{1}{2}}\tilde{\mathcal{Z}}^T\mathbf{v}\|_2} \leq \frac{\sqrt{2\beta}}{\sqrt{1+\gamma}\left(1-\frac{1}{\mu}\right)\tau}\kappa(M^{\frac{1}{2}}).$$

Proof. By the definition of θ and the triangle inequality, for any $\mathbf{v}\neq 0$, we have

$$\left|\|W^{\frac{1}{2}}\tilde{\mathcal{Z}}^T\mathbf{v}\|_2 - \|W^{\frac{1}{2}}(C\otimes\mathcal{M})^T\mathbf{v}\|_2\right| \leq \|W^{\frac{1}{2}}\bar{\mathcal{Z}}^T\mathbf{v}\|_2 \leq \|W^{\frac{1}{2}}\tilde{\mathcal{Z}}^T\mathbf{v}\|_2+\|W^{\frac{1}{2}}(C\otimes\mathcal{M})^T\mathbf{v}\|_2.$$

Using the bound $\|W^{\frac{1}{2}}(C\otimes\mathcal{M})^T\mathbf{v}\|_2 \leq \theta\|W^{\frac{1}{2}}\tilde{\mathcal{Z}}^T\mathbf{v}\|_2$, we obtain

$$(1-\theta)\|W^{\frac{1}{2}}\tilde{\mathcal{Z}}^T\mathbf{v}\|_2 \leq \|W^{\frac{1}{2}}\bar{\mathcal{Z}}^T\mathbf{v}\|_2 \leq (1+\theta)\|W^{\frac{1}{2}}\tilde{\mathcal{Z}}^T\mathbf{v}\|_2.$$

Squaring these inequalities leads to the spectral bounds

$$(1-\theta)^2 \leq \lambda(\bar{\mathcal{S}}^{-1}\tilde{\mathcal{S}}) \leq (1+\theta)^2.$$

Next, it remains to analyze the upper bound of θ. To this end, observe that

$$\theta = \sup_{\mathbf{v}\neq 0} \frac{\|W^{\frac{1}{2}}(C\otimes\mathcal{M})^T\mathbf{v}\|_2}{\|W^{\frac{1}{2}}\tilde{\mathcal{Z}}^T\mathbf{v}\|_2} \leq \frac{\|W^{\frac{1}{2}}(C\otimes\mathcal{M})^T\|_2}{\sigma_{\min}(W^{\frac{1}{2}}\tilde{\mathcal{Z}}^T)},$$

from which we bound the numerator and the denominator separately. First, for the numerator

$$\begin{aligned}\|W^{\frac{1}{2}}(C\otimes\mathcal{M})^T\|_2 &= \|(D^{-\frac{1}{2}}\otimes\mathcal{M}_\gamma^{-\frac{1}{2}})(C^T\otimes\mathcal{M}^T)\|_2\\ &= \|(D^{-\frac{1}{2}}C^T)\otimes(\mathcal{M}_\gamma^{-\frac{1}{2}}\mathcal{M}^T)\|_2\\ &= \|D^{-\frac{1}{2}}C^T\|_2\cdot\|\mathcal{M}_\gamma^{-\frac{1}{2}}\mathcal{M}^T\|_2.\end{aligned}$$

Also, note that simple calculation yields $\|D^{-\frac{1}{2}}C^T\|_2\cdot\|\mathcal{M}_\gamma^{-\frac{1}{2}}\mathcal{M}^T\|_2 \leq \frac{\sqrt{2}}{\sqrt{1+\gamma}}\|M^{\frac{1}{2}}\|_2$, so that the numerator satisfies

$$\|W^{\frac{1}{2}}(C\otimes\mathcal{M})^T\|_2 \leq \frac{\sqrt{2}}{\sqrt{1+\gamma}}\|M^{\frac{1}{2}}\|_2.$$

For the denominator, using the result from the previous lemma, we have

$$\sigma_{\min}(W^{\frac{1}{2}}\tilde{\mathcal{Z}}^T) = \sigma_{\min}(\tilde{\mathcal{Z}}W^{\frac{1}{2}}) \geq \left(1-\frac{1}{\mu}\right)\frac{\tau}{\sqrt{\beta}}\sigma_{\min}(M^{\frac{1}{2}}).$$

Combining the bounds for the numerator and denominator,

$$\begin{aligned}\theta &\leq \frac{\frac{\sqrt{2}}{\sqrt{1+\gamma}}\|M^{\frac{1}{2}}\|_2}{\left(1-\frac{1}{\mu}\right)\frac{\tau}{\sqrt{\beta}}\sigma_{\min}(M^{\frac{1}{2}})}\\ &= \frac{\sqrt{2\beta}}{\sqrt{1+\gamma}\left(1-\frac{1}{\mu}\right)\tau}\cdot\frac{\sigma_{\max}(M^{\frac{1}{2}})}{\sigma_{\min}(M^{\frac{1}{2}})}\\ &= \frac{\sqrt{2\beta}}{\sqrt{1+\gamma}\left(1-\frac{1}{\mu}\right)\tau}\kappa(M^{\frac{1}{2}}),\end{aligned}$$

thereby completing the proof of the theorem. □

Note that the bound on θ implies that, for sufficiently large τ (e.g., $\tau \gtrsim \sqrt{\beta}\,\kappa(M)^{1/2}$), we have $\theta < 1$; hence $\bar{\mathcal{S}}$ and $\tilde{\mathcal{S}}$ are spectrally close. In practice, $\kappa(M)$ can be kept $\mathcal{O}(1)$ via appropriate basis choices and mass lumping, so it suffices to require $\tau \gg \sqrt{\beta}$.

Next, we prove the spectral equivalence between $\tilde{\mathcal{S}}$ and its truncated form $\tilde{\mathcal{S}}_r$. First, note from [6, Theorem 3.8], that

$$(1-\varepsilon_1)\left(\sum_{\ell=1}^{n_A} H_\ell \otimes A_\ell\right) \preceq \sum_{\ell=1}^{r} H_\ell \otimes A_\ell \preceq (1+\varepsilon_2)\left(\sum_{\ell=1}^{n_A} H_\ell \otimes A_\ell\right),$$

where $\varepsilon_1, \varepsilon_2$ indicate the importance of the truncated residual; the more terms to be calculated, the smaller $\varepsilon_1, \varepsilon_2$ we have, the tighter the preconditioned spectrum is clustered around unity, and the more accurate and the better preconditioner we have. Applying Lemma 4.2, and $H_1 \otimes M \succeq 0$, we know that

$$(1-\varepsilon_1)\tilde{\mathcal{Z}} \preceq \tilde{\mathcal{Z}}_r \preceq (1+\varepsilon_2)\tilde{\mathcal{Z}}.$$

From Lemma 4.3, we know that the generalized spectra of $(\tilde{\mathcal{S}}_r, \tilde{\mathcal{S}})$ and $\left((W^{\frac{1}{2}}\tilde{\mathcal{Z}}_r W^{\frac{1}{2}})^2, (W^{\frac{1}{2}}\tilde{\mathcal{Z}} W^{\frac{1}{2}})^2\right)$ coincide. Therefore we get

$$(1-\varepsilon_1)^2 \le \lambda(\tilde{\mathcal{S}}^{-1}\tilde{\mathcal{S}}_r) \le (1+\varepsilon_2)^2.$$

This establishes the spectral equivalence of $\tilde{\mathcal{S}}_r$ and $\tilde{\mathcal{S}}$.

Finally, we proceed to establish the spectral equivalence between the truncated preconditioner $\tilde{\mathcal{S}}_r$ and the hierarchical symmetric block Gauss-Seidel approximation $\tilde{\mathcal{S}}_{\text{hGS-PINT}}$. To this end, following the idea from [6], we can rewrite it as $H_\ell = L_\ell + L_\ell^T$, $\ell = 2, 3, \ldots, n_A$, and matrices L_ℓ have at most one nonzero entry per row and per column.

Now, define

$$X_1 = I_{N_t} \otimes \left(H_1 \otimes (1+\tau\sqrt{\frac{1+\gamma}{\beta}})M + \tau A_1\right)$$
$$X_r = I_{N_t} \otimes \left(\sum_{\ell=2}^{r} L_\ell \otimes A_\ell\right),\ r = 2, 3, \ldots n_A,$$

so we know $\tilde{\mathcal{Z}}_{\text{hGS-PINT}} = (X_1 + X_r)X_1^{-1}(X_1 + X_r^T) = \tilde{\mathcal{Z}}_r + X_r X_1^{-1} X_r^T$, the Rayleigh quotient

$$\frac{\mathbf{v}^T \tilde{\mathcal{Z}}_{\text{hGS-PINT}}\mathbf{v}}{\mathbf{v}^T \tilde{\mathcal{Z}}_r \mathbf{v}} = 1 + \underbrace{\frac{\mathbf{v}^T X_r X_1^{-1} X_r^T \mathbf{v}}{\mathbf{v}^T(X_1 + X_r + X_r^T)\mathbf{v}}}_{\zeta}.$$

Let $\mathbf{u} = X_1^{\frac{1}{2}}\mathbf{v}$; then

$$\zeta(\mathbf{u}) = \frac{\mathbf{u}^T Y Y^T \mathbf{u}}{\mathbf{u}^T(I + Y + Y^T)\mathbf{u}},$$

where $Y = X_1^{-\frac{1}{2}} X_r X_1^{-\frac{1}{2}} = I_{N_t} \otimes (\sum_{\ell=2}^{r} L_\ell \otimes \mathfrak{A}_\ell)$, and

$$\mathfrak{A}_\ell = \left((I + \tau\sqrt{\frac{1+\gamma}{\beta}})M + \tau A_1\right)^{-\frac{1}{2}} A_\ell \left((I + \tau\sqrt{\frac{1+\gamma}{\beta}})M + \tau A_1\right)^{-\frac{1}{2}}.$$

Hence,

$$\zeta(\mathbf{u}) \le \max_{\mathbf{u}\neq 0} \frac{\mathbf{u}^T Y Y^T \mathbf{u}}{u^T u} \cdot \max_{\mathbf{u}\neq 0} \frac{\mathbf{u}^T \mathbf{u}}{\mathbf{u}^T (I + Y + Y^T)\mathbf{u}} = \frac{\sigma^2_{\max}(Y)}{\lambda_{\min}(I + Y + Y^T)}. \tag{4.5}$$

The following result holds.

LEMMA 4.8. *Define*

$$\Delta_r \;:=\; \sum_{\ell=2}^{r} \|H_\ell\|_2 \, \rho_\ell, \qquad \rho_\ell \;:=\; \|\mathfrak{A}_\ell\|_2,$$

where

$$\mathfrak{A}_\ell := \Big((I + \tau\sqrt{\tfrac{1+\gamma}{\beta}})\, M + \tau A_1\Big)^{-\frac{1}{2}} \, A_\ell \, \Big((I + \tau\sqrt{\tfrac{1+\gamma}{\beta}})\, M + \tau A_1\Big)^{-\frac{1}{2}}.$$

Let

$$Y := X_1^{-1/2} X_r X_1^{-1/2} = I_{N_t} \otimes \Big(\sum_{\ell=2}^{r} L_\ell \otimes \mathfrak{A}_\ell\Big), \qquad E := Y + Y^\top = I_{N_t} \otimes \Big(\sum_{\ell=2}^{r} H_\ell \otimes \mathfrak{A}_\ell\Big),$$

with $H_\ell = L_\ell + L_\ell^\top$. *Then*

$$\lambda_{\min}\big(I + Y + Y^\top\big) \;\ge\; 1 - \Delta_r, \quad \sigma_{\max}(Y) \;\le\; \sum_{\ell=2}^{r} \|L_\ell\|_2 \, \rho_\ell \;\le\; \Delta_r, \quad \textit{and} \quad \rho_\ell \;\le\; \frac{\|\sqrt{\theta_\ell}\, \kappa_\ell\|_{L^\infty(\Omega)}}{\mathbb{k}_{\min}},$$

where $\sigma_{\max}(\cdot)$ *denotes the largest singular values.*

Proof. Using

$$I + Y + Y^\top = X_1^{-1/2}\Big(X_1 + I_{N_t} \otimes \sum_{\ell=2}^{r} H_\ell \otimes A_\ell\Big) X_1^{-1/2} = X_1^{-1/2}\, \bar{\mathcal{Z}}_r \, X_1^{-1/2},$$

the spectrum of $I + Y + Y^\top$ coincides with the generalized spectrum of the pair $(\bar{\mathcal{Z}}_r, X_1)$. Since E is symmetric, $\lambda_{\min}(I + E) \ge 1 - \|E\|_2$. By the Kronecker product norm rule and the triangle inequality,

$$\|E\|_2 = \Big\| I_{N_t} \otimes (\sum_{\ell=2}^{r} H_\ell \otimes \mathfrak{A}_\ell) \Big\|_2 \le \sum_{\ell=2}^{r} \|H_\ell\|_2 \, \|\mathfrak{A}_\ell\|_2 = \sum_{\ell=2}^{r} \|H_\ell\|_2 \, \rho_\ell = \Delta_r,$$

which yields $\lambda_{\min}(I + Y + Y^\top) \ge 1 - \Delta_r$.

Similarly,

$$\|Y\|_2 = \Big\| I_{N_t} \otimes \sum_{\ell=2}^{r} L_\ell \otimes \mathfrak{A}_\ell \Big\|_2 \le \sum_{\ell=2}^{r} \|L_\ell\|_2 \, \rho_\ell.$$

Since $H_\ell = L_\ell + L_\ell^\top$ and each L_ℓ has at most one nonzero per row and per column, we have $\|L_\ell\|_2 \le \|H_\ell\|_2$. Hence $\sigma_{\max}(Y) = \|Y\|_2 \le \Delta_r$.

For the explicit bound on ρ_ℓ, recall that $\rho_\ell = \|\mathfrak{A}_\ell\|_2$ is the maximum eigenvalue of the generalized eigenvalue problem $A_\ell \mathbf{v} = \lambda K \mathbf{v}$, where $K = (I + \tau\sqrt{\frac{1+\gamma}{\beta}})\, M + \tau A_1$.

In terms of the associated finite element function v_h, the Rayleigh quotient is given by

$$\frac{\mathbf{v}^\top A_\ell \mathbf{v}}{\mathbf{v}^\top K \mathbf{v}} = \frac{\displaystyle\int_\Omega a_\ell \, |\nabla v_h|^2 \, dx}{\displaystyle\int_\Omega \left((1 + \tau\sqrt{\tfrac{1+\gamma}{\beta}})|v_h|^2 + \tau a_0 \, |\nabla v_h|^2 \right) dx}.$$

Since the mass term is non-negative, we can bound this ratio by neglecting the L^2-term in the denominator:

$$\frac{\mathbf{v}^\top A_\ell \mathbf{v}}{\mathbf{v}^\top K \mathbf{v}} \;\le\; \frac{\displaystyle\int_\Omega a_\ell \, |\nabla v_h|^2 \, dx}{\tau \displaystyle\int_\Omega a_0 \, |\nabla v_h|^2 \, dx} \;\le\; \frac{1}{\tau} \frac{\|\sqrt{\theta_\ell}\, \kappa_\ell\|_{L^\infty(\Omega)}}{\Bbbk_{\min}},$$

and taking the supremum gives the stated bound on ρ_ℓ; substituting it into $\|Y\|_2$ yields the last inequality. □

Observe from above that, with $\Delta_r < 1$,

$$\zeta(\mathbf{v}) \;=\; \frac{\mathbf{v}^\top Y Y^\top \mathbf{v}}{\mathbf{v}^\top (I + Y + Y^\top)\mathbf{v}} \;\le\; \frac{\sigma_{\max}^2(Y)}{\lambda_{\min}(I + Y + Y^\top)} \;\le\; \frac{\Delta_r^2}{1 - \Delta_r}.$$

Consequently,

$$1 \;\le\; \frac{\mathbf{v}^\top \mathcal{Z}_{\text{hGS}-\text{PINT}} \mathbf{v}}{\mathbf{v}^\top \bar{\mathcal{Z}}_r \mathbf{v}} \;\le\; 1 + \frac{\Delta_r^2}{1 - \Delta_r}.$$

From Lemma 4.3, we know that the generalized spectra of $(\tilde{\mathcal{S}}_{\text{hGS-PINT}}, \tilde{\mathcal{S}}_r)$ and $\left((W^{\frac{1}{2}} \mathcal{Z}_{\text{hGS-PINT}} W^{\frac{1}{2}})^2, (W^{\frac{1}{2}} \tilde{\mathcal{Z}}_r W^{\frac{1}{2}})^2 \right)$ coincide. Therefore we get

$$1 \le \lambda(\tilde{\mathcal{S}}_r^{-1} \tilde{\mathcal{S}}_{\text{hGS-PINT}}) \le (1 + \frac{\Delta_r^2}{1 - \Delta_r})^2.$$

COROLLARY 4.9 (Steady-State Case). *Since the steady-state optimal control problem corresponds to the special case $N_t = 1$ of the time-dependent formulation, all preceding results apply directly with the simplified notation. The spectral equivalence chain for the steady-state Schur complement preconditioner is*

$$\mathcal{S}_{exact} \sim \mathcal{S} \sim \mathcal{S}_r \sim \mathcal{S}_{hGS},$$

where $\mathcal{S}_{exact}$ is defined in (2.17), $\mathcal{S}$ in (2.18), and

$$\mathcal{S}_r = \mathcal{Z}_r \mathcal{M}_\gamma^{-1} \mathcal{Z}_r^\top, \qquad \mathcal{Z}_r = \sum_{\ell=1}^{r} H_\ell \otimes \tilde{A}_\ell,$$

with $\tilde{A}_1 = A_1 + \sqrt{\frac{1+\gamma}{\beta}} M$ and $\tilde{A}_\ell = A_\ell$ for $\ell = 2, \ldots, r$. When $r = 1$, $\mathcal{S}_r$ reduces to the mean-based preconditioner, and when $r = n_A$, it recovers the full operator $\mathcal{S}$. Finally, $\mathcal{S}_{hGS}$ represents the computationally feasible approximation of $\mathcal{S}_r$ in which the linear systems $\mathcal{Z}_r \mathbf{x} = \mathbf{b}$ are solved approximately via the hierarchical Gauss-Seidel method (Algorithm 1), as implemented in Algorithm 2–3.

5. Numerical experiments. This section validates the theoretical findings of Sections 2–3 through comprehensive numerical experiments. We pursue two primary objectives: (i) verifying the mesh-independence and spectral bounds established in the preceding sections, and (ii) demonstrating the computational efficiency of the proposed hierarchical Gauss-Seidel (hGS) preconditioner across varying truncation strategies. Experiments are presented for both steady-state problems (Section 5.1) and time-dependent problems (Section 5.2). The numerical experiments were performed on a system running AlmaLinux-9 with 40GB RAM, and the proposed algorithms were implemented using MATLAB 23.2.

The random input $\Bbbk$ is characterized by the covariance function (2.2)–(2.3),

$$C_{\Bbbk}(x,y) = \sigma_{\Bbbk}^2 \exp\left(-\frac{|x_1-y_1|}{\ell_1} - \frac{|x_2-y_2|}{\ell_2}\right) \quad \forall (x,y) \in [-1,1]^2,$$

where $\sigma_{\Bbbk}$ is the variance coefficient, determining the randomness of the input. In our simulations, we set the correlation lengths as $\ell_1 = \ell_2 = 1$ and the mean of the data as $\mathbb{E}[\Bbbk] = 1$. For gPC setting in (2.9), we consider the case of the log-normal distribution with Hermite polynomials. This problem has been extensively studied in [19]. Also, we used $\gamma = 1$ in both cases, which means we only consider the case with standard deviation. To discretize the spatial domain, we implemented our code based on IFISS 3.7 [23], using $\mathbf{Q}_1$ approximation. For temporal discretization, we apply the all-at-once technique proposed in [20] and set the terminal time as $T = 1$. In all numerical experiments, the spatial mesh size h and the time step τ are chosen as 2^{-i}, with $i = 4, 5, 6, 7$. We solve the linear systems (2.15) and (3.8) using the preconditioners given by (2.16) and (3.9), respectively, employing the flexible GMRES method (without restarting) [22]. The stopping criterion is defined in terms of the relative residual $\|r_k\|/\|b\|$, with thresholds 10^{-8} for the steady-state experiments and 10^{-6} or 10^{-4} for the time-dependent runs, where r_k denotes the residual at iteration k and b is the right-hand side vector. To assess the effectiveness of the hierarchical preconditioning strategy, we systematically compare three truncation settings for the (3,3)-block preconditioner: $n_\tau = 1$ (mean-based approximation), $n_\tau = m_\xi + 1$ (hGS truncated at the first-order stochastic terms), and $n_\tau = n_A$ (full expansion retaining all c_{ijk} coefficients). Both iteration counts and computational times (in seconds) are reported for each configuration. To verify the efficiency of the hierarchical Gauss-Seidel method, we compare both iteration counts and computational costs under different truncation settings: $n_\tau = 1$, $m_\xi + 1$, and n_ξ.

We consider homogeneous Dirichlet conditions, corresponding to Example 2 in [9, Chapter 5]. This example is defined on a square domain $\Omega_\square$ with a discontinuous target function and inconsistent boundary data

$$y_d = \begin{cases} 1 & \text{in } \Omega_1 := [-1,0]^2, \\ 0 & \text{in } \Omega \backslash \Omega_1. \end{cases} \tag{5.1}$$

We subsequently present numerical experiments for both steady-state and time-dependent problems to illustrate and verify the efficiency of our proposed hGS method.

Here are some details about the implementation for preconditioners (2.16), (3.9). Since time-dependent problems can be seen as a series of steady-state problems, and also because of the diagonal structure of matrix D and matrix I_{N_t}, we can just focus on the steady-state preconditioner.

The practical implementation of the preconditioner $\mathcal{P}$ in (2.16) involves different strategies for its constituent blocks. For the (1,1) and (2,2) blocks, which are

based on Kronecker products involving the mass matrix M, applying their inverses requires solving linear systems with M. These solves are handled efficiently by either a direct Cholesky decomposition or the iterative Chebyshev semi-iteration method (Algorithm 5) [11, 29].

For the more complex (3,3) block, which represents the approximate Schur complement $\mathcal{S}$, we employ an outer iterative scheme. Specifically, we use the Preconditioned Richardson method, outlined in Algorithm 6 [7, Chapter 7], where the core of our proposalâ the hierarchical Gauss-Seidel (hGS) method from Algorithm 1 serves as the preconditioner for each Richardson step.

Algorithm 5 Chebyshev semi-iteration for mass matrix preconditioning [11, 29]

Given mass matrix M, vectors b, $x^{(0)} = 0$, $x^{(-1)} = 0$, and parameter $\omega_0 = 1$.
Set $\lambda_{\min} = 1/4$, and $\lambda_{\max} = 9/4$
Calculate $\gamma = (\lambda_{\min} + \lambda_{\max})/2$ and $\rho = (\lambda_{\max} - \lambda_{\min})/(\lambda_{\max} + \lambda_{\min})$.
Set $D = \gamma \cdot \mathrm{diag}(M)$. (a diagonal matrix)
for $k = 0, 1, \ldots, N-1$ **do**
    $\omega_{k+1} = 1/\left(1 - \frac{\omega_k \rho^2}{4}\right)$
    $r^{(k)} = b - Mx^{(k)}$
    $p^{(k)} = D^{-1} r^{(k)}$
    $x^{(k+1)} = \omega_{k+1} \cdot (p^{(k)} + x^{(k)} - x^{(k-1)}) + x^{(k-1)}$
end for
Return $x = x^{(N)}$.

Algorithm 6 Preconditioned Richardson iteration with hGS preconditioner

Given matrix $\mathcal{Z}$, vector b, and initial guess $x^{(1)}$.
$r^{(1)} = b - \mathcal{Z}x^{(1)}$. (initial residual)
for $k = 1, 2, \ldots, N$ **do**
    Solve $\mathcal{Z}z^{(k)} = r^{(k)}$. (apply Algorithm 1)
    $x^{(k+1)} = x^{(k)} + z^{(k)}$ (update solution)
    $r^{(k+1)} = b - \mathcal{Z}x^{(k+1)}$. (update residual)
end for
Return $x^{(N+1)}$

5.1. Steady-state case. This subsection focuses on the steady-state optimal control problem (2.5). We examine the performance of the proposed preconditioner (2.16) under systematic variations in: (i) the variance coefficient $\sigma_{\mathbb{k}}$ (Tables 5.1–5.3), (ii) the regularization parameter β (Table 5.4), and (iii) the spatial and stochastic discretization levels. For each configuration, we compare three solvers for the (1,1) and (2,2) blocks—Chebyshev semi-iteration with 5 or 10 steps (Algorithm 5) and direct Cholesky factorization—combined with the three truncation strategies for the (3,3)-block described above. For the (3,3)-block, we apply the Richardson iteration (Algorithm 6) with $N = 1$, i.e., one application of the hGS preconditioner per outer GMRES iteration. All tests use a fixed tolerance of 10^{-8}.

Next, by fixing the parameter β, we perform further tests summarized in Tables (5.1-5.3), employing different step settings for the Chebyshev smoother and the Cholesky decomposition for blocks (1,1) and (2,2), as well as various truncation strate-

TABLE 5.1

Simulation results showing the total number of iterations from low-rank preconditioned GMRES and the total CPU times (in seconds) using preconditioner with $\beta = 10^{-4}$, $\sigma = 0.01$, and selected spatial (N_h) and stochastic (N_ξ) degrees of freedom

	# iter(t)			# iter(t)			# iter(t)			# iter(t)		
N_ξ \ N_h	289($h=\frac{1}{2^4}$)			1089($h=\frac{1}{2^5}$)			4225($h=\frac{1}{2^6}$)			16641($h=\frac{1}{2^7}$)		
n_τ	1	m_ξ+1	n_A	1	m_ξ+1	n_A	1	m_ξ+1	n_A	1	m_ξ+1	n_A
$\sigma_a = 0.01$												
Chebyshev-5+hGS-1												
(3,3) 20	32(5.7)	32(4.0)	32(5.0)	36(9.2)	35(10.0)	35(15.6)	36(23.2)	27(26.9)	27(46.5)	35(173.4)	33(144.9)	27(194.8)
(4,4) 70	32(129.1)	32(129.5)	32(182.9)	36(243.4)	35(239.1)	35(409.2)	36(752.3)	34(724.9)	34(1237.8)	35(2585.4)	33(2524.6)	33(5253.9)
(6,3) 84	32(317.5)	32(318.3)	32(403.0)	36(508.9)	35(498.2)	35(601.3)	36(1059.3)	34(864.8)	34(1632.5)	35(3099.0)	33(3798.1)	33(5630.2)
Chebyshev-10+hGS-1												
(3,3) 20	26(2.8)	26(2.6)	26(4.4)	30(6.9)	29(7.1)	29(12.0)	31(20.2)	31(21.1)	31(37.6)	32(145.1)	31(130.3)	31(154.5)
(4,4) 70	26(111.7)	26(112.2)	26(153.7)	30(191.6)	29(189.0)	29(323.8)	32(550.7)	31(383.1)	31(830.2)	33(1745.9)	31(2273.0)	31(3341.3)
(6,3) 84	26(263.9)	26(266.1)	26(329.2)	30(457.7)	29(432.2)	29(504.1)	32(959.1)	31(800.4)	31(1515.3)	33(3015.1)	31(3643.9)	31(5318.7)
Cholesky+hGS-1												
(3,3) 20	25(3.5)	25(4.0)	25(4.3)	29(7.4)	27(6.7)	27(10.9)	29(19.1)	29(26.5)	29(44.5)	31(143.6)	29(162.7)	29(182.9)
(4,4) 70	25(104.2)	25(107.6)	25(144.3)	29(177.3)	27(168.9)	27(283.9)	29(419.0)	29(480.8)	29(937.1)	31(2390.7)	29(2046.0)	29(3299.9)
(6,3) 84	25(245.4)	25(247.4)	25(316.8)	29(429.9)	27(393.6)	27(675.5)	29(831.9)	29(1008.3)	29(1756.8)	31(3563.3)	29(3393.6)	29(6436.3)

TABLE 5.2

Simulation results showing the total number of iterations from low-rank preconditioned GMRES and the total CPU times (in seconds) using preconditioner with $\beta = 10^{-4}$, $\sigma = 0.1$, and selected spatial (N_h) and stochastic (N_ξ) degrees of freedom

	# iter(t)			# iter(t)			# iter(t)			# iter(t)		
N_ξ \ N_h	289($h=\frac{1}{2^4}$)			1089($h=\frac{1}{2^5}$)			4225($h=\frac{1}{2^6}$)			16641($h=\frac{1}{2^7}$)		
n_τ	1	m_ξ+1	n_A	1	m_ξ+1	n_A	1	m_ξ+1	n_A	1	m_ξ+1	n_A
$\sigma_a = 0.1$												
Chebyshev-5+hGS-1												
(3,3) 20	39(3.6)	33(3.3)	33(3.7)	35(11.1)	35(8.7)	29(14.2)	34(29.7)	36(40.6)	35(42.3)	34(162.3)	28(84.0)	34(191.2)
(4,4) 70	41(199.9)	34(173.7)	33(136.4)	45(218.5)	35(190.5)	35(299.8)	46(841.0)	36(670.3)	35(903.4)	45(2093.8)	35(2917.1)	34(4099.5)
(6,3) 84	41(460.9)	34(238.3)	33(325.3)	45(475.0)	35(529.5)	35(610.9)	44(1332.7)	36(1028.1)	35(1708.7)	44(4940.4)	35(4008.7)	34(5843.5)
Chebyshev-10+hGS-1												
(3,3) 20	34(2.8)	26(2.2)	26(3.1)	38(7.1)	30(7.0)	30(10.5)	40(25.7)	31(32.8)	31(37.8)	42(108.7)	31(86.2)	31(152.0)
(4,4) 70	36(191.4)	26(139.4)	26(114.9)	40(186.6)	30(234.7)	30(264.6)	42(488.8)	31(518.1)	31(894.2)	44(1746.1)	32(1518.3)	31(3608.8)
(6,3) 84	35(411.8)	26(193.3)	26(283.1)	40(436.9)	30(354.9)	30(558.3)	42(1221.2)	31(970.6)	31(1529.2)	43(4988.0)	32(3801.3)	31(5417.9)
Cholesky+hGS-1												
(3,3) 20	33(2.6)	25(2.2)	25(3.2)	37(7.9)	27(6.9)	27(11.1)	39(28.4)	29(34.0)	29(34.8)	39(129.7)	29(103.2)	29(150.1)
(4,4) 70	35(177.5)	25(127.6)	25(177.2)	39(282.8)	29(216.6)	27(320.4)	41(603.6)	29(581.4)	29(818.1)	41(1928.1)	29(1582.8)	29(3003.8)
(6,3) 84	35(401.1)	25(332.4)	25(407.1)	39(657.8)	29(445.5)	27(665.0)	41(1156.5)	29(869.2)	29(1706.1)	41(4696.6)	29(2848.4)	29(22402.2)

gies for block (3,3) within the hGS method. The numerical experiments were conducted using multiple mesh sizes and stochastic parameter configurations, with a solver tolerance set to 10^{-8}.

Tables 5.1–5.3 present results for $\beta = 10^{-4}$ with $\sigma_{\mathbb{k}} \in \{0.01, 0.1, 0.4\}$, covering a range from near-deterministic to highly stochastic regimes. Tables 5.1–5.3 demonstrate three key theoretical properties. First, regarding *mesh independence*, iteration counts grow sub-linearly with spatial refinement for fixed stochastic dimension N_ξ, consistent with the spectral bounds established in Section 2. Second, concerning *truncation efficiency*, the $n_\tau = m_\xi + 1$ strategy achieves iteration counts comparable to the full expansion ($n_\tau = n_A$) while avoiding the computational overhead of summing over all c_{ijk} coefficients, thereby validating the hierarchical approximation framework. Third, regarding *smoother comparison*, the 5-step Chebyshev semi-iteration balances convergence rate and per-iteration cost more effectively than either the 10-step variant or direct Cholesky factorization. Across all configurations, the $n_\tau = m_\xi + 1$ truncation consistently delivers performance intermediate between the mean-based approximation ($n_\tau = 1$) and the full expansion, confirming the practical value of the proposed hierarchical preconditioning strategy.

Table 5.4 examines the sensitivity to the regularization parameter β, which bal-

TABLE 5.3

Simulation results showing the total number of iterations from low-rank preconditioned GMRES and the total CPU times (in seconds) using preconditioner with $\beta = 10^{-4}$, $\sigma = 0.4$, and selected spatial (N_h) and stochastic (N_ξ) degrees of freedom

	# iter(t)			# iter(t)			# iter(t)			# iter(t)		
N_h / N_ξ	289($h=\frac{1}{2^4}$)			1089($h=\frac{1}{2^5}$)			4225($h=\frac{1}{2^6}$)			16641($h=\frac{1}{2^7}$)		
n_τ	1	m_ξ+1	n_A	1	m_ξ+1	n_A	1	m_ξ+1	n_A	1	m_ξ+1	n_A
$\sigma_a = 0.4$												
Chebyshev-5+hGS-1												
(3,3) 20	59(5.7)	36(3.6)	27(4.5)	84(13.3)	39(10.1)	29(13.2)	88(50.9)	38(37.8)	36(60.4)	88(398.2)	38(163.3)	28(203.0)
(4,4) 70	100(275.7)	38(198.3)	34(140.1)	112(498.4)	42(299.3)	37(316.8)	118(1792.6)	41(697.2)	36(929.6)	117(7712.6)	41(2655.2)	36(3742.7)
(6,3) 84	86(720.2)	36(414.0)	34(463.0)	96(986.0)	40(596.4)	37(624.8)	102(3051.9)	40(1452.6)	37(1838.1)	102(9946.2)	41(4576.9)	36(7289.7)
Chebyshev-10+hGS-1												
(3,3) 20	69(6.4)	30(2.8)	28(3.3)	77(13.9)	33(8.2)	30(9.9)	81(50.5)	35(34.4)	31(38.1)	83(208.6)	36(148.7)	33(162.3)
(4,4) 70	92(278.4)	32(96.5)	28(125.9)	105(508.4)	36(175.3)	30(269.3)	110(1318.5)	37(781.9)	32(857.9)	110(5212.5)	38(2920.1)	33(4546.2)
(6,3) 84	79(931.3)	30(355.0)	28(411.9)	89(1373.9)	34(529.2)	30(531.0)	95(3528.4)	36(1337.8)	32(1576.4)	97(11001.9)	38(4413.7)	33(5769.8)
Cholesky+hGS-1												
(3,3) 20	69(6.8)	29(3.0)	27(4.0)	77(17.5)	31(8.2)	29(11.7)	81(57.5)	33(35.2)	29(50.3)	83(279.8)	33(180.9)	31(181.8)
(4,4) 70	31(159.1)	31(156.3)	29(350.5)	105(730.2)	33(238.6)	29(331.6)	109(1903.1)	35(726.1)	31(953.3)	109(7600.0)	35(2737.8)	31(3350.7)
(6,3) 84	33(753.1)	33(353.2)	27(425.5)	89(1325.8)	33(493.8)	29(634.8)	95(3122.6)	35(1055.9)	31(1837.7)	97(10923.5)	35(4023.8)	31(6429.3)

ances the tracking term and control cost in the objective functional (2.5). As β decreases from 10^{-2} to 10^{-5}, the optimization problem becomes increasingly dominated by the tracking term. The iteration counts remain remarkably stable across this range, demonstrating that the hierarchical preconditioner effectively handles varying parameter regimes without requiring problem-specific tuning. The $n_\tau = m_\xi + 1$ truncation consistently performs comparably to the full expansion while maintaining reduced computational cost.

TABLE 5.4

Simulation results using the preconditioner with m_ξ=3 p=3, $\sigma_{\mathbb{k}} = 0.2$, $\beta \in \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$ and $N_h = 1089 (h = \frac{1}{2^5})$.

	# iter(t)			# iter(t)			# iter(t)		
N_ξ	(3,3) 20			(4,4) 70			(6,3) 84		
n_τ	1	m_ξ+1	n_A	1	m_ξ+1	n_A	1	m_ξ+1	n_A
Chebyshev-5+hGS-1									
$\beta = 10^{-2}$	52(10.0)	32(7.9)	30(13.5)	60(272.3)	32(149.2)	30(363.0)	56(596.9)	32(338.9)	30(626.9)
$\beta = 10^{-3}$	52(10.3)	34(8.3)	34(15.5)	60(273.8)	35(161.7)	34(411.4)	56(594.0)	34(360.7)	34(709.7)
$\beta = 10^{-4}$	43(11.7)	37(7.6)	35(15.6)	62(279.7)	37(172.1)	35(299.2)	58(629.0)	37(393.4)	35(731.8)
$\beta = 10^{-5}$	41(11.2)	38(7.9)	37(17.0)	60(270.6)	38(175.9)	37(317.9)	56(589.0)	38(413.3)	38(795.5)
Chebyshev-10+hGS-1									
$\beta = 10^{-2}$	48(8.7)	30(6.1)	29(11.9)	56(420.8)	29(147.9)	29(359.7)	52(577.6)	34(370.8)	24(653.1)
$\beta = 10^{-3}$	48(8.8)	31(6.4)	30(12.6)	56(432.3)	31(151.6)	30(406.9)	52(554.8)	31(339.6)	30(675.8)
$\beta = 10^{-4}$	48(8.8)	30(7.1)	30(19.4)	56(419.0)	31(151.2)	30(390.6)	52(575.0)	31(339.6)	30(676.1)
$\beta = 10^{-5}$	47(8.7)	30(7.0)	30(13.2)	54(406.4)	30(146.0)	30(383.9)	51(554.0)	30(333.0)	30(675.1)
Chol+hGS-1									
$\beta = 10^{-2}$	47(12.4)	29(6.9)	27(11.9)	55(380.1)	29(214.4)	27(354.4)	51(975.8)	29(483.9)	27(595.1)
$\beta = 10^{-3}$	47(12.0)	29(6.4)	27(11.9)	55(396.0)	29(215.9)	27(364.5)	51(876.9)	29(455.6)	27(596.0)
$\beta = 10^{-4}$	47(10.4)	29(7.4)	27(12.2)	55(387.9)	29(216.0)	27(343.9)	53(847.9)	29(432.8)	27(595.5)
$\beta = 10^{-5}$	45(10.0)	29(6.8)	27(12.0)	53(369.6)	29(205.5)	27(359.3)	51(829.3)	29(438.9)	27(594.1)

5.2. Time-dependent case. This subsection evaluates the all-at-once preconditioner (3.9) for time-dependent optimal control problems. The discretization results in KKT systems of dimension $N_t \times N_\xi \times N_h$, where N_t denotes the number of time steps. We investigate the scalability with respect to: (i) mesh refinement (Table 5.5), (ii) regularization parameter β (Table 5.6), (iii) variance coefficient $\sigma_{\mathbb{k}}$ (Table 5.7), (iv) temporal discretization τ (Table 5.8), and (v) stochastic dimension (m_ξ, p) (Ta-

ble 5.9). Based on the steady-state findings, we employ the Chebyshev-5+hGS-1 configuration unless otherwise noted, reporting results for both strict (10^{-6}) and moderate (10^{-4}) tolerances to illustrate practical convergence behavior. As in the steady-state case, we use $N = 1$ in the Richardson iteration (Algorithm 6) for the (3,3)-block.

From our observations in the steady-state problem, a combination of a 5-step Chebyshev smoother with one step of our hGS method achieves a good balance between the computational cost of matrix operations and GMRES iterations; thus, we typically adopt this combination when testing time-dependent cases as well.

As indicated in Table 5.5, the 5-step Chebyshev smoother yields consistent iteration counts compared to either the 10-step smoother or direct Cholesky decomposition. As the spatial discretization is refined from $nc = 3$ to $nc = 6$, representing a growth from 116,640 to 6,084,000 DoF, the iteration count for $n_\tau = m_\xi + 1$ exhibits sub-linear growth consistent with the near mesh-independence predicted by the spectral theory in Section 3. The $n_\tau = m_\xi + 1$ truncation achieves iteration counts comparable to the full expansion ($n_\tau = n_A$) while reducing the cost of assembling and applying the (3,3)-block preconditioner—a trade-off that becomes increasingly favorable as the problem dimension grows.

TABLE 5.5

Simulation results using hGS Method with different truncation settings n_A with the model with time-dependent diffusion constraint for different tolerance and mesh size at the random setting with m_ξ=3, $p = 3$ $\sigma_{\mathbb{k}}$=0.2, $\beta = 10^{-4}$ and number of steps=8 ($\tau = \frac{1}{2^3}$).

nc	DoF	tol=10^{-6}			tol=10^{-4}		
n_τ		1	m_ξ+1	n_A	1	m_ξ+1	n_A
3	116,640	57(11.9)	39(8.8)	39(15.7)	41(6.8)	31(5.7)	29(8.8)
4	416,160	75(39.3)	47(29.6)	45(39.6)	55(21.7)	35(14.9)	35(24.7)
5	1,568,160	83(150.9)	53(103.3)	51(158.6)	66(90.6)	43(58.8)	39(92.1)
6	6,084,000	84(570.8)	55(383.3)	53(585.1)	68(401.4)	42(252.2)	42(424.1)

Table 5.6 examines four orders of magnitude for β, ranging from 10^{-2} (control-dominant) to 10^{-8} (tracking-dominant). The mean-based preconditioner ($n_\tau = 1$) exhibits strong dependence on β, with iteration counts decreasing as β decreases (since smaller β yields problems dominated by the simpler tracking term). In contrast, the $n_\tau = m_\xi + 1$ truncation maintains stable iteration counts across all tested values, demonstrating that the hGS preconditioner automatically adapts to the problem structure without manual parameter tuning. This robustness confirms the theoretical framework's applicability across diverse parameter regimes.

TABLE 5.6

Simulation results using hGS Method with different truncations setting n_A with the model with time-dependent diffusion constraint for different tolerance and β at mesh size $N_h = \frac{1}{2^5}$ random setting with m_ξ=3, $p = 3$, $\sigma_{\mathbb{k}}$=0.2, and number of steps=8 ($\tau = \frac{1}{2^3}$).

β	DoF	tol=10^{-6}			tol=10^{-4}		
n_τ		1	m_ξ+1	n_A	1	m_ξ+1	n_A
10^{-2}	1,568,160	107(206.1)	69(136.3)	69(214.8)	74(102.0)	48(70.4)	47(112.6)
10^{-3}	1,568,160	116(164.4)	74(109.4)	74(174.6)	66(90.6)	42(60.7)	42(100.0)
10^{-6}	1,568,160	97(137.0)	77(115.6)	77(182.5)	57(77.4)	45(65.3)	44(103.7)

Table 5.7 explores the range $\sigma_{\mathbb{k}} \in \{1\%, 2\%, 5\%, 10\%, 20\%, 40\%\}$, spanning from nearly deterministic to highly uncertain regimes. The mean-based preconditioner ($n_\tau = 1$) exhibits significant degradation as uncertainty increases, whereas the $n_\tau = m_\xi + 1$ truncation maintains stable iteration counts across the entire range. Especially when $\sigma_{\mathbb{k}}$ increases from 20% to 40%, the mean-based preconditioner performs poorly with a large number of iterations, but the hGS method maintains robust performance. This robustness confirms that the hierarchical preconditioner effectively captures the essential stochastic structure without requiring full expansion of all coupling coefficients. Table 5.8 investigates the all-at-once system scalability by varying N_t from 4

TABLE 5.7

Simulation results using hGS Method with different truncations setting n_A with the model with time-dependent diffusion constraint for different tolerance and $\sigma_{\mathbb{k}}$ at $\beta = 10^{-4}$, mesh size $N_h = \frac{1}{2^5}$ random setting with m_ξ=3, $p = 3$, and number of steps=8 ($\tau = \frac{1}{2^3}$).

$\sigma_{\mathbb{k}}$	DoF	tol=10^{-6}			tol=10^{-4}		
n_τ		1	m_ξ+1	n_A	1	m_ξ+1	n_A
1%	1,568,160	51(77.7)	51(85.3)	51(133.9)	43(68.6)	41(72.7)	41(112.4)
2%	1,568,160	53(80.8)	51(82.2)	51(130.0)	43(70.2)	41(71.1)	41(113.7)
5%	1,568,160	59(90.2)	51(82.1)	51(130.9)	47(78.3)	41(72.6)	41(113.1)
10%	1,568,160	65(99.2)	51(82.4)	51(130.4)	53(86.2)	42(70.6)	41(116.5)
20%	1,568,160	83(150.9)	53(103.3)	51(158.6)	66(90.6)	43(58.8)	39(92.1)
40%	1,568,160	129(200.9)	57(92.3)	51(130.5)	100(179.3)	47(82.3)	43(119.2)

to 256 (time steps $\tau \in \{1/4, 1/16, 1/64, 1/256\}$), corresponding to total system sizes ranging from 784,080 to over 12.5 million DoF. As the temporal resolution increases, the coupled space-time-stochastic system grows proportionally, yet the $n_\tau = m_\xi + 1$ truncation maintains sub-linear iteration growth relative to system size. The computational time scales approximately linearly with DoF, confirming the efficiency of the all-at-once preconditioner for massively coupled systems.

TABLE 5.8

Simulation results using hGS Method with different truncations setting n_A with the model with time-dependent diffusion constraint for different tolerance and $\sigma_{\mathbb{k}}$ at $\beta = 10^{-4}$, mesh size $N_h = \frac{1}{2^5}$ random setting with $m_\xi = 3$, $p = 3$, and number of steps=8 ($\tau = \frac{1}{2^3}$).

τ	DoF	tol=10^{-6}			tol=10^{-4}		
n_τ		1	m_ξ+1	n_A	1	m_ξ+1	n_A
$1/2^2$	784,080	81(78.6)	51(51.1)	51(71.1)	66(53.9)	43(37.6)	42(57.2)
$1/2^4$	3,136,320	85(238.2)	55(166.4)	53(258.7)	68(174.2)	44(124.3)	43(203.7)
$1/2^6$	12,545,280	101(1029.8)	67(733.3)	65(1205.8)	78(840.5)	53(604.2)	52(1028.6)

Finally, Table 5.9 compares three gPC configurations: $(m_\xi, p) \in \{(3,3), (4,4), (6,3)\}$, yielding $N_\xi \in \{20, 70, 84\}$ basis functions. Table 5.9 varies the stochastic discretization parameters (m_ξ, p), exploring both the number of random variables and polynomial order. As N_ξ increases from 20 to 84, the iteration count for $n_\tau = m_\xi + 1$ grows modestly, demonstrating near-independence from the stochastic discretization level. This behavior confirms the effectiveness of the hierarchical truncation strategy in maintaining spectral properties across varying gPC expansion settings.

The time-dependent experiments establish that the proposed all-at-once preconditioner maintains robust performance across a wide range of problem parameters

TABLE 5.9

Simulation results using hGS Method with different truncations setting n_A with the model with time-dependent diffusion constraint for different tolerance and stochastic setting at $\beta = 10^{-4}$, mesh size $N_h = \frac{1}{2^5}$ random setting with $\sigma_{\mathbb{k}}$=0.2, and number of steps=8 ($\tau = \frac{1}{2^3}$).

(m_ξ,P)	DoF	tol=10^{-6}			tol=10^{-4}		
n_τ		1	m_ξ+1	n_A	1	m_ξ+1	n_A
20 (3,3)	1,568,160	83(150.9)	53(103.3)	51(158.6)	66(90.6)	43(58.8)	39(92.1)
70 (4,4)	5,488,560	95(2435.3)	53(1397.4)	52(2964.7)	74(1809.2)	43(1044.9)	43(2476.2)
84 (6,3)	6,586,272	93(4092.4)	56(2670.3)	52(5207.4)	70(3059.4)	43(2107.7)	43(4413.4)

and discretization levels. Four key findings emerge from these results. First, regarding *near mesh-independence*, iteration growth remains sub-linear with spatial refinement (Table 5.5), consistent with the spectral bounds derived in Section 4. Second, concerning *parameter robustness*, the hGS method adapts automatically to varying β (Table 5.6) and $\sigma_{\mathbb{k}}$ (Table 5.7) without manual tuning, demonstrating the preconditioner's insensitivity to problem-specific parameters. Third, in terms of *scalability*, the preconditioner handles systems with over 12.5 million DoF (Table 5.8) and high stochastic dimensions (Table 5.9) efficiently, validating the computational feasibility for large-scale applications. Fourth, regarding *truncation efficiency*, the $n_\tau = m_\xi + 1$ strategy consistently delivers performance comparable to the full expansion at significantly reduced cost, validating the theoretical analysis in Section 4. These results demonstrate that the hierarchical preconditioning framework extends seamlessly from steady-state to time-dependent problems, providing a practical and theoretically-grounded solution for large-scale stochastic optimal control.

6. Conclusions. This paper has successfully designed, analyzed, and implemented a novel hierarchical preconditioning strategy for large-scale stochastic optimal control problems. Our approach leverages a truncated stochastic expansion within a block-structured preconditioner for the Karush-Kuhn-Tucker (KKT) system, striking an effective balance between computational cost and preconditioning quality. Numerical results confirm that the proposed hGS method consistently outperforms both standard mean-based preconditioners and computationally intensive full-expansion methods across a wide range of problem parameters.

A key contribution of this work is the extension of this framework to time-dependent problems. We developed and tested a tailored hGS preconditioner within an all-at-once discretization scheme, demonstrating the versatility and effectiveness of our approach for these more challenging, large-scale scenarios. Comprehensive numerical experiments on benchmark problems with Dirichlet boundary conditions have validated the robustness and numerical efficiency of the proposed algorithms. Future research could involve extending this preconditioning framework to problems with more complex PDE constraints, such as the Navier-Stokes equations, or investigating its application to optimal control problems with inequality constraints.

REFERENCES

[1] H. ANTIL, S. DOLGOV, AND A. ONWUNTA, *TTRisk: Tensor train decomposition algorithm for risk averse optimization*, Numerical Linear Algebra with Applications, 30 (2023), p. 2481.

[2] I. BABUŠKA, R. TEMPONE, AND G. E. ZOURARIS, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM J. Numer. Anal., 42 (2004), pp. 800–825, https://doi.org/10.1137/S0036142902418680.

[3] P. Benner, S. Dolgov, A. Onwunta, and M. Stoll, *Low-rank solvers for unsteady Stokes–Brinkman optimal control problem with random data*, Computational Methods in Applied Mechanics, 304 (2016), pp. 26–54.
[4] P. Benner, S. Dolgov, A. Onwunta, and M. Stoll, *Low-rank solution of an optimal control problem constrained by random Navier-Stokes equations*, International Journal for Numerical Methods in Fluids, 92 (2020), pp. 1653–1678, https://doi.org/10.1002/fld.4843.
[5] P. Benner, A. Onwunta, and M. Stoll, *Block-diagonal preconditioning for optimal control problems constrained by PDEs with uncertain inputs*, SIAM Journal on Matrix Analysis and Applications, 37 (2016), pp. 491–518, https://doi.org/10.1137/15M1018502.
[6] A. Bespalov, D. Loghin, and R. Youngnoi, *Truncation preconditioners for stochastic Galerkin finite element discretizations*, SIAM Journal on Scientific Computing, 43 (2021), pp. S92–S116, https://doi.org/10.1137/20M1345645.
[7] A. k. Björck, *Numerical methods for least squares problems*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996, https://doi.org/10.1137/1.9781611971484, https://doi.org/10.1137/1.9781611971484.
[8] P. Chen and A. Quarteroni, *Weighted reduced basis method for stochastic optimal control problems with elliptic PDE constraint*, SIAM/ASA J. Uncertainty Quantification, 2 (2014), pp. 364–396, https://doi.org/10.1137/130940517.
[9] H. Elman, D. Silvester, and A. Wathen, *Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics*, Oxford University Press, 2014, https://doi.org/10.1093/acprof:oso/9780199678792.001.0001.
[10] R. G. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag New York, Inc., New York, NY, USA, 1991. (Revised edition by Dover Publications, 2003).
[11] G. H. Golub and R. S. Varga, *Chebyshev semi-iterative methods, successive overrelaxation iterative methods, and second order richardson iterative methods*, Numerische Mathematik, 3 (1961), pp. 157–168.
[12] D. Kouri, M. Heinkenschloss, D. Ridzal, and B. van Bloemen Waanders, *A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty*, SIAM Journal on Scientific Computing, 35 (2013), pp. A1847–A1879, https://doi.org/10.1137/120892362.
[13] D. P. Kouri, *A multilevel stochastic collocation algorithm for optimization of pdes with uncertain coefficients*, SIAM/ASA Journal on Uncertainty Quantification, 2 (2014), pp. 55–81.
[14] O. Le Maître and O. M. Knio, *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Scientific Computation, Springer, 2010.
[15] H.-C. Lee and J. Lee, *A stochastic Galerkin method for stochastic control problems*, Commun. Comput. Phys., 14 (2013), pp. 77–106, https://doi.org/10.4208/cicp.241011.150612a.
[16] G. J. Lord, C. E. Powell, and T. Shardlow, *An Introduction to Computational Stochastic PDEs*, Cambridge Texts in Applied Mathematics, Cambridge University Press, 2014.
[17] J. Martínez-Frutos and F. Periago Esparza, *Optimal Control of PDEs under Uncertainty: An Introduction with Application to Optimal Shape Design of Structures*, Springer Briefs in Mathematics, Springer Cham, 2018, https://doi.org/10.1007/978-3-319-98210-6.
[18] J. W. Pearson, M. Stoll, and A. J. Wathen, *Regularization-robust preconditioners for time-dependent PDE-constrained optimization problems*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 1126–1152, https://doi.org/10.1137/110847949, https://doi.org/10.1137/110847949.
[19] C. E. Powell and H. C. Elman, *Block-diagonal preconditioning for spectral stochastic finite-element systems*, IMA J. Numer. Anal., 29 (2009), pp. 350–375, https://doi.org/10.1093/imanum/drn014.
[20] T. Rees, M. Stoll, and A. Wathen, *All-at-once preconditioning in pde-constrained optimization*, Kybernetika, 46 (2010), pp. 341–360.
[21] E. Rosseel and G. N. Wells, *Optimal control with stochastic PDE constraints and uncertain controls*, Comput. Methods Appl. Mech. Engrg., 213–216 (2012), pp. 152–167, https://doi.org/10.1016/j.cma.2011.11.026.
[22] Y. Saad, *A flexible inner-outer preconditioned GMRES algorithm*, SIAM J. Sci. Comput., 14 (1993), pp. 461–469, https://doi.org/10.1137/0914028.
[23] D. Silvester, H. Elman, and A. Ramage, *Incompressible Flow and Iterative Solver Software (IFISS) version 3.5*, September 2016, http://www.manchester.ac.uk/ifiss/.
[24] B. Sousedík and R. G. Ghanem, *Truncated hierarchical preconditioning for the stochastic Galerkin FEM*, International Journal for Uncertainty Quantification, 4 (2014), pp. 333–348, https://doi.org/10.1615/Int.J.UncertaintyQuantification.2014007353.
[25] B. Sousedík, R. G. Ghanem, and E. T. Phipps, *Hierarchical Schur complement precondi-*

tioner for the stochastic Galerkin finite element methods, Numerical Linear Algebra with Applications, 21 (2014), pp. 136–151, https://doi.org/10.1002/nla.1869.

[26] H. Tiesler, R. Kirby, D. Xiu, and T. Preusser, *Stochastic collocation for optimal control problems with stochastic PDE constraints*, SIAM J. Control. Optim., 50 (2012), pp. 2659–2682, https://doi.org/10.1137/110835438.

[27] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, American Mathematical Society, Providence, Rhode Island, 2010. (Graduate Studies in Mathematics, Volume 112).

[28] H. Wang, D. Wu, C. Yu, and K. L. Teo, *A duality-based approach for linear parabolic optimal control problems*, Optimal Control Appl. Methods, 45 (2024), pp. 1140–1165, https://doi.org/10.1002/oca.3094, https://doi-org.ezproxy.lib.lehigh.edu/10.1002/oca.3094.

[29] A. Wathen and T. Rees, *Chebyshev semi-iteration in preconditioning for problems including the mass matrix*, Electron. Trans. Numer. Anal., 34 (2008/09), pp. 125–135.

[30] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, 2010.

[31] R. Youngnoi, *Improving convergence in stochastic Galerkin finite element methods via truncation preconditioners*, SIAM Journal on Scientific Computing, 47 (2025), pp. A1937–A1963, https://doi.org/10.1137/24M1685699.