Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

RESEARCH ARTICLE

Preconditioners for robust optimal control problems under uncertainty

Fabio Nobile[1] | Tommaso Vanzan[1]

[1]CSQI Chair, Institute de Mathématiques, École Polytechnique Fédérale de Lausanne

Correspondence
Tommaso Vanzan, CSQI Chair, Institute de Mathématiques, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. Email: tommaso.vanzan@epfl.ch

Summary

The discretization of robust quadratic optimal control problems under uncertainty using the finite element method and the stochastic collocation method leads to large saddle-point systems, which are fully coupled across the random realizations. Despite its relevance for numerous engineering problems, the solution of such systems is notoriously challenging. In this manuscript, we study efficient preconditioners for all-at-once approaches using both an algebraic and an operator preconditioning framework. We show in particular that for values of the regularization parameter not too small, the saddle-point system can be efficiently solved by preconditioning in parallel all the state and adjoint equations. For small values of the regularization parameter, robustness can be recovered by the additional solution of a small linear system, which however couples all realizations. A mean approximation and a Chebyshev semi-iterative method are proposed to solve this reduced system. We consider a random elliptic partial differential equation whose diffusion coefficient $\kappa(x,\omega)$ is modeled as an almost surely continuous and positive random field, though not necessarily uniformly bounded and coercive. We further provide estimates of the dependence of the spectrum of the preconditioned system matrix on the statistical properties of the random field and on the discretization of the probability space. Such estimates involve either the first or second moment of the random variables $1/\min_{x\in\overline{D}}\kappa(x,\omega)$ and $\max_{x\in\overline{D}}\kappa(x,\omega)$, where D is the spatial domain. The theoretical results are confirmed by numerical experiments, and implementation details are further addressed.

KEYWORDS:
Optimal control problems under uncertainty, parameter robust preconditioners, random PDE, lognormal fields

ACKNOWLEDGMENTS

The two authors acknowledge funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N. 800898, project ExaQUte – EXAscale Quantification of Uncertainties for Technology and Science Simulation. The first author acknowledge funding also from the Swiss National Science Foundation under the Project n. 172678 "Uncertainty Quantification techniques for PDE constrained optimization and random evolution equations".

1 | INTRODUCTION

Optimal Control Problems (OCPs) constrained by deterministic Partial Differential Equations (PDEs) have been extensively studied in the last decades since they are essential tools in the design of complex engineering systems, see, e.g., the monographs [1–3]. However, the physical system under study is often affected by uncertainties, either due to a lack of knowledge on some parameters defining the model or due to an intrinsic variability usually characterized in probabilistic terms, thus leading to random PDEs. To have more reliable results, it is important to account for the ubiquitous uncertainty in nature by constraining an OCP to such random PDEs, and optimizing statistical measures, often called risk measures [4, Chapter 6.3], of a quantity of interest [5, 6]. This leads to an OCP Under Uncertainty (OCPUU), and in this work we focus on problems whose structure is

$$\begin{aligned} &\min_{u\in U} \mathcal{R}\left[Q(y(\omega),u)\right] \\ &\text{s.t. } y(\omega)\in V \text{ satisfies} \\ &a_\omega(y(\omega),v)=(u,v) \quad \forall v\in V, \text{ a.e. } \omega\in\Omega. \end{aligned} \tag{1}$$

where u is the unknown deterministic control, $y(\omega)$ is the state variable which satisfies the random PDE constraint expressed in a weak form for almost every realization ω of the randomness, Q is the quantity of interest and $\mathcal{R}$ is a risk measure.

There are two possible paradigms to minimize numerically such functionals involving statistical measures. The first one, called Stochastic Approximation (SA) method [4, Chapter 5.9], includes iterative methods that at each iteration draw new realizations of the underlying randomness, independent from the previous ones. Examples of such approaches are the stochastic gradient method and its variants, which have been recently studied for OCPUU [7–10].

We adopt here the second approach called Sample Average Approximation (SAA) method [4, Chapter 5.1], in which the original objective functional is replaced by an accurate approximation obtained discretizing once and for all the probability space using Stochastic Collocation methods (SC), with Monte Carlo, Quasi-Monte Carlo [11], or Gaussian quadrature formulae. We do not consider approximations based on Multilevel Monte Carlo [12] and sparse grids formulae [13, 14], since they may not preserve the convexity of the objective functional. After discretization, the optimality conditions become an extremely large (possibly nonlinear) global system which requires efficient tailored solvers.

The goal of this manuscript is to analyze optimal preconditioners for the linear optimality system obtained from a SC discretization of a robust quadratic OCP constrained by a random PDE. Although we restrict to a quadratic OCP, we remark that the preconditioners developed in this work will also be useful for more complex, non-quadratic, OCPs to precondition the linear system obtained at each iteration of a nonlinear optimization algorithm (e.g. Newton's method) see, e.g., Kouri et al. [15]. We are in particular interested in studying the dependence of the spectrum of the preconditioned system with respect to the regularization parameter of the control, the underlying randomness of the system, and the level of discretization in probability.

We stress that the saught control is *deterministic*, since the properties of the global optimality system depend strongly on whether the control is *stochastic* or *deterministic*. In the first case, one assumes that the realization of the randomness is observable, and thus an optimal control can be established for each single random realization, leading to a stochastic optimal control $u(\omega)$.

On the one hand, such problems are easier to solve using SC methods since the global linear system is often decoupled across all random samples, so that one actually needs to solve a sequence of independent, deterministic OCPs, one for each sample, for which optimal preconditioners are available [16–22]. On the other hand, a discretization based on Stochastic Galerkin methods (SG) [23, 24] leads to a fully coupled saddle point system across the random components, even when the optimal control is stochastic. Optimal preconditioners for an OCP with stochastic control combined with a SG discretization have been analyzed by Benner et al. [25].

Nevertheless, the use of a stochastic control may not be realistic either because the randomness is not observable, as in the case of subsurface flows (unless one is willing to drill for geotechnical investigation), or because a unique control has to be designed before the randomness can be observed (e.g. the optimal shape of a building is designed before it is eventually subjected to random incoming wind conditions). In these cases, that is, the ones we are interested in, one computes a unique *deterministic* control valid for all random realizations. This setting is often called robust OCPUU. These problems are harder, since the global system fully couples all the random samples. For a SC discretization, the global optimality system involves N state equations, N adjoint equations and a single optimality equation, where N is the number of collocation points.

Let us now review previous works concerning solution strategies for robust OCPUU. Gradient-based approaches, which permit to obtain the solution iteratively solving the three sets of equations (state, adjoint, optimality) sequentially at each iteration, have been combined with a Multilevel Monte Carlo estimators in Van Barel et al. [12]. Rosseel et al. [26] derive optimality conditions for a general quadratic OCP constrained by a random elliptic PDE. Their control can be either stochastic or deterministic.

They interestingly remark that a discretization using SC leads to a global system coupling all realizations, unless the control is fully stochastic, as previously mentioned. Hence, they focus on SG and present numerical results using either a mean-based preconditioner[27] or collective smoothing multigrid[21]. An MG/OPT algorithm based on a hierarchy of sparse grid approximations of the objective functional has been proposed in Kouri et al.[28] Another MG/OPT algorithm based on a classical hierarchy of geometric meshes has been analyzed in Vandewalle et al.[29] Other works aimed to reduce the computatational costs based on trust-region algorithms are Kouri et al.[13] and Zahr et al.[30].

Despite the remarks of Rosseel et al.[26] about the loss of non-intrusivity of SC methods for robust OCPUU, we are interested in analysing SC for the following reasons. First, SCM maintains its advantages in terms of applicability with respect to general parameter distributions and ease of implementation[31, Chapter 10]. Second, one can construct preconditioners whose action can be fully parallelized across the realizations of the randomness, one example being the preconditioner proposed by Kouri et al.[15] That is, while a global system involving all realizations has to be solved, the preconditioner does not couple the realizations, as it requires to solve approximately (i.e. to precondition) independently each forward and adjoint problem. In this perspective, this preconditioner has favourable properties in terms of parallelization and memory distribution in a high performance setting. In this manuscript, we analyse, among others, the performance of the preconditioner proposed in Kouri et al.[15], by providing theoretical estimates for the spectrum of the preconditioned system.

As the regularization parameter on the control, denoted by β, gets smaller, the preconditioner introduced by Kouri et al.[15] becomes inefficient. Thus, we introduce a first new preconditioner, named P_{LR}, which still preconditions each state and adjoint equation in parallel, but requires the additional solution of a small linear system. We partially characterize the spectrum of the preconditioned system and show numerically its β-robustness. Finally, to derive a provably β-robust preconditioner, we study the optimality system at the fully-continuous level, and our analysis leads to a second new preconditioner, named P_{OP}, for which a complete theory is available. Both the first and second preconditioner require the inversion of the sum of all inverses of the stiffness matrices. A mean approximation, combined possibly with a Chebyshev semi-iteration, is shown to be sufficient to efficiently approximate this inverse for quite a wide range of parameters, leading to *practical* $P_{\mathrm{LRM}}, P_{\mathrm{LRC}}, P_{\mathrm{OPM}}, P_{\mathrm{OPC}}$ preconditioners, where the subscript M stands for "mean" and C for Chebyshev.

We remark that the development of robust preconditioners for small values of the regularization parameter is not obvious and poses some interesting mathematical and computational challenges which, surprisingly, are similar to those encountered in deterministic OCP when the control acts locally, either on a portion of the domain[32], or on a portion of the boundary[33].

We further stress that the analysis does not assume that the random bilinear form is uniformly bounded and coercive with respect to the randomness, which is a frequent simplifying hypothesis in the literature[25, 27, 34]. Hence, the results will also cover the case of log-normally distributed random fields, which are common models in engineering applications, and they will cast light on how the preconditioners' performance is affected by the variance of the random fields and by the level of discretization in the probability space.

To develop optimal preconditioners for robust OCPUU, we rely on two different approaches. The first one used to derive the preconditioner P_{LR} is algebraic and has its roots in the seminal work of Murphy et al.[35], who proposed an optimal, but expensive, preconditioner for saddle point matrices which relies on the exact Schur complement. For deterministic OCP, several preconditioners based on approximations of the exact Schur complements have been studied in the last decade[16–18, 36–38]. This first approach is suitable for a L^2 penalization on the norm of the control. The second approach, used to derive the preconditioner P_{OP}, consists in the so-called "operator preconditioning" paradigm, and is based on identifying the saddle point system as a linear operator acting between Hilbert spaces, and finding proper weighted-norms such that the continuity constants of the map and of its inverse are independent of the parameters of interest. We refer the interested reader to works of Malek et al.[39], Zulehner[19], Mardal et al.[20], Kirby[40] and Khan et al.[41]. While studying this approach, we will discuss the well-posedness of the OCPUU and the development of robust preconditioners at the continuous level for log-normal fields, without relying on the framework developed in Gittelson et al.[42, 43]. This second approach requires a H^1 penalization on the norm of the control.

The manuscript is organized as follows. In Section 2 we introduce the notation, while in Section 3 we define the model problem, provide sufficient conditions for well-posedness and derive the optimality conditions. Section 4 introduces the discretization both in probability and in physical space. Section 5 deals with algebraic preconditioners for saddle point matrices based on approximations of the Schur complement. Section 6 derives preconditioners using the operator preconditioning approach. Finally, Section 7 presents numerical experiments validating the theoretical results.

2 | NOTATION

Let $D \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, be a Lipschitz bounded domain and $(\Omega, \mathcal{F}, \mathbb{P})$ a complete probability space. For every $p \in [1, \infty]$, $L^p(D)$ denotes the space of $p-$Lebesgue integrable functions over D and $H^1(D)$ is the Sobolev space

$$H^1(D) := \left\{ v \in L^2(D) : \partial_{x_i} v \in L^2(D), \quad \text{for } i = 1, \dots, d \right\}.$$

The natural space for the analysis is $H_0^1(D)$, which is the subspace of $H^1(D)$ containing functions that vanish on ∂D, equipped with the norm $\|y\|_{H_0^1(D)} := \|\nabla y\|_{L^2(D)}$. The topological dual of $H_0^1(D)$ is $H^{-1}(D)$. We denote by C_P the Poincaré constant so that $\|v\|_{L^2(D)} \leq C_P \|v\|_{H_0^1(D)}$, $\forall v \in H_0^1(D)$. For the sake of brevity, we will denote $H_0^1(D)$ and $H^{-1}(D)$ by Y and Y'. Given an integer $N \in \mathbb{N}$ and a Hilbert space V, we denote by $\underline{V} := \prod_{i=1}^N V$ the Cartesian product of N copies of V. Given a Banach space U, the duality pairing between U and U' is denoted by $\langle \cdot, \cdot \rangle$. The specific choice of U will be clear from the context. Further, let $L^p(\Omega, \mathcal{F}, \mathbb{P}; V)$ be the Bochner space[44]

$$L^p(\Omega, \mathcal{F}, \mathbb{P}; V) := \left\{ v : \Omega \to V, \ v \text{ strongly measurable}, \int_\Omega \|v(\cdot, \omega)\|_V^p d\mathbb{P}(\omega) < +\infty \right\},$$

henceforth noted $L^p(\Omega, V)$, and equipped with the norm $\|v\|_{L^p(\Omega,V)} := (\int_\Omega \|v(\cdot, \omega)\|_V^p d\mathbb{P}(\omega))^{\frac{1}{p}}$. For a Hilbert space V, $L^2(\Omega, V)$ is a Hilbert space as well, equipped with the scalar product $(u, v)_{L^2(\Omega,V)} := \int_\Omega (u(\cdot, \omega), v(\cdot, \omega))_V d\mathbb{P}(\omega)$. To stress better the dependence of function-valued random variables on an elementary random event ω, we will use the notation $v_\omega = v(\cdot, \omega)$ for almost every (a.e.) $\omega \in \Omega$. The expectation operator $\mathbb{E} : L^1(\Omega) \to \mathbb{R}$ is defined as

$$\mathbb{E}[X] = \int_\Omega X(\omega) d\mathbb{P}(\omega), \quad \forall X \in L^1(\Omega).$$

For $X \in L^2(\Omega)$, the variance $\mathbb{V} : L^2(\Omega) \to \mathbb{R}^+$ and standard deviation $\mathbb{S} : L^2(\Omega) \to \mathbb{R}^+$ are defined as

$$\mathbb{V}[X] := \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \int_\Omega (X - \mathbb{E}[X])^2 d\mathbb{P}(\omega), \quad \text{and} \quad \mathbb{S}[X] := \sqrt{\mathbb{V}[X]}.$$

We will use repeatedly the Woodbury identity,

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1},$$

where $A \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{r \times r}$, $U \in \mathbb{R}^{n \times r}$, $V \in \mathbb{R}^{r \times n}$, with A and C invertible. Finally, the spectrum of a matrix H is denoted with $\sigma(H)$.

3 | PROBLEM SETTING

We consider the elliptic random Partial Differential Equation (PDE)

$$\begin{aligned} -\text{div}(\kappa(x, \omega) \nabla y(x, \omega)) &= \phi(x), \qquad x \in D, \ \omega \in \Omega, \\ y(x, \omega) &= 0, \qquad x \in \partial D, \ \omega \in \Omega, \end{aligned} \tag{2}$$

where $\phi(x)$ is a deterministic force term and ω is an elementary random event. Equation (2) is commonly used to describe subsurface flows and heat diffusion in random media. The random field $\kappa(x, \omega)$ models the statistical properties of the medium. For instance, interpreting (2) as a subsurface flow model, $\kappa(x, \omega)$ represents a random permeability field.

Assumption 1 (On the random diffusion field)**.** The random diffusion field κ has almost surely (a.s.) continuous and positive realizations and the map $\omega \mapsto \kappa(\cdot, \omega) \in C^0(\overline{D})$ is measurable. Thus, the random variables $\kappa_{\min}(\omega) := \min_{x \in \overline{D}} \kappa(x, \omega)$ and $\kappa_{\max}(\omega) := \max_{x \in \overline{D}} \kappa(x, \omega)$ are well-defined. Further, there exists a $p \in [1, \infty]$ such that both $\kappa_{\max}$ and $\frac{1}{\kappa_{\min}}$ are in $L^p(\Omega)$.

These assumptions are clearly verified with $p = \infty$ by a continuous and uniformly bounded random field, i.e. if there exist K_1, $K_2 \in \mathbb{R}^+$ such that

$$K_1 \leq \kappa(x, \omega) \leq K_2, \quad \forall x \in D, \text{ a.e. } \omega \in \Omega.$$

Another instance is the log-normal random field $\kappa(x,\omega) = \exp(g(x,\omega))$, where $g(x,\omega)$ is a Gaussian field with covariance function $\mathrm{cov}[g](x,y) := k(\|x-y\|)$, and $k(\cdot)$ is a Lipschitz function. Both $\kappa_{\min}$ and $\kappa_{\max}$ are in $L^p(\Omega)$ for every $p \in [1,\infty)$[45]. The log-normal field is commonly used in hydrology applications[46, 47] and has been extensively studied[42, 43, 45, 48–50].
For a.e. $\omega \in \Omega$, $a_\omega(\cdot,\cdot) : Y \times Y \to \mathbb{R}$, $a_\omega(u,v) := \int_D \kappa(x,\omega)\nabla u(x)\nabla v(x)dx$ is a symmetric, continuous and coercive bilinear form, but not necessarily uniformly in ω due to Assumption 1. It holds

$$\kappa_{\min}(\omega)\|u\|_Y^2 \le a_\omega(u,u) \le \kappa_{\max}(\omega)\|u\|_Y^2. \tag{3}$$

The weak formulation of (2) on Y for a.e. $\omega \in \Omega$ is

$$\text{find } y_\omega \in Y \text{ s.t. } a_\omega(y_\omega, v) = \langle \phi, v\rangle, \quad \text{for every } v \in Y, \text{for a.e. } \omega \text{ in } \Omega. \tag{4}$$

Due to Assumption 1, the following classical result holds[45, 48, 51].

Lemma 1. Problem (4) has a unique solution y_ω for a.e. $\omega \in \Omega$. Further,

$$\|y_\omega\|_Y \le \frac{\|\phi\|_{Y'}}{\kappa_{\min}(\omega)}, \quad \text{for a.e. } \omega \in \Omega,$$
$$\|y\|_{L^p(\Omega,Y)} \le \|\phi\|_{Y'} \left\| \frac{1}{\kappa_{\min}} \right\|_{L^p(\Omega)}.$$

As an alternative to the a.e. formulation (4), a global weak formulation in both physical and probability spaces can be considered. Defining the bilinear form

$$a(u,v) := \int_\Omega \int_D \kappa(x,\omega)\nabla u(x,\omega)\nabla v(x,\omega)dx d\mathbb{P}(\omega) = \mathbb{E}\left[a_\omega(u_\omega, v_\omega)\right], \tag{5}$$

the energy space $\mathscr{Y} := \left\{v : \Omega \to Y : \mathcal{F}/\mathcal{B}(Y)\text{-measurable}, \|v\|_{\mathcal{A}}^2 := a(v,v) < \infty\right\}$, and the functional $\Phi(v) := \int_\Omega \langle \phi, v_\omega\rangle d\mathbb{P}(\omega) = \mathbb{E}\left[\langle \phi, v_\omega\rangle\right]$, the global weak formulation reads

$$\text{find } y \in \mathscr{Y} \text{ s.t. } a(y,v) = \Phi(v), \quad \forall v \in \mathscr{Y}. \tag{6}$$

We further introduce the operators associated with the bilinear forms $a_\omega(\cdot,\cdot)$ and $a(\cdot,\cdot)$, namely

$$\mathcal{A}_\omega : Y \to Y' \quad \langle \mathcal{A}_\omega u, v\rangle := a_\omega(u,v), \tag{7}$$
$$\mathcal{A} : \mathscr{Y} \to \mathscr{Y}' \quad \langle \mathcal{A}u, v\rangle := a(u,v). \tag{8}$$

The link between the global weak formulation (6) and the a.e. formulation (4) is provided in the following Lemma.

Lemma 2. The solution of (4), interpreted as the representative element of the equivalence class of functions coinciding $\mathbb{P}$-a.s. with it, is the unique solution of the linear variational problem (6) and lies in $L^p(\Omega, Y)$.

Proof. Since the energy space $\mathscr{Y}$ is a Hilbert space,[42, Proposition 3.6], the existence and uniqueness of the solution of (6) follows from Riesz's theorem if $\Phi \in \mathscr{Y}'$. Due to the specific form of Φ, this is easily verified since for any $\phi \in Y'$,

$$|\Phi(v)| = \left| \int_\Omega \langle \phi, v(\cdot,\omega)\rangle d\mathbb{P}(\omega) \right| \le \|\phi\|_{Y'} \int_\Omega \|v(\cdot,w)\|_Y d\mathbb{P}(\omega) \le \|\phi\|_{Y'} \sqrt{\mathbb{E}\left[\frac{1}{\kappa_{\min}(\omega)}\right]} \|v\|_{\mathcal{A}}.$$

Further, Corollary 3.8 in Gittelson et al.[42] shows that the solution of (4) coincides $\mathbb{P}$-a.e with the unique solution of (6). Finally, using Lemma 1, we obtain the desired regularity. □

In this manuscript, we are interested in solving OCPs constrained by the state equation (6), the applications in mind being the optimal control of heat diffusion processes or inverse problems in subsurface flows. We suppose that the deterministic force term ϕ can be decomposed in a given deterministic part called f, and a deterministic control $\widetilde{u}$. We suppose that $\widetilde{u}$ lies in the dual of a Hilbert space U, which will be either $L^2(D)$ or Y. In both cases, we use the Riesz operator $\Lambda_U : U \to U' \subset Y'$, such that $\widetilde{u} = \Lambda_U u$ and $\langle \widetilde{u}, v\rangle = (u,v)_U, \forall v \in U$. The quantity we aim to compute is the Riesz representative u.
We focus on the quadratic objective functional

$$\begin{aligned} J &= \frac{1}{2}\mathbb{E}\left[\|y_\omega - y_d\|_{L^2(D)}^2\right] + \frac{\gamma}{2}\|\mathbb{S}\left[y_\omega\right]\|_{L^2(D)}^2 + \frac{\beta}{2}\|u\|_U^2 \\ &= \frac{1}{2}\left(y - y_d, y - y_d\right)_{L^2(\Omega,L^2(D))} + \frac{\gamma}{2}\left(y - \mathbb{E}\left[y_\omega\right], y - \mathbb{E}\left[y_\omega\right]\right)_{L^2(\Omega,L^2(D))} + \frac{\beta}{2}(u,u)_U, \end{aligned}$$

where $\gamma \geq 0$, $\beta > 0$, and $y_d \in L^2(D)$ is a deterministic target state. The optimization of J consists in a trade-off between how close y is to the target state y_d, and how large are the L^2 norm of the pointwise standard deviation of y and the energy of the control u. The relative importance of the latter two terms is measured by γ and β. The whole OCP can be formulated as

$$\begin{cases} \min_{u \in U} J(u) = \frac{1}{2}\mathbb{E}\left[\|y_\omega(u) - y_d\|_{L^2}^2\right] + \frac{\gamma}{2}\|\mathbb{S}\left[y_\omega(u)\right]\|_{L^2}^2 + \frac{\beta}{2}\|u\|_U^2, \\ \text{where } y_\omega(u) \in \mathscr{Y} \text{ solves} \\ \mathbb{E}\left[\langle \mathcal{A}_\omega y_\omega(u), v_\omega\rangle\right] = \mathbb{E}\left[\langle f + \Lambda_U u, v_\omega\rangle\right], \quad \forall v \in \mathscr{Y}. \end{cases} \tag{9}$$

We emphasize the dependence of y on the control u through the notation $y(u)$.

Lemma 3 (Well posedness of the OCP)**.** If Assumption 1 holds with $p \geq 2$, then the OCP (9) admits a unique solution $u^* \in U$.

Proof. The proof is standard and it is a straightforward generalization of the classical theory of Lions[1], see also the monographs[3, 52]. The case $p = 2$ is detailed in Theorem 3.4 in Martínez-Frutos et al.[53] □

To derive the optimality conditions, we rely on an optimize-then-discretize paradigm and a Lagrangian approach. For the sake of brevity, we omit the calculations of the directional derivatives evaluated in (y, u, p), where $p \in L^2(\Omega, Y)$ is the adjoint variable, along the directions δy, δp, and δu. We refer the interested reader to Van Barel et al.[12, Section 4] and Ayoul-Guilmard et al.[54] The optimality system reads

$$\begin{aligned} \mathbb{E}\left[\langle \mathcal{A}_\omega p_\omega, v_\omega\rangle\right] + \mathbb{E}\left[\langle \Lambda_{L^2}\left(y_\omega + \gamma(y_\omega - \mathbb{E}\left[y_\omega\right])\right), v_\omega\rangle\right] &= \mathbb{E}\left[\langle \Lambda_{L^2} y_d, v_\omega\rangle\right], \ \forall v \in \mathscr{Y}, \\ \langle \beta \Lambda_U u - \Lambda_U \mathbb{E}\left[p_\omega\right], v\rangle &= 0, \qquad \forall v \in U, \\ \mathbb{E}\left[\langle \mathcal{A}_\omega y_\omega, v_\omega\rangle\right] - \mathbb{E}\left[\langle \Lambda_U u, v_\omega\rangle\right] &= \mathbb{E}\left[\langle f, v_\omega\rangle\right], \qquad \forall v \in \mathscr{Y}. \end{aligned} \tag{10}$$

Notice that we tacitly used the self-adjointness of the state equation (2). The analysis of Section 5 can be naturally extended to the non-symmetric case, introducing the discretization of the adjoint operator of $\mathcal{A}_\omega$ where appropriate, see, e.g., Rees et al.[17]. More effort is needed to extend the results of Section 6, as the analysis relies on the choice of a proper norm which is problem dependent.

4 | DISCRETIZATION

4.1 | Discretization in probability

To numerically approximate the solution of (9), we rely on a Sample Average Approximation (SAA)[4]. We replace the exact expectation operator $\mathbb{E}[\cdot]$ with a suitable quadrature formula $\widehat{\mathbb{E}}[\cdot]$ with N nodes. Given a random variable $X \in L^2(\Omega)$ we approximate,

$$\begin{aligned} \mathbb{E}[X(\omega)] &= \int_\Omega X(\omega) d\mathbb{P}(\omega) \approx \sum_{i=1}^N \zeta_i X(\omega_i) =: \widehat{\mathbb{E}}[X(\omega)], \\ \mathbb{S}[X(\omega)] &= \sqrt{\mathbb{E}\left[(X(\omega) - \mathbb{E}[X(\omega)])^2\right]} \approx \sqrt{\widehat{\mathbb{E}}\left[(X(\omega) - \widehat{\mathbb{E}}[X(\omega)])^2\right]} =: \widehat{\mathbb{S}}[X(\omega)], \end{aligned}$$

where ζ_i and ω_i are, respectively, the weights and nodes of the quadrature formula with $\sum_{j=1}^N \zeta_j = 1$. We restrict ourselves to quadrature formulae with positive weights, such as Monte Carlo, Quasi-Monte Carlo and Gaussian formulae. We exclude sparse grids and Multilevel Monte Carlo approximations, since the presence of negative weights may compromize the convexity of the OCP. The construction of Gaussian quadrature formulae requires that the probability space can be parametrized by a sequence (finite or countable) of independent random variables $\{\xi_j\}_j$, each with distribution μ_j, and the existence of a complete basis of tensorized $L^2_{\mu_j}$-orthonormal polynomials. This assumption can be either the consequence of a modelling hypothesis or mathematically justified as the truncation of a Karhunen-Loève expansion of the diffusion field κ or of a transformation of it, $\widetilde{\kappa} = \psi(\kappa)$ (as, e.g., in the log-normal case with $\widetilde{\kappa} = \log(\kappa)$)[51, Section 7.4]. Concerning the quadrature error, we refer to Martin et al.[7] for Monte Carlo, to Guth et al.[11] for Quasi Monte Carlo, and to Martin et al.[7, Appendix A] for SC discretizations.
Once the probability space has been discretized, the vectors

$$y(x) = (y_{\omega_1}(x), \dots, y_{\omega_N}(x))^\top \in \underline{Y} \text{ and } p(x) = (p_{\omega_1}(x), \dots, p_{\omega_N}(x))^\top \in \underline{Y},$$

contain snapshots of the function-valued random variables $\omega \mapsto y(\cdot,\omega)$ and $\omega \mapsto p(\cdot,\omega)$ at the N collocation points. We now introduce the operator $\widehat{\mathcal{A}} : \underline{Y} \to \underline{Y}'$, which approximates the bilinear form $a(\cdot,\cdot)$ in (5),

$$\langle \widehat{\mathcal{A}}\underline{u}, \underline{v}\rangle = \sum_{i=1}^{N} \zeta_i \langle \mathcal{A}_{\omega_i} u_{\omega_i}, v_{\omega_i}\rangle = \widehat{\mathbb{E}}\left[\langle \mathcal{A}_\omega \underline{u}, \underline{v}\rangle\right], \quad \forall \underline{u} = (u_{\omega_1}, \dots, u_{\omega_N}), \underline{v} = (v_{\omega_1}, \dots, v_{\omega_N}) \in \underline{Y}, \tag{11}$$

the constant extension operator $\mathcal{I} : Y' \to \underline{Y}'$ such that $\mathcal{I} f = (f, \dots, f)^\top \; \forall f \in Y'$, and its adjoint $\mathcal{I}' : \underline{Y} \to Y$ as $\mathcal{I}'\underline{v} = \sum_{i=1} v_{\omega_i}$ $\forall \underline{v} \in \underline{Y}$, so that $\langle \mathcal{I} f, \underline{v}\rangle_{\underline{Y}',\underline{Y}} = \langle f, \mathcal{I}'\underline{v}\rangle_{Y',Y}$. Notice that $\widehat{\mathbb{E}}\left[y_\omega\right] = \sum_{i=1}^{N} \zeta_i y_{\omega_i} = \mathcal{I}'\mathcal{Z}\underline{y}$, where $\mathcal{Z} = \mathrm{diag}(\zeta_1, \dots, \zeta_N)$ is a diagonal matrix containing the quadrature weights. Finally, the operator $\underline{\Lambda}_{L^2}$ is defined as $\underline{\Lambda}_{L^2}\underline{v} = (\Lambda_{L^2} v_{\omega_1}, \dots, \Lambda_{L^2} v_{\omega_N})^\top$. The semi-discrete matrix formulation of (10)[1], written as an equality in dual spaces, is

$$\begin{pmatrix} \underline{\Lambda}_{L^2}\left((1+\gamma)\mathcal{Z} - \gamma\mathcal{Z}\mathcal{I}\mathcal{I}'\mathcal{Z}\right) & 0 & \widehat{\mathcal{A}} \\ 0 & \beta\Lambda_U & -\Lambda_U \mathcal{I}'\mathcal{Z} \\ \widehat{\mathcal{A}} & -\mathcal{Z}\mathcal{I}\Lambda_U & 0 \end{pmatrix} \begin{pmatrix} \underline{y} \\ \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \mathcal{Z}\mathcal{I}\Lambda_{L^2} y_d \\ 0 \\ \mathcal{Z}\mathcal{I} f \end{pmatrix}. \tag{12}$$

which corresponds to the set of equations

$$\begin{aligned} \mathcal{A}_{\omega_i} p_{\omega_i} + (1+\gamma)\Lambda_{L^2} y_{\omega_i} - \gamma\widehat{\mathbb{E}}\left[y_\omega\right] &= \Lambda_{L^2} y_d, && i = 1, \dots, N, \\ \beta\Lambda_U u - \Lambda_U \widehat{\mathbb{E}}\left[p_\omega\right] &= 0, && \\ \mathcal{A}_{\omega_i} y_{\omega_i} - \Lambda_U u &= f, && i = 1, \dots, N. \end{aligned} \tag{13}$$

4.2 | Discretization in space

Let us denote by $\left\{\mathcal{T}_h\right\}_{h>0}$ a family of regular triangulations of D. Y^h denotes the space of continuous piecewise polynomial functions of degree r over $\mathcal{T}_h$ that vanish on ∂D, that is $Y^h := \left\{ v_h \in C^0(\overline{D}) : v_h|_K \in \mathbb{P}_r(K), \quad \forall K \in \mathcal{T}_h,\ y|_{\partial D} = 0 \right\} \subset Y$. N_h is the number of degrees of freedom associated with the space Y^h. We consider a finite element discretization of system (12). The vectors $\mathbf{y} = (\mathbf{y}_1, \dots, \mathbf{y}_N) \in \mathbb{R}^{N\cdot N_h}$ and $\mathbf{p} = (\mathbf{p}_1, \dots, \mathbf{p}_N) \in \mathbb{R}^{N\cdot N_h}$ are the discretization of the vector functions $\underline{y}$ and $\underline{p}$. To discretize the control u, we use the same finite element space Y^h [7, Remark 3.1]. Further, the matrices $A_{\omega_i} \in \mathbb{R}^{N_h\times N_h}$ are the stiffness matrices corresponding to the elliptic operators $\mathcal{A}_{\omega_i}$, and $A_0 := \sum_{i=1}^{N} \zeta_i A_{\omega_i}$ is the empirical mean. $M_s \in \mathbb{R}^{N_h\times N_h}$ is the standard mass matrix. The identity matrices are $I_s \in \mathbb{R}^{N_h\times N_h}$ and $I \in \mathbb{R}^{N\cdot N_h\times N\cdot N_h}$. According to the choice of the control space, that is $U = L^2(D)$ or $U = Y'$, the representation of the Riesz operator Λ_U is either $\Lambda_U = M_s$ or $\Lambda_U = K$, where K is the stiffness matrix associated with the standard scalar product in Y. In the following, we will suppose the control u lies in $L^2(D)$, i.e. $U = L^2(D)$.

At the fully discrete level, system (12) reads $\mathscr{S}\mathbf{x} = \mathbf{b}$,

$$\mathscr{S} = \begin{pmatrix} M\left((1+\gamma)Z - \gamma Z\mathbb{1}\mathbb{1}^\top Z\right) & 0 & A \\ 0 & \beta M_s & -M_s\mathbb{1}^\top Z \\ A & -Z\mathbb{1}M_s & 0 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} Z\mathbb{1}M_s\mathbf{y}_d \\ 0 \\ Z\mathbb{1}\mathbf{f} \end{pmatrix}, \tag{14}$$

where $A \in \mathbb{R}^{N\cdot N_h\times N\cdot N_h}$, $M \in \mathbb{R}^{N\cdot N_h\times N\cdot N_h}$, $\mathbb{1} \in \mathbb{R}^{N\cdot N_h\times N_h}$ are defined as

$$A := \begin{pmatrix} \zeta_1 A_{\omega_1} & & & \\ & \zeta_2 A_{\omega_2} & & \\ & & \ddots & \\ & & & \zeta_N A_{\omega_N} \end{pmatrix}, \quad M := \begin{pmatrix} M_s & & & \\ & M_s & & \\ & & \ddots & \\ & & & M_s \end{pmatrix}, \quad \mathbb{1} := \begin{pmatrix} I_s \\ I_s \\ \vdots \\ I_s \end{pmatrix},$$

and $Z = \mathrm{diag}(\zeta_1 I_s, \dots, \zeta_N I_s) \in \mathbb{R}^{N\cdot N_h, N\cdot N_h}$ is the discretization of $\mathcal{Z}$. Since M has constant diagonal blocks the following equalities hold true and will be extensively used,

$$M\mathbb{1}\mathbb{1}^\top = \mathbb{1}M_s\mathbb{1}^\top = \mathbb{1}\mathbb{1}^\top M \quad \text{and} \quad MZ = ZM. \tag{15}$$

[1] The same semi-discrete optimality system can be derived using a discrete-then-optimize paradigm in probability, that is by replacing $\mathbb{E}[\cdot]$ with $\widehat{\mathbb{E}}[\cdot]$ into (9), and then by calculating the directional derivatives.

It is convenient to rewrite (14) in a compact form as

$$\mathscr{S} = \begin{pmatrix} C & B^\top \\ B & 0 \end{pmatrix}, \tag{16}$$

where

$$C := \begin{pmatrix} M_\gamma & 0 \\ 0 & \beta M_s \end{pmatrix} \quad \text{with } M_\gamma := M\left((1+\gamma)Z - \gamma Z \mathbb{1}\mathbb{1}^\top Z\right), \quad \text{and } B := \begin{pmatrix} A & -Z\mathbb{1}M_s \end{pmatrix}.$$

The matrix M_γ plays a key role in the following, thus we discuss its properties in the next Lemma.

Lemma 4. The matrix $M_\gamma = M\left((1+\gamma)Z - \gamma Z \mathbb{1}\mathbb{1}^\top Z\right)$ is symmetric and positive definite for any $\gamma \geq 0$. Its inverse is equal to

$$M_\gamma^{-1} = \left(M\left((1+\gamma)Z - \gamma Z \mathbb{1}\mathbb{1}^\top Z\right)\right)^{-1} = \left(\frac{1}{1+\gamma} Z^{-1} + \frac{\gamma}{1+\gamma}\mathbb{1}\mathbb{1}^\top\right) M^{-1}.$$

Proof. A straightforward calculation shows that

$$\begin{aligned}(I - \mathbb{1}\mathbb{1}^\top Z)^\top M Z (I - \mathbb{1}\mathbb{1}^\top Z) &= MZ - MZ\mathbb{1}\mathbb{1}^\top Z + Z\mathbb{1}\mathbb{1}^\top M Z \mathbb{1}\mathbb{1}^\top Z - Z\mathbb{1}\mathbb{1}^\top MZ \\ &= MZ - ZM\mathbb{1}\mathbb{1}^\top Z + Z\mathbb{1}M_s\mathbb{1}^\top Z - Z\mathbb{1}\mathbb{1}^\top MZ = MZ - Z\mathbb{1}M_s\mathbb{1}^\top Z,\end{aligned}$$

where we used (15) and $\mathbb{1}^\top M Z \mathbb{1} = \sum_{j=1}^N \zeta_j M_s = M_s$. Hence, M_γ is symmetric since it can be written as

$$M_\gamma = MZ + \gamma MZ\left(I - \mathbb{1}\mathbb{1}^\top Z\right) = MZ + \gamma (I - \mathbb{1}\mathbb{1}^\top Z)^\top M Z (I - \mathbb{1}\mathbb{1}^\top Z).$$

The positive definiteness follows from the positiveness of the weights of the quadrature formulae. Finally using the Woodbury identity the claim follows,

$$\begin{aligned}\left((1+\gamma)Z - \gamma Z\mathbb{1}\mathbb{1}^\top Z\right)^{-1} &= \frac{1}{1+\gamma}Z^{-1} - \left(\frac{1}{1+\gamma}\right)^2 \mathbb{1}\left(-\frac{1}{\gamma}I_s + \frac{1}{1+\gamma}\mathbb{1}^\top Z \mathbb{1}\right)^{-1}\mathbb{1}^\top = \\ &= \frac{1}{1+\gamma}Z^{-1} - \left(\frac{1}{1+\gamma}\right)^2 \mathbb{1}\left(-\frac{1}{\gamma} + \frac{1}{1+\gamma}\right)^{-1}\mathbb{1}^\top = \frac{1}{1+\gamma}Z^{-1} + \frac{\gamma}{1+\gamma}\mathbb{1}\mathbb{1}^\top.\end{aligned}$$

□

5 | ALGEBRAIC PRECONDITIONERS

In this section we study algebraic preconditioners for the saddle point matrix (16) based on the seminal work by Murphy et al.[35], where the authors showed that a general saddle point matrix $\begin{pmatrix} C & B_2^\top \\ B_1 & 0 \end{pmatrix}$ can be optimally preconditioned by $\mathcal{P} := \mathrm{diag}(C, S)$, where $S = B_1 C^{-1} B_2^\top$ is the Schur complement. However, since inverting S is too expensive, one needs to use suitable approximations. Let us consider the preconditioner $\widetilde{\mathcal{P}} := \mathrm{diag}(C, \widetilde{S})$, obtained replacing the exact Schur complement S with a symmetric positive definite approximation $\widetilde{S}$. A characterization of the spectrum of $\widetilde{\mathcal{P}}^{-1}\mathscr{S}$ is provided by the following lemma.

Lemma 5 (Spectrum of $\widetilde{\mathcal{P}}^{-1}\mathscr{S}$)**.** The matrix $\widetilde{\mathcal{P}}^{-1}\mathscr{S}$ has eigenvalue 1 with multiplicity N_h. The remaining $2N \cdot N_h$ eigenvalues are distinct and equal to

$$\lambda_j = \frac{1 + \sqrt{1 + 4\sigma_j}}{2}, \quad \lambda_{j+N\cdot N_h} = \frac{1 - \sqrt{1 + 4\sigma_j}}{2},$$

where σ_j are the eigenvalues of $\widetilde{S}^{-1}S$, and $j = 1, \dots, N \cdot N_h$.

Proof. See, for instance, Section 6.2.1 of Elman et al.[55] □

Lemma 5 reduces the problem of estimating $\sigma(\widetilde{\mathcal{P}}^{-1}\mathscr{S})$ to the problem of estimating $\sigma(\widetilde{S}^{-1}S)$.
For the simpler deterministic OCP, the exact Schur complement is $S = A_s M_s^{-1} A_s + \frac{1}{\beta} M_s$, where A_s is the stiffness matrix, M_s the mass matrix and β is the control regularization parameter. Rees et al.[16,17] approximate S with $A_s M_s^{-1} A_s$, and obtain eigenvalue estimates for the preconditioned system which clearly show a β dependence. An identical approximation in the context of robust OCPUU has been first proposed in Kouri et al.[15], though without a theoretical analysis. In subsection 5.1, we provide a full characterization of the spectrum of the preconditioned system for the robust OCPUU which highlights both the dependence on β and on the random field extremal values.

As the performance of the preconditioner analyzed in Subsection 5.1 deteriorates as $\beta \to 0$, in subsection 5.2, we propose a β-robust preconditioner based on a more involved approximation of the Schur complement of (16), inspired by works on deterministic OCPs [16–18, 36–38].

5.1 | A Schur complement approximation

The exact Schur complement of the saddle point matrix $\mathscr{S}$ in (16) is

$$S := BC^{-1}B^\top = AM_\gamma^{-1}A + \frac{1}{\beta} Z\mathbb{1}M_s\mathbb{1}^\top Z. \tag{17}$$

The term $AM_\gamma^{-1}A$ is block diagonal if and only if $\gamma = 0$, in which case it is the direct generalization of the matrix which appears in a deterministic OCP, except that the diagonal blocks are multiplied by the weights of quadrature formula. On the other hand, the term $\frac{1}{\beta} Z\mathbb{1}M_s\mathbb{1}^\top Z$ is difficult to handle as it has significantly different properties from the corresponding $\frac{1}{\beta}M_s$ term of deterministic OCPs. First, $\frac{1}{\beta} Z\mathbb{1}M_s\mathbb{1}^\top Z$ is a block-dense matrix, where each block is given by a mass matrix. Second it is a relatively low-rank term. Its effect is to couple all the equations, increasing the difficulties to construct β robust preconditioners. We remark that a similar low-rank perturbation appears in deterministic OCP with a control acting on a subset of the boundary [33]. The first approximation $\widetilde{S}$, and corresponding preconditioner $\widetilde{P}$, we consider is obtained dropping the β-dependent low-rank term,

$$\widetilde{S} := AM_\gamma^{-1}A \approx S, \quad \widetilde{P} := \begin{pmatrix} C & 0 \\ 0 & \widetilde{S} \end{pmatrix}. \tag{18}$$

Computing $\widetilde{S}^{-1}S$ we obtain

$$\widetilde{S}^{-1}S = (AM_\gamma^{-1}A)^{-1}\left(AM_\gamma^{-1}A + \frac{1}{\beta} Z\mathbb{1}M_s\mathbb{1}^\top Z\right) = I + \frac{1}{\beta} A^{-1}M_\gamma A^{-1} Z\mathbb{1}M_s\mathbb{1}^\top Z =: I + \frac{1}{\beta}\widetilde{H},$$

that is, $\widetilde{S}^{-1}S$ is the identity plus a β-dependent low-rank term. Hence, $\widetilde{S}^{-1}S$ will have at most N_h eigenvalues different from one since rank $\left(\mathbb{1}\mathbb{1}^\top Z\right) = N_h$. To study the spectrum of $\widetilde{H}$, we consider the similar matrix

$$H := Z\widetilde{H}Z^{-1} = ZA^{-1}M_\gamma A^{-1}ZM\mathbb{1}\mathbb{1}^\top = (1+\gamma)ZA^{-1}MZA^{-1}ZM\mathbb{1}\mathbb{1}^\top - \gamma ZA^{-1}MZ\mathbb{1}\mathbb{1}^\top ZA^{-1}ZM\mathbb{1}\mathbb{1}^\top. \tag{19}$$

A characterization of the spectrum of $\widetilde{S}^{-1}S$ is provided in the following Lemma.

Lemma 6 (Spectrum of $\widetilde{S}^{-1}S$)**.** The spectrum of $\widetilde{S}^{-1}S$ satisfies $\sigma(\widetilde{S}^{-1}S) = \{1\} \cup \left\{1 + \frac{1}{\beta}\mu_j\right\}_{j=1}^{N_h}$, where μ_j are the eigenvalues of $\widehat{\mathbb{E}}\left[K_\omega^2\right] + \gamma\left(\widehat{\mathbb{E}}\left[K_\omega^2\right] - \widehat{\mathbb{E}}\left[K_\omega\right]^2\right)$, with $K_\omega := A_\omega^{-1}M_s$.

Proof. As $ZA^{-1} = \mathrm{diag}(A_{\omega_1}^{-1}, A_{\omega_2}^{-1}, \dots, A_{\omega_N}^{-1})$, direct calculations show that

$$ZA^{-1}MZA^{-1}ZM\mathbb{1}\mathbb{1}^\top = \begin{pmatrix} \zeta_1 K_{\omega_1}^2 & \zeta_1 K_{\omega_1}^2 & \cdots & \zeta_1 K_{\omega_1}^2 \\ \zeta_2 K_{\omega_2}^2 & \zeta_2 K_{\omega_2}^2 & \cdots & \zeta_2 K_{\omega_2}^2 \\ \vdots & \vdots & \vdots & \vdots \\ \zeta_N K_{\omega_N}^2 & \zeta_N K_{\omega_N}^2 & \cdots & \zeta_N K_{\omega_N}^2 \end{pmatrix},$$

and

$$ZA^{-1}MZ\mathbb{1}\mathbb{1}^\top ZA^{-1}ZM\mathbb{1}\mathbb{1}^\top = \begin{pmatrix} \zeta_1 K_{\omega_1}\left(\sum_{i=1}^N \zeta_i K_{\omega_i}\right) & \cdots & \zeta_1 K_{\omega_1}\left(\sum_{i=1}^N \zeta_i K_{\omega_i}\right) \\ \zeta_2 K_{\omega_2}\left(\sum_{i=1}^N \zeta_i K_{\omega_i}\right) & \cdots & \zeta_2 K_{\omega_2}\left(\sum_{i=1}^N \zeta_i K_{\omega_i}\right) \\ \vdots & \vdots & \vdots \\ \zeta_N K_{\omega_N}\left(\sum_{i=1}^N \zeta_i K_{\omega_i}\right) & \cdots & \zeta_N K_{\omega_N}\left(\sum_{i=1}^N \zeta_i K_{\omega_i}\right) \end{pmatrix}$$

The matrix H ((19)) is then equal to

$$H = \begin{pmatrix} H_1 & \dots & H_1 \\ H_2 & \dots & H_2 \\ \vdots & \vdots & \vdots \\ H_N & \dots & H_N \end{pmatrix} = \widehat{H}\mathbb{1}\mathbb{1}^\top,$$

where $H_i := \zeta_i K_{\omega_i}^2 + \gamma\left(\zeta_i K_{\omega_i}^2 - \zeta_i K_{\omega_i}\left(\sum_{j=1}^N \zeta_j K_{\omega_j}\right)\right)$, and $\widehat{H} := \mathrm{diag}(H_1, H_2, \dots, H_N)$. Clearly, the rank of H is N_h, being N_h the size of K_{ω_i}, i.e. the number of degrees of freedom in the finite element discretization.

We then look for an eigenpair $(\lambda, \underline{\mathbf{v}})$, where $\underline{\mathbf{v}} = (\mathbf{v}_1, \dots, \mathbf{v}_N) \in \mathbb{R}^{N \cdot N_h}$. Notice that if $(\lambda, \underline{\mathbf{v}})$ is an eigenpair of H, then $(\lambda, \mathbf{w})$ with $\mathbf{w} := \sum_{j=1}^{N} \mathbf{v}_j = \mathbb{1}^\top \mathbf{v}$ is an eigenpair of the reduced matrix $\mathbb{1}^\top \widehat{H} \mathbb{1} = \sum_{i=1}^{N} H_i$, since

$$\mathbb{1}^\top \widehat{H} \mathbb{1}\mathbf{w} = \mathbb{1}^\top \widehat{H} \mathbb{1}\mathbb{1}^\top \mathbf{v} = \lambda \mathbb{1}^\top \mathbf{v} = \lambda \mathbf{w}.$$

Thus, we can first compute the eigenpair $(\lambda, \mathbf{w})$ of $\sum_{i=1}^{N} H_i$, and then recover the eigenpair $(\lambda, \mathbf{v})$ of H setting $\mathbf{v} = \frac{1}{\lambda} \widehat{H} \mathbb{1}\mathbf{w}$, since

$$H\mathbf{v} = \widehat{H} \mathbb{1}\mathbb{1}^\top \mathbf{v} = \frac{1}{\lambda} \widehat{H} \mathbb{1}\mathbb{1}^\top \widehat{H} \mathbb{1}\mathbf{w} = \widehat{H} \mathbb{1}\mathbf{w} = \lambda \mathbf{v}.$$

Calculating explicitly $(\sum_{i=1}^{N} H_i)$ we obtain,

$$\sum_{i=1}^{N} H_i = \sum_{i=1}^{N} \zeta_i K_{\omega_i}^2 + \gamma \left[\sum_{i=1}^{N} \zeta_i K_{\omega_i}^2 - \left(\sum_{j=1}^{N} \zeta_j K_{\omega_j} \right)^2 \right] = \widehat{\mathbb{E}}\left[K_\omega^2\right] + \gamma \left(\widehat{\mathbb{E}}\left[K_\omega^2\right] - \widehat{\mathbb{E}}\left[K_\omega\right]^2 \right),$$

and the claim follows. □

Lemmas 5 and 6 guarantee that the spectrum of $\widetilde{\mathcal{P}}^{-1}\mathcal{S}$ is well clustered around 1 and $\frac{1\pm\sqrt{5}}{2}$, except for $2N_h$ eigenvalues which depend on β and on the spectrum of $\widehat{\mathbb{E}}\left[K_\omega^2\right] + \gamma \left(\widehat{\mathbb{E}}\left[K_\omega^2\right] - \widehat{\mathbb{E}}\left[K_\omega\right]^2 \right)$. We formalize this statement in the following Theorem.

Theorem 1. The matrix $\widetilde{\mathcal{P}}^{-1}\mathcal{S}$ has eigenvalue $\lambda = 1$ with multiplicity N_h, $\lambda = \frac{1+\sqrt{5}}{2}$ and $\lambda = \frac{1-\sqrt{5}}{2}$ each with multiplicity $(N-1) \cdot N_h$. The remaining $2N_h$ eigenvalues are equal to

$$\lambda_j = \frac{1}{2}\left(1 + \sqrt{5 + \frac{4\mu_j}{\beta}}\right), \quad \lambda_{j+N_h} = \frac{1}{2}\left(1 - \sqrt{5 + \frac{4\mu_j}{\beta}}\right),$$

where μ_j, $j = 1, \dots, N_h$ are the eigenvalues of $\widehat{\mathbb{E}}\left[K_\omega^2\right] + \gamma \left(\widehat{\mathbb{E}}\left[K_\omega^2\right] - \widehat{\mathbb{E}}\left[K_\omega\right]^2 \right)$, with $K_\omega = A_\omega^{-1} M_s$.

Proof. The claim follows directly using the characterization of the spectrum of $\widetilde{S}^{-1}S$ (provided by Lemma 6) in Lemma 5. □

We now study how the eigenvalues of $\widehat{\mathbb{E}}\left[K_\omega^2\right] + \gamma \left(\widehat{\mathbb{E}}\left[K_\omega^2\right] - \widehat{\mathbb{E}}\left[K_\omega\right]^2 \right)$ depend on the mesh size h and on the statistical properties of the random field κ under the assumption of a quasi-uniform triangulation $\mathcal{T}_h$. To do so, we briefly recall some useful results, namely, an inverse inequality [56, Proposition 6.3.2], an equivalence of norms [57, Lemma 9.7], and a characterization of the spectra of the mass matrix [57, Theorem 9.8] and stiffness matrix [55, Theorem 1.32] concerning the finite element space Y_h and the associated Lagrangian basis:

$$\exists C_I > 0 \quad \|\nabla v_h\|_{L^2(D)} \le C_I h^{-1} \|v_h\|_{L^2(D)}, \quad \forall v_h \in Y_h, \tag{20}$$

$$\exists C_1, C_2 > 0 \quad C_1 h^d |\mathbf{v}_h|^2 \le \|v_h\|_{L^2(D)}^2 \le C_2 h^d |\mathbf{v}_h|^2, \quad \forall v_h \in Y_h, \tag{21}$$

$$\sigma(M_s) \subset [C_1 h^d, C_2 h^d] \tag{22}$$

$$\sigma(A_\omega) \subset \left[\kappa_{\min}(\omega) \frac{C_1 h^d}{C_P^2}, \kappa_{\max}(\omega) C_I^2 C_2 h^{d-2} \right], \tag{23}$$

where $\mathbf{v}_h$ is the vector collecting the nodal degrees of freedom of v_h, $|\cdot|$ is the vector euclidean norm and d is the spatial dimension. The constants C_I, C_1, C_2 may depend on the polynomial order r. We first derive an auxiliary result concerning the matrices $L_\omega := A_\omega^{-1} M_s A_\omega^{-1}$ and $\widehat{\mathbb{E}}\left[L_\omega\right]$:

Lemma 7. Defining $c_L := \frac{C_1}{C_I^4 C_2^2}$ and $C_L := \frac{C_2 C_P^4}{C_1^2}$, the following inclusions hold:

$$\begin{aligned} \sigma\left(L_\omega\right) &\subset \left[\frac{c_L h^{4-d}}{\kappa_{\max}^2(\omega)}, \frac{h^{-d} C_L}{\kappa_{\min}^2(\omega)} \right], && \text{for a.e. } \omega \in \Omega, \\ \sigma\left(\widehat{\mathbb{E}}\left[L_\omega\right]\right) &\subset \left[c_L h^{4-d} \widehat{\mathbb{E}}\left[\frac{1}{\kappa_{\max}^2(\omega)}\right], C_L h^{-d} \widehat{\mathbb{E}}\left[\frac{1}{\kappa_{\min}^2(\omega)}\right] \right] \end{aligned}$$

Proof. The matrix L_ω is symmetric and positive definite. Its extremal eigenvalues are characterized by the Raleigh quotients,

$$\begin{aligned} \lambda_{\max}(L_\omega) &= \sup_{\mathbf{v} \in \mathbb{R}^{N_h}} \frac{\mathbf{v}^\top A_\omega^{-1} M_s A_\omega^{-1} \mathbf{v}}{\mathbf{v}^\top \mathbf{v}} = \sup_{\mathbf{w} \in \mathbb{R}^{N_h}} \frac{\mathbf{w}^\top M_s \mathbf{w}}{\mathbf{w}^\top A_\omega^2 \mathbf{w}} \le \frac{C_2 h^d C_P^4}{\kappa_{\min}^2(\omega) C_1^2 h^{2d}} = \frac{h^{-d}}{\kappa_{\min}^2(\omega)} \frac{C_2 C_P^4}{C_1^2}, \\ \lambda_{\min}(L_\omega) &= \inf_{\mathbf{v} \in \mathbb{R}^{N_h}} \frac{\mathbf{v}^\top A_\omega^{-1} M_s A_\omega^{-1} \mathbf{v}}{\mathbf{v}^\top \mathbf{v}} = \inf_{\mathbf{w} \in \mathbb{R}^{N_h}} \frac{\mathbf{w}^\top M_s \mathbf{w}}{\mathbf{w}^\top A_\omega^2 \mathbf{w}} \ge \frac{C_1 h^d}{\kappa_{\max}^2(\omega) C_I^4 C_2^2 h^{2d-4}} = \frac{h^{4-d}}{\kappa_{\max}^2(\omega)} \frac{C_1}{C_I^4 C_2^2}. \end{aligned} \tag{24}$$

The matrix $\widehat{\mathbb{E}}\left[L_\omega\right]$ is positive definite as well, being the convex combination of positive definite matrices. Using (24), its extremal eigenvalues are bounded by

$$\lambda_{\max}(\widehat{\mathbb{E}}\left[L_\omega\right]) = \sup_{\mathbf{v}\in\mathbb{R}^{N_h}} \frac{\sum_{i=1}^N \zeta_i \mathbf{v}^\top L_{\omega_i}\mathbf{v}}{\mathbf{v}^\top\mathbf{v}} \leq \sum_{i=1}^N \zeta_j \sup_{\mathbf{v}\in\mathbb{R}^{N_h}} \frac{\mathbf{v}^\top L_{\omega_i}\mathbf{v}}{\mathbf{v}^\top\mathbf{v}} \leq h^{-d} C_L \widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\min}(\omega)}\right],$$
$$\lambda_{\min}(\widehat{\mathbb{E}}\left[L_\omega\right]) = \inf_{\mathbf{v}\in\mathbb{R}^{N_h}} \frac{\sum_{i=1}^N \zeta_i \mathbf{v}^\top L_{\omega_i}\mathbf{v}}{\mathbf{v}^\top\mathbf{v}} \geq \sum_{i=1}^N \zeta_j \inf_{\mathbf{v}\in\mathbb{R}^{N_h}} \frac{\mathbf{v}^\top L_{\omega_i}\mathbf{v}}{\mathbf{v}^\top\mathbf{v}} \geq h^{4-d} c_L \widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\max}(\omega)}\right].$$

□

Thanks to Lemma 7, we can give bounds on the spectrum of the matrix $\widehat{\mathbb{E}}\left[K^2_\omega\right] + \gamma\left(\widehat{\mathbb{E}}\left[K^2_\omega\right] - \widehat{\mathbb{E}}\left[K_\omega\right]^2\right)$, thus leading to the final result.

Theorem 2 (Characterization of the spectrum of $\widetilde{\mathcal{P}}^{-1}\mathcal{S}$)**.** If the triangulation $\mathcal{T}_h$ is quasi-uniform, then the eigenvalues of $\widetilde{\mathcal{P}}^{-1}\mathcal{S}$ satisfy one of

$$\lambda \in \left\{1, \frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}\right\},$$
$$\frac{1}{2}\left(1+\sqrt{5+\frac{4d_L h^4}{\beta}\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\max}(\omega)}\right]}\right) \leq \lambda \leq \frac{1}{2}\left(1+\sqrt{5+\frac{4D_L(1+\gamma)}{\beta}\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\min}(\omega)}\right]}\right), \tag{25}$$
$$\text{or } \frac{1}{2}\left(1-\sqrt{5+\frac{4D_L(1+\gamma)}{\beta}\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\min}(\omega)}\right]}\right) \leq \lambda \leq \frac{1}{2}\left(1-\sqrt{5+\frac{4d_L h^4}{\beta}\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\max}(\omega)}\right]}\right),$$

where $d_L = c_L C_1$, $D_L = C_L C_2$ are constants independent of N, β, γ, h and on the random set of realizations $\{\omega_i\}_{i=1}^N$.

Proof. Due to Theorem 1, we study the extremal eigenvalues of $\widehat{\mathbb{E}}\left[K^2_\omega\right] + \gamma\left(\widehat{\mathbb{E}}\left[K^2_\omega\right] - \widehat{\mathbb{E}}\left[K_\omega\right]^2\right)$. We first suppose $\gamma = 0$ and remark that $\widehat{\mathbb{E}}\left[K^2_\omega\right] = \widehat{\mathbb{E}}\left[L_\omega M_s\right] = \widehat{\mathbb{E}}\left[L_\omega\right] M_s$, which is similar to $M_s^{-1/2}\widehat{\mathbb{E}}\left[L_\omega\right] M_s^{-1/2}$. Using Lemma 7, we have

$$\lambda_{\max}(\widehat{\mathbb{E}}\left[K^2_\omega\right]) = \sup_{\mathbf{v}\in\mathbb{R}^{N_h}} \frac{\mathbf{v}^\top\widehat{\mathbb{E}}\left[L_\omega\right]\mathbf{v}}{\mathbf{v}^\top M_s^{-1}\mathbf{v}} \leq C_L C_2 \widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\min}(\omega)}\right].$$
$$\lambda_{\min}(\widehat{\mathbb{E}}\left[K^2_\omega\right]) = \inf_{\mathbf{v}\in\mathbb{R}^{N_h}} \frac{\mathbf{v}^\top\widehat{\mathbb{E}}\left[L_\omega\right]\mathbf{v}}{\mathbf{v}^\top M_s^{-1}\mathbf{v}} \geq c_L C_1 h^4 \widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\max}(\omega)}\right].$$

Next, we consider the γ dependent term and observe that

$$\widehat{\mathbb{E}}\left[K^2_\omega\right] - \widehat{\mathbb{E}}\left[K_\omega\right]^2 = \widehat{\mathbb{E}}\left[A_\omega^{-1} M_s A_\omega^{-1} M_s\right] - \widehat{\mathbb{E}}\left[A_\omega^{-1} M_s\right]\widehat{\mathbb{E}}\left[A_\omega^{-1} M_s\right] = \left(\widehat{\mathbb{E}}\left[A_\omega^{-1} M_s A_\omega^{-1}\right] - \widehat{\mathbb{E}}\left[A_\omega^{-1}\right] M_s \widehat{\mathbb{E}}\left[A_\omega^{-1}\right]\right) M_s$$
$$= \widehat{\mathbb{E}}\left[\left(A_\omega^{-1} M_s^{\frac{1}{2}} - \widehat{\mathbb{E}}\left[A_\omega^{-1} M_s^{\frac{1}{2}}\right]\right)\left(A_\omega^{-1} M_s^{\frac{1}{2}} - \widehat{\mathbb{E}}\left[A_\omega^{-1} M_s^{\frac{1}{2}}\right]\right)^\top\right] M_s.$$

Thus, $\widehat{\mathbb{E}}\left[K^2_\omega\right] - \widehat{\mathbb{E}}\left[K_\omega\right]^2$ can be written as the product between an expectation of a semi positive definite matrix and M_s, hence its eigenvalues are real and non-negative. Sharp estimates of the eigenvalues of $\widehat{\mathbb{E}}\left[K^2_\omega\right] - \widehat{\mathbb{E}}\left[K_\omega\right]^2$ rely on bounds of the spectrum of $A_\omega^{-1} - \widehat{\mathbb{E}}\left[A_\omega^{-1}\right]$, which however are not available in terms of $\kappa_{\min}(\omega)$ and $\kappa_{\max}(\omega)$. To obtain an upper bound, we rely on the following estimates,

$$\lambda_{\max}\left(\widehat{\mathbb{E}}\left[K^2_\omega\right] + \gamma\widehat{\mathbb{E}}\left[K^2_\omega\right] - \gamma\widehat{\mathbb{E}}\left[K_\omega\right]^2\right) \leq C_L C_2(1+\gamma)\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\min}(\omega)}\right]. \tag{26}$$

To obtain a lower bound, we simply ignore the γ-dependent term,

$$\lambda_{\min}\left(\widehat{\mathbb{E}}\left[K^2_\omega\right] + \gamma\left(\widehat{\mathbb{E}}\left[K^2_\omega\right] - \widehat{\mathbb{E}}\left[K_\omega\right]^2\right)\right) \geq \lambda_{\min}\left(\widehat{\mathbb{E}}\left[K^2_\omega\right]\right) = c_L C_1 h^4 \widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\max}(\omega)}\right]. \tag{27}$$

Combining (26) and (27) with Theorem 1 concludes the proof. □

5.1.1 | Comments on the influence of the random field and computational considerations

A few comments are in order about Theorem 2. First, $\widetilde{P}$ will not lead to β-robust convergence, as the spectrum clearly spreads as $\beta \to 0$. Second, for a given κ satisfying Assumption 1 with $p \geq 2$, $\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\min}(\omega)}\right]$ and $\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\max}(\omega)}\right]$ converge to the finite continuous

expectations $\mathbb{E}\left[\frac{1}{\kappa^2_{\min}(\omega)}\right]$ and $\mathbb{E}\left[\frac{1}{\kappa^2_{\max}(\omega)}\right]$ as we increase the number N of collocation points. Hence, the preconditioner is robust with respect to the level of discretization of the probability space. Third, $\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\min}(\omega)}\right]$ and $\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\max}(\omega)}\right]$ represent the dependence of $\sigma(\widetilde{\mathcal{P}}^{-1}\mathcal{S})$ on the random field. Thus, the spectrum spreads when considering random fields κ with smaller values of $\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\min}(\omega)}\right]$. Estimates on the moments of $\kappa_{\max}(\omega)$ and $\frac{1}{\kappa_{\min}(\omega)}$ are available for the log-normal random field of the form[45, 49],

$$\kappa_L(x,\omega) = e^{g(x,\omega)} = e^{\sigma\sum_{j=1}^{\infty}\sqrt{\lambda_j}b_j(x)N_j(\omega)}, \tag{28}$$

where $g(x,\omega)$ is a mean zero Gaussian field with covariance function $\mathrm{Cov}_g(x,y)$, $(b_j(x),\sigma^2\lambda_j)$ are the eigenpairs of $\mathcal{T}: L^2(D)\to L^2(D)$, $(\mathcal{T}f)(x) := \int_D \mathrm{Cov}_g(x,y)f(y)dy$, and $N_j(\omega)\sim\mathcal{N}(0,1)$. Assuming that $b_j(x)$ are Hölder continuous with exponent $0<\alpha\le 1$ $\forall j\ge 1$, and that $R_\alpha := \sum_{j=1}^{N}\lambda_j\|b_j\|_{C^{0,\alpha}(\overline{D})}<\infty$, it holds

$$\|g\|_{L^p(\Omega,C^{0,\alpha}(\overline{D}))} \le \widetilde{C}^{\frac{1}{p}}\sqrt{R_\alpha}\sigma((p-1)!!)^{\frac{1}{p}},$$

for every even p, where $(p-1)!!$ is the bi-factorial and $\widetilde{C}$ is indipendent on σ and p. Further, using Fernique's Theorem, one can show[45, Proposition 3.10] that

$$\left\|\frac{1}{\kappa_{\min}}\right\|_{L^p(\Omega)} = \left(\mathbb{E}\left[\frac{1}{\kappa^p_{\min}(\omega)}\right]\right)^{\frac{1}{p}} \le De^{Cp\sigma^2} =: B_p, \quad \text{and} \quad \|\kappa_{\max}\|_{L^p(\Omega)} = \left(\mathbb{E}\left[\kappa^p_{\max}(\omega)\right]\right)^{\frac{1}{p}} \le B_p,$$

where D and C are constants independent of p and σ. The exponential dependence over p and σ^2 is not dramatic, as p can be chosen equal to 2, and in physical applications σ^2 is usually small: for instance setting $\sigma^2=1.5$, one can already model random fields which vary up to four orders of magnitude inside the domain, see Section 7.2.

To better understand the behaviour of the N_h, β-dependent, eigenvalues of $\widetilde{S}^{-1}S$, we consider two different random models and corresponding OCPs (9). The first one is a log-normal random diffusion field with $\mathrm{Cov}_g(x,y)=\sigma^2 e^{\frac{-\|x-y\|_2^2}{L^2}}$, where L is the correlation length. With $L^2=0.5$, retaining the first $M=3$ terms in (28) is enough to preserve 99% of the variance. The second random field is defined as

$$\kappa_B(x,y,\boldsymbol{\xi}) := 1+\exp(\sigma^2(\xi_1\cos(1.1\pi x)+\xi_2\cos(1.2\pi x)+\xi_3\sin(1.3\pi y)+\xi_4\sin(1.4\pi y))), \tag{29}$$

where $\boldsymbol{\xi}_i(\omega)\sim\mathcal{U}([-1,1])$, $i=1,\dots,4$, and independent. We remark that $1\le\kappa_B(x,y,\boldsymbol{\xi}(\omega))\le 1+\exp(4\sigma^2)$ for all $\omega\in\Omega$, thus (29) is a uniformly bounded random field. We discretize the probability space using SC with a tensorized Gauss-Hermite quadrature, for the log-normal field, and a tensorized Gauss-Legendre quadrature for the bounded random field. The number of nodes for each component is denoted with m. The total number of collocation points is $N=m^M$, with $M=3$ for the log-normal field and $M=4$ for the bounded random field.

Table 1 shows the behaviour of smallest and largest eigenvalues of $\widetilde{S}^{-1}S$ for different values of β, σ^2 and m. As Theorem 2 predicts, $\sigma(\widetilde{S}^{-1}S)$ is well clustered for β large, but it definitely spreads for small values of β. The random field (29) is bounded from below by one, thus the constant $\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\min}(\omega)}\right]$ does not deteriorate as σ^2 increases, and this results in a $\sigma(\widetilde{S}^{-1}S)$ which is bounded uniformly with respect to σ^2. The log-normal field shows instead a weak dependence on σ^2 as $\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\min}(\omega)}\right]$ gets larger as σ^2 increases. The third subtable shows that the preconditioner is robust with respect to the number of collocation points, as expected, since the estimates of Theorem 2 do not involve pointwise quantities such as, e.g., $\min_\omega\kappa_{\min}(\omega)$, but rely on empirical expectations which converge to finite quantities as m increases.

Finally, we remark that $\widetilde{S}$ has favourable properties from the implementational point of view. The major cost when applying $\widetilde{S}^{-1}$ is the matrix-vector multiplication between A^{-1} and a vector v, which is commonly approximated using a spectrally equivalent preconditioner $\widehat{A}^{-1}$. Due to its block diagonal structure, one can compute $\widehat{A}^{-1}v$ in parallel and in distributed way on a cluster. Further, the multiplication with the matrix M_γ can be similarly performed in parallel if $\gamma=0$, and thus the action of the whole preconditioner $\widetilde{P}^{-1}$ is fully parallelizable. In contrast, if $\gamma\neq 0$, all nodes must communicate once at each application of $\widetilde{S}^{-1}$, as one needs to compute the expectation of $\widehat{A}^{-1}v$. However, the major cost of $\widetilde{S}^{-1}$, that is the application of $\widehat{A}^{-1}$ onto a vector, can still be performed in parallel.

TABLE 1 Smallest and largest eigenvalues $\lambda_{\min}, \lambda_{\max}$ of $\widetilde{S}^{-1}S$ for several values of β, σ^2 and m. The number of collocation points is $N = m^4$ for κ_B and $N = m^3$ for κ_L.

β	10^{-2}	10^{-4}	10^{-6}	10^{-8}
$\kappa_B(x,\omega)$	1 - 1.06	1 - 7.12	1 - 613	1 - 61263
$\kappa_L(x,\omega)$	1 - 1.46	1 - 47.64	1 - 4.6e3	1 - 4.66e5

$N_h = 225$, $m = 3$, $\sigma^2 = 0.5$, $\gamma = 0.1$, $L^2 = 0.5$.

σ^2	0.1	0.5	1	1.5
$\kappa_B(x,\omega)$	1 - 1.06	1 - 1.06	1 - 1.05	1 - 1.05
$\kappa_L(x,\omega)$	1 - 1.28	1 - 1.46	1 - 1.83	1 - 2.44

$N_h = 225$, $m = 3$, $\beta = 10^{-2}$, $\gamma = 0.1$, $L^2 = 0.5$.

m	2	3	4	5
$\kappa_B(x,\omega)$	1 - 1.06	1 - 1.06	1 - 1.06	1 - 1.06
$\kappa_L(x,\omega)$	1 - 1.42	1 - 1.46	1 - 1.47	1 - 1.47

$N_h = 225$, $\sigma^2 = 0.5$, $\beta = 10^{-2}$, $\gamma = 0.1$, $L^2 = 0.5$.

5.2 | Matching Schur complement technique

Despite being computationally attractive and presenting a favourable dependence on the random field κ, the spectrum of the preconditioned Schur complement analyzed in Section 5.1 spreads as $\beta \to 0$ and $\gamma \to +\infty$. The second limit is physically less relevant. Indeed, as $\gamma \to +\infty$, the optimal control u tends to $-f$, since $y_\omega = 0$ for a.e. $\omega \in \Omega$ is the only solution of the PDE minimizing the L^2-norm of the pointwise variance. In constrast, small values of β are used in practice to find a control such that the state is as close as possible to the target state, regardless of its cost. To get a clustered spectrum for small values of β, we consider the matching Schur complement technique which consists in looking for a preconditioner $\widehat{S}$ of the Schur complement factorized as

$$\widehat{S} = (A + \alpha\widehat{X})M_\gamma^{-1}(A + \alpha\widehat{X}^\top) = AM_\gamma^{-1}A + \alpha^2\widehat{X}M_\gamma^{-1}\widehat{X}^\top + \alpha\widehat{X}M_\gamma^{-1}A + \alpha AM_\gamma^{-1}\widehat{X}^\top,$$

where $\alpha \in \mathbb{R}$, $\widehat{X} \in \mathbb{R}^{N\cdot N_h \times N\cdot N_h}$ are chosen such that $\alpha^2 X M_\gamma^{-1}\widehat{X}^\top = \frac{1}{\beta}Z\mathbb{1}M_s\mathbb{1}^\top Z$. In other words, $\widehat{S}$ is equal to S, once the cross terms $\alpha\widehat{X}M_\gamma^{-1}A$ and $\alpha AM_\gamma^{-1}\widehat{X}^\top$ are neglected. For some simple deterministic OCPs, it has been proven that this approximation is sufficient to obtain β robustness[18, 38], without essentially increasing the computational cost compared to the approximation $S \approx A_s M_s^{-1} A_s$. Nevertheless, theoretical results are not available for several problems[37, 58], even though improved β robustness has been confirmed by numerical examples. In this subsection, we apply this technique to the model problem and we partially characterize the spectrum of the preconditioned Schur complement. Finally, we present numerical experiments confirming the improved β robustness, and discuss the additional computational costs compared to $\widetilde{S}$.

Defining $\alpha := \frac{1}{\sqrt{\beta}}$ and $\widehat{X} := Z\mathbb{1}M_s\mathbb{1}^\top Z$, a direct calculation shows

$$\alpha^2\widehat{X}M_\gamma^{-1}\widehat{X}^\top = \alpha^2 Z\mathbb{1}M_s\mathbb{1}^\top Z\left(\frac{Z^{-1}}{1+\gamma} + \frac{\gamma}{1+\gamma}\mathbb{1}\mathbb{1}^\top\right)M^{-1}Z\mathbb{1}M_s\mathbb{1}^\top Z =$$
$$\alpha^2 Z\mathbb{1}M_s\mathbb{1}^\top Z\left(\frac{Z^{-1}}{1+\gamma} + \frac{\gamma}{1+\gamma}\mathbb{1}\mathbb{1}^\top\right)Z\mathbb{1}\mathbb{1}^\top Z = \alpha^2 Z\mathbb{1}M_s\left(\frac{\mathbb{1}^\top Z\mathbb{1}}{1+\gamma} + \frac{\gamma(\mathbb{1}^\top Z\mathbb{1})^2}{1+\gamma}\right)\mathbb{1}^\top Z = \frac{1}{\beta}Z\mathbb{1}M_s\mathbb{1}^\top Z,$$

where we used $M^{-1}Z\mathbb{1}M_s\mathbb{1}^\top = Z\mathbb{1}\mathbb{1}^\top$ as M has constant diagonal blocks and $M^{-1}Z = ZM^{-1}$, and $\mathbb{1}^\top Z\mathbb{1} = 1$. Note further that $\widehat{X}^\top = \widehat{X}$. We thus study the Schur complement preconditioner S_{LR} and associated preconditioner P_{LR},

$$S_{\mathrm{LR}} := \left(A + \frac{1}{\sqrt{\beta}}Z\mathbb{1}M_s\mathbb{1}^\top Z\right)M_\gamma^{-1}\left(A + \frac{1}{\sqrt{\beta}}Z\mathbb{1}M_s\mathbb{1}^\top Z\right), \quad P_{\mathrm{LR}} := \begin{pmatrix} C & 0 \\ 0 & S_{\mathrm{LR}} \end{pmatrix} \tag{30}$$

which is symmetric and positive definite. The subscript LR stands for *Low-Rank*, as the matrix in parentheses in the expression for S_{LR} in (30) involves a low-rank perturbation. We partially characterize the spectrum of $S_{\mathrm{LR}}^{-1}S$ in the following theorem.

Theorem 3 (Spectrum of $S_{\mathrm{LR}}^{-1}S$)**.** The matrix $S_{\mathrm{LR}}^{-1}S$ has the eigenvalue $\lambda = 1$ with geometric multiplicity equal to $(N-2)N_h$. The remaining $2N_h$ eigenvalues are real and greater than $\frac{1}{2}$.

Proof. To study the spectrum of $S_{\text{LR}}^{-1}S$ we consider the generalized eigenvalue problem $S\mathbf{v} = \lambda S_{\text{LR}}\mathbf{v}$, and we define the subspaces $\mathcal{H} := \left\{\mathbf{v} \in \mathbb{R}^{N\cdot N_h} : \mathbb{1}^\top Z\mathbf{v} = \sum_{j=1}^N \zeta_j \mathbf{v}_j = 0\right\}$, and $\mathcal{K} := \left\{\mathbf{v} \in \mathbb{R}^{N\cdot N_h} : \mathbb{1}^\top A\mathbf{v} = \sum_{j=1}^N \zeta_j A_{\omega_j}\mathbf{v}_j = 0\right\}$. Both $\mathcal{H}$ and $\mathcal{K}$ have dimension $(N-1)N_h$, and their intersection $\mathcal{H}\cap\mathcal{K}$ has dimension $(N-2)N_h$. We claim that any $\mathbf{v} \in \mathcal{H}\cap\mathcal{K}$ satisfies $S\mathbf{v} = 1\cdot S_{\text{LR}}\mathbf{v}$ and thus it is an eigenvector of $S_{\text{LR}}^{-1}S$ associated with $\lambda = 1$. Indeed,

$$\begin{aligned} S\mathbf{v} &= (AM_\gamma^{-1}A + \frac{1}{\beta}Z\mathbb{1}M_s\mathbb{1}^\top Z)\mathbf{v} = (AM_\gamma^{-1}A + \frac{1}{\beta}Z\mathbb{1}M_s\mathbb{1}^\top Z + \frac{1}{\sqrt{\beta}}\left(Z\mathbb{1}\mathbb{1}^\top MZM_\gamma^{-1}A + AM_\gamma^{-1}ZM\mathbb{1}\mathbb{1}^\top Z\right)\mathbf{v} \\ &= (AM_\gamma^{-1}A + \frac{1}{\beta}Z\mathbb{1}M_s\mathbb{1}^\top Z + \frac{1}{\sqrt{\beta}}\left(Z\mathbb{1}\mathbb{1}^\top A + A\mathbb{1}\mathbb{1}^\top Z\right)\mathbf{v} = S_{\text{LR}}\mathbf{v}, \end{aligned}$$

where we used the equality

$$\mathbb{1}^\top MZM_\gamma^{-1}A = \mathbb{1}^\top MZ\left(\frac{1}{1+\gamma}Z^{-1} + \frac{\gamma}{1+\gamma}\mathbb{1}\mathbb{1}^\top\right)M^{-1}A = \mathbb{1}^\top A.$$

To get the second part of the claim, we define $\widetilde{X} := M_\gamma^{-\frac{1}{2}}A$ and $\widetilde{Y} := \frac{1}{\sqrt{\beta}}M_\gamma^{-\frac{1}{2}}Z\mathbb{1}M_s\mathbb{1}^\top Z$. $\widetilde{X}$ is an invertible matrix, while $\widetilde{Y}$ has rank N_h. Algebraic manipulations show that

$$S = \widetilde{X}^\top\widetilde{X} + \widetilde{Y}^\top\widetilde{Y} \quad \text{and} \quad S_{\text{LR}} = (\widetilde{X}+\widetilde{Y})^\top(\widetilde{X}+\widetilde{Y}).$$

Theorem 1 of Pearson et al.[58] guarantees then that the eigenvalues of $S_{\text{LR}}^{-1}S$ are real and larger than $\frac{1}{2}$. □

Theorem 3 guarantees that $S_{\text{LR}}^{-1}S$ has $(N-2)N_h$ eigenvalues equal to 1, but does not provide estimates for the remaining $2N_h$ eigenvalues. For a deterministic OCP[18], $\widetilde{X} = M_s^{-\frac{1}{2}}A_s$ and $\widetilde{Y} = \frac{1}{\sqrt{\beta}}M_s^{\frac{1}{2}}$, and one can show, using the positive definiteness of $\widetilde{X}^\top\widetilde{Y} + \widetilde{Y}^\top\widetilde{X} = \frac{2}{\sqrt{\beta}}A_s$ that the spectrum of the preconditioned Schur complement lies in $\left[\frac{1}{2}, 1\right]$. Unfortunately in our case, similarly to Pearson et al.[37, 58], $\widetilde{X}^\top\widetilde{Y}+\widetilde{Y}^\top\widetilde{X} = \frac{1}{\sqrt{\beta}}\left(Z\mathbb{1}\mathbb{1}^\top A + A\mathbb{1}\mathbb{1}^\top Z\right)$ is indefinite. Numerically, we have observed that $S_{\text{LR}}^{-1}S$ has N_h eigenvalues in the interval $[\frac{1}{2}, 1]$ and the remaining N_h are larger than 1, but grow very mildly as $\beta \to 0$. We refer to Tables 2 for a further discussion.

5.2.1 | Mean and Chebyshev semi-iterative approximations

The application of S_{LR}^{-1} requires the inversion of the symmetric and positive definite matrix $\left(A + \frac{1}{\sqrt{\beta}}Z\mathbb{1}M_s\mathbb{1}^\top Z\right)$, which consists in a full-rank matrix plus a low-rank perturbation. To do so, we use the Woodbury identity

$$\begin{aligned} \left(A + \frac{1}{\sqrt{\beta}}Z\mathbb{1}M_s\mathbb{1}^\top Z\right)^{-1} &= \left(A + Z\mathbb{1}\frac{1}{\sqrt{\beta}}M_s\mathbb{1}^\top Z\right)^{-1} = A^{-1}\left(I - Z\mathbb{1}\left[\sqrt{\beta}M_s^{-1} + \mathbb{1}^\top ZA^{-1}Z\mathbb{1}\right]^{-1}\mathbb{1}^\top ZA^{-1}\right) \\ &= A^{-1}\left(I - Z\mathbb{1}\left[I + \frac{1}{\sqrt{\beta}}M_s\mathbb{1}^\top ZA^{-1}Z\mathbb{1}\right]^{-1}\frac{1}{\sqrt{\beta}}M_s\mathbb{1}^\top ZA^{-1}\right). \end{aligned} \tag{31}$$

Unfortunately, (31) is of no practical use as it requires the solution of a linear system with $L := \left[I + \frac{1}{\sqrt{\beta}}M_s\mathbb{1}^\top ZA^{-1}Z\mathbb{1}\right]$, which involves $\mathbb{1}^\top ZA^{-1}Z\mathbb{1} = \sum_{i=1}^N \zeta_i A_{\omega_i}^{-1}$. To make the approach feasible, we propose two different approximations.

The first one is based on the mean approximation $\sum_{i=1}^N \zeta_i A_{\omega_i}^{-1} \approx A_0^{-1}$, that is we replace the weighted average of the inverses with the inverse of the mean matrix $A_0 = \sum_{i=1}^N \zeta_i A_{\omega_i}$. Then,

$$\begin{aligned} \left(A + \frac{1}{\sqrt{\beta}}Z\mathbb{1}M_s\mathbb{1}^\top Z\right)^{-1} &\approx A^{-1}\left(I - Z\mathbb{1}\left[I + \frac{1}{\sqrt{\beta}}M_sA_0^{-1}\right]^{-1}\frac{1}{\sqrt{\beta}}M_s\mathbb{1}^\top ZA^{-1}\right) \\ &= A^{-1}\left(I - Z\mathbb{1}A_0\left[A_0 + \frac{1}{\sqrt{\beta}}M_s\right]^{-1}\frac{1}{\sqrt{\beta}}M_s\mathbb{1}^\top ZA^{-1}\right). \end{aligned} \tag{32}$$

We will denote with S_{LRM} the Schur complement preconditioner (30), where the matrices in the parentheses are approximately inverted through (32), and the associated preconditioner by

$$P_{\text{LRM}} := \begin{pmatrix} C & 0 \\ 0 & S_{\text{LRM}} \end{pmatrix}. \tag{33}$$

As for forward problems[27, 59], this approximation is satisfactory for small variances and uniformly bounded random fields, while it is definitely poor if the variance is large and for random fields with unbounded random variables, e.g. the log-normal field.
As an alternative approximation, it would be tempting to use a Krylov method to approximate the inverse of L. However, any Krylov method is a non-linear map with respect to the right hand side and the initial vector[60]. Hence, we instead approximate the solution of $L\mathbf{v} = \mathbf{z}$ using N_{it} iterations of the damped preconditioned stationary iterative method that, starting from an initial guess $\mathbf{v}^0$, computes

$$\mathbf{v}^k = \mathbf{v}^{k-1} + \alpha P_0^{-1}(\mathbf{z} - L\mathbf{v}^{k-1}), \quad k = 1, \dots, N_{\text{it}},$$

accelerated by the Chebyshev Semi-Iterative method[61, Section 10.1.5]. We will denote with S_{LRC} the Schur complement preconditioner (30) obtained by approximating the inverse of L in (31) with such iterative procedure, and the associated preconditioner

$$P_{\text{LRC}} := \begin{pmatrix} C & 0 \\ 0 & S_{\text{LRC}} \end{pmatrix}. \tag{34}$$

As we will see in Tables 2, P_{LRC} allows us to recover robustness with respect to β, but the cost of each iteration is larger compared to an iteration of P_{LRM}. In our experiments, we set $P_0^{-1} = (I + \frac{1}{\sqrt{\beta}} M_s A_0^{-1})^{-1} = A_0 \left(A_0 + \frac{1}{\sqrt{\beta}} M_s \right)^{-1}$. The Chebyshev Semi-Iterative method requires two parameters $\underline{\lambda}$ and $\overline{\lambda}$ such that $-1 < \underline{\lambda} \leq \lambda_1 \leq \cdots \leq \lambda_N \leq \overline{\lambda} < 1$, where λ_j are the eigenvalues of $I - \alpha P_0^{-1} L$[60]. To estimate the spectrum of $(I - \alpha P_0^{-1} L)$, we rely on the following Lemma.

Lemma 8. The spectrum of $P_0^{-1} L$ is real and bounded from below by 1.

Proof. Algebraic manipulations lead to

$$P_0^{-1} L = I + \left(M_s^{-1} + \frac{1}{\sqrt{\beta}} A_0^{-1} \right)^{-1} \frac{1}{\sqrt{\beta}} \left(\mathbb{1}^\top Z A^{-1} Z \mathbb{1} - A_0^{-1} \right).$$

Hence, if $\sum_{j=1}^N \zeta_j A_{\omega_j}^{-1} - A_0^{-1}$ is semi-positive definite, then $P_0^{-1} L$ has real eigenvalues and $\lambda_{\min} > 1$. To show this, take an arbitrary $0 \neq \mathbf{v} \in \mathbb{R}^{N_h}$, and consider the map $\phi_{\mathbf{v}} : S_{++}^n \to \mathbb{R}$, where S_{++}^n is the set of positive definite matrices in $\mathbb{R}^{n \times n}$, defined as $\phi_{\mathbf{v}} := \mathbf{v}^\top A^{-1} \mathbf{v}$. The map $\phi_{\mathbf{v}}$ is convex[62, Lemma 1]. Thus due to Jensen's inequality

$$\mathbf{v}^\top \left(\sum_{j=1}^N \zeta_j A_{\omega_j}^{-1} \right) \mathbf{v} - \mathbf{v}^\top \left(\sum_{j=1}^N \zeta_j A_{\omega_j} \right)^{-1} \mathbf{v} = \sum_{j=1}^N \zeta_j \phi_{\mathbf{v}}(A_{\omega_j}) - \phi_{\mathbf{v}} \left(\sum_{j=1}^N \zeta_j A_{\omega_j} \right) \geq 0,$$

hence, due to the arbitrariness of $\mathbf{v}$, $\mathbb{1}^\top Z A^{-1} Z \mathbb{1} - A_0^{-1}$ is semi-positive definite. □

Let $\lambda_{\min}$ and $\lambda_{\max}$ be the minimum and maximum eigenvalues of $P_0^{-1} L$. From Lemma 8, it follows that $\sigma(I - \alpha P_0^{-1} L) \subset [1 - \alpha\lambda_{\max}, 1 - \alpha\lambda_{\min}]$ as $P_0^{-1} L$ has real and positive spectrum. The parameter α is needed to guarantee the convergence of the stationary method, that is $\rho(I - \alpha P_0^{-1} L) < 1$. The optimal α which minimizes $\rho(I - \alpha P_0^{-1} L)$ is $\alpha_{\text{opt}} = \frac{2}{\lambda_{\min}(P_0^{-1} L) + \lambda_{\max}(P_0^{-1} L)}$. However, α_{opt} leads to a spread spectrum, while the Chebyshev Semi-Iterative method takes advantage of clustered spectra, like Krylov methods[61, Section 10.1.5]. We therefore set $\alpha := \frac{1}{1 + \lambda_{\max}(P_0^{-1} L)}$, where $\lambda_{\max}(P_0^{-1} L)$ is approximated, once and for all, using few iterations of the power method. This choice guarantees the convergence of the iterative method since $\alpha \leq \frac{2}{\lambda_{\max}(P_0^{-1} L)}$. Finally, in the Chebyshev Semi-Iterative method we take $\underline{\lambda} = 1 - \alpha\lambda_{\max}$ and $\overline{\lambda} = 1 - \alpha$.
Table 2 compares the behaviour of the extremal eigenvalues of the Schur complement preconditioned by S_{LR}, S_{LRM} and S_{LRC} in different regimes. S_{LR} has a very high computational cost and is of no practical use. It is included in Table 2 as a reference, in order to assess how well the approximated versions S_{LRM} and S_{LRC} perform, compared to S_{LR}. Notice that S_{LRM} and S_{LRC} have different cost per iteration. We defer to Section 7 a comparison of the two in terms of computational efficiency, and we focus here only on the clustering of the spectrum of the preconditioned Schur complements.
From the first two tables, we observe that S_{LR} shows a (very weak) dependence on β and on σ^2, emphasized in the case of the log-normal field, but the spectrum still remains sufficiently clustered. Notice that $\sigma(S_{\text{LR}}^{-1} S)$ is not contained in the interval $[\frac{1}{2}, 1]$, as in the deterministic case[18]. The third table shows that S_{LR} is robust with respect to finer discretizations of the probability space.

Let us now consider the approximations S_{LRM} and S_{LRC}. On the one hand, S_{LRM} is a valid choice for the uniformly bounded random field and for values of β not too small. It definitely performs poorly for the log-normal field. On the other hand, S_{LRC}^{-1} matches the performance of the exact preconditioner S_{LR}, both for the bounded and log-normal fields, with a small number N_{it} of Chebyshev semi-iterations. However, to obtain good performances, N_{it} has to increase as σ^2 increases, especially for the log-normal field, due to the poorer performance of P_0 as a preconditioner in the inner Chebyshev semi-iterations. From the

TABLE 2 Smallest and largest eigenvalues $\lambda_{\min} - \lambda_{\max}$ of $S_{\text{LR}}^{-1}S$, $S_{\text{LRM}}^{-1}S$ and $S_{\text{LRC}}^{-1}S$. The number of collocation points is $N = m^4$ for κ_B and $N = m^3$ for κ_L.

β		10^{-2}	10^{-4}	10^{-6}	10^{-8}
$S_{\text{LR}}^{-1}S$	$\kappa_B(x,\omega)$	0.68 - 1.00	0.50 - 1.02	0.50 - 1.17	0.50 - 1.30
$S_{\text{LRM}}^{-1}S$	$\kappa_B(x,\omega)$	0.68 - 1.00	0.48 - 1.02	0.13 - 1.07	8.7e-3 - 30.59
$S_{\text{LRC}}^{-1}S$	$\kappa_B(x,\omega)$	0.68 - 1.00	0.50 - 1.02	0.50 - 1.17	0.50 - 1.30
$S_{\text{LR}}^{-1}S$	$\kappa_L(x,\omega)$	0.52 - 1.11	0.50 - 1.74	0.50 - 2.39	0.52 - 2.61
$S_{\text{LRM}}^{-1}S$	$\kappa_L(x,\omega)$	0.45 - 1.12	0.02 - 2.13	1e-4 - 7.73e2	1.3e-5 - 8.98e4
$S_{\text{LRC}}^{-1}S$	$\kappa_L(x,\omega)$	0.52 - 1.11	0.50 -1.74	0.50 - 2.39	0.52 - 2.61

$N_h = 225$, $m = 3$, $\sigma^2 = 0.5$, $\gamma = 0.1$, $L^2 = 0.5$. $N_{\text{it}} = 2$ for $\kappa_B(x,\omega)$ and $N_{\text{it}} = 4$ for $\kappa_L(x,\omega)$.

σ^2		0.1	0.5	1	1.5
$S_{\text{LR}}^{-1}S$	$\kappa_B(x,\omega)$	0.50 - 1.04	0.50 - 1.30	0.50 - 1.66	0.50 - 1.98
$S_{\text{LRM}}^{-1}S$	$\kappa_B(x,\omega)$	0.49 - 1.01	8.7e-3 - 30.59	3.7e-8 - 1.00e3	4.4e-5 - 1.03e4
$S_{\text{LRC}}^{-1}S$	$\kappa_B(x,\omega)$	0.50 - 1.04	0.50 - 1.30	0.49 - 1.67	0.09 - 1.97
$S_{\text{LR}}^{-1}S$	$\kappa_L(x,\omega)$	0.52 - 1.43	0.52 - 2.61	0.52 - 4.35	0.52 - 6.54
$S_{\text{LRM}}^{-1}S$	$\kappa_L(x,\omega)$	5.9e-4 - 1.52e3	1.3e-5 - 8.98e4	0.23 - 9.82e5	0.70 - 5.87e6
$S_{\text{LRC}}^{-1}S$	$\kappa_L(x,\omega)$	0.52 - 1.43	0.52 - 2.61	0.52 - 4.34	0.51 - 6.54

$N_h = 225$, $m = 3$, $\beta = 10^{-8}$, $\gamma = 0.1$, $L^2 = 0.5$. N_{it} is equal to 2 for $\kappa_B(x,\omega)$ and equal to $2, 4, 6, 8$ for $\sigma^2 = 0.1, 0.5, 1, 1.5$ respectively for $\kappa_L(x,\omega)$.

m		2	3	4	5
$S_{\text{LR}}^{-1}S$	$\kappa_B(x,\omega)$	0.50 - 1.17	0.50 - 1.17	0.50 - 1.17	0.50 - 1.17
$S_{\text{LRM}}^{-1}S$	$\kappa_B(x,\omega)$	0.13 - 1.07	0.13 - 1.07	0.13 - 1.07	0.13 - 1.07
$S_{\text{LRC}}^{-1}S$	$\kappa_B(x,\omega)$	0.50 - 1.17	0.50 - 1.17	0.50 - 1.17	0.50 - 1.17
$S_{\text{LR}}^{-1}S$	$\kappa_L(x,\omega)$	0.50 - 2.13	0.50 - 2.39	0.50 - 2.43	0.50 - 2.43
$S_{\text{LRM}}^{-1}S$	$\kappa_L(x,\omega)$	0.0025 - 634	1e-4 - 773	1e-4 - 784	2.9e-3 - 785
$S_{\text{LRC}}^{-1}S$	$\kappa_L(x,\omega)$	0.50 - 2.13	0.50 - 2.39	0.5 - 2.43	0.5 - 2.43

$N_h = 225$, $\sigma^2 = 0.5$, $\beta = 10^{-6}$, $\gamma = 0.1$, $L^2 = 0.5$, $N_{\text{it}} = 2$ for κ_B and $N_{\text{it}} = 4$ for κ_L.

computational point of view, both S_{LRM} and S_{LRC} require the inversion of four times (approximately and possibly in parallel) the matrix A at each outer Krylov iteration, in constrast with $\widetilde{S}$ which requires the inversion (possibly, approximately) of A only twice per iteration. There is further a synchronization step where the reduced size system involving the matrix L, or its mean approximation, is approximately solved.

6 | PRECONDITIONING IN A HILBERT SETTING

Another technique to develop robust preconditioners for parameter-dependent saddle point problems is called "operator preconditioning", which has its foundation in the analysis of iterative methods in Hilbert spaces[20, 39, 40]. In our setting, the parameters will be the couple (β, γ). In a nutshell, let $\mathscr{T}$ be a self-adjoint operator from $\mathcal{V} \to \mathcal{V}'$, and suppose we want to solve the linear equation $\mathscr{T}x = f$ in $\mathcal{V}'$. As $\mathscr{T}$ is a map between two different Hilbert spaces, we may identify an isomorphism $\mathcal{R} : \mathcal{V}' \to \mathcal{V}$, and consider the equivalent problem $\mathcal{R}\mathscr{T}x = \mathcal{R}f$ in $\mathcal{V}$. To choose an operator $\mathcal{R}$, one can define first a scalar product on $\mathcal{V}$,

and then set $\mathcal{R}$ equal the Riesz isomorphism such that $\langle \mathscr{T}x, y\rangle = (\mathcal{R}\mathscr{T}x, y)_{\mathcal{V}}$ and $\|\mathcal{R}\mathscr{T}x\|_{\mathcal{V}} = \|\mathscr{T}x\|_{\mathcal{V}'}$, so that

$$\begin{aligned}\|\mathcal{R}\mathscr{T}\|_{\mathcal{L}(\mathcal{V},\mathcal{V})} &= \sup_{0\neq x\in\mathcal{V}} \frac{\|\mathcal{R}\mathscr{T}x\|_{\mathcal{V}}}{\|x\|_{\mathcal{V}}} = \sup_{0\neq x\in\mathcal{V}} \frac{\|\mathscr{T}x\|_{\mathcal{V}'}}{\|x\|_{\mathcal{V}}} = \|\mathscr{T}\|_{\mathcal{L}(\mathcal{V},\mathcal{V}')},\\ \|\left(\mathcal{R}\mathscr{T}\right)^{-1}\|_{\mathcal{L}(\mathcal{V},\mathcal{V})} &= \sup_{0\neq x\in\mathcal{V}} \frac{\|\left(\mathcal{R}\mathscr{T}\right)^{-1}x\|_{\mathcal{V}}}{\|x\|_{\mathcal{V}}} = \left(\inf_{0\neq x\in\mathcal{V}} \frac{\|\mathcal{R}\mathscr{T}x\|_{\mathcal{V}}}{\|x\|_{\mathcal{V}}}\right)^{-1} = \left(\inf_{0\neq x\in\mathcal{V}} \frac{\|\mathscr{T}x\|_{\mathcal{V}'}}{\|x\|_{\mathcal{V}}}\right)^{-1} = \|\mathscr{T}^{-1}\|_{\mathcal{L}(\mathcal{V}',\mathcal{V})}.\end{aligned} \tag{35}$$

Hence, if one finds an appropriate, e.g., (β,γ)-dependent, scalar product $(\cdot,\cdot)_{\mathcal{V}}$ (hence, a norm on $\mathcal{V}$), so that $\|\mathscr{T}\|_{\mathcal{L}(\mathcal{V},\mathcal{V}')} \leq C$ and $\|\mathscr{T}^{-1}\|_{\mathcal{L}(\mathcal{V}',\mathcal{V})} \leq \alpha$, with C and α parameter-independent, then considering the Riesz isomorphism $\mathcal{R}$ associated with $(\cdot,\cdot)_{\mathcal{V}}$, one obtains using (35), $\kappa(\mathcal{R}\mathscr{T}) = \|\mathcal{R}\mathscr{T}\|_{\mathcal{L}(\mathcal{V},\mathcal{V})}\|\left(\mathcal{R}\mathscr{T}\right)^{-1}\|_{\mathcal{L}(\mathcal{V},\mathcal{V})} \leq C\alpha$, that is, the condition number of the the preconditioned system $\mathcal{R}\mathscr{T}$ is independent of (β,γ).

In this section, we apply the operator preconditioning paradigm to the robust optimal control problem (9), the final goal being to find the two (β,γ)-independent constants C and α. To do so, we first consider the continuous optimality system ((38)), and exploit its structure (Lemma 9 and Theorem 4) to derive an equivalent formulation involving reduced spaces ((43)). Second, we define (β,γ)-dependent norms on these reduced spaces ((44)) and prove the well-posedness of the optimality system at the continuous level in Theorem 5. Finally, we identify the constants C and α ((48) and (49)). Notice that the functional space of control functions U is now set equal to Y.

Let us consider the optimality conditions in (10). We introduce the space $\widehat{\mathscr{X}} := \mathscr{Y} \times U$ and the bilinear forms

$$\mathcal{I} : Y' \to \mathscr{Y}' \quad \text{such that} \quad \langle \mathcal{I}f, v\rangle := \int_\Omega \langle f, v(\cdot,\omega)\rangle d\mathbb{P}(\omega) = \mathbb{E}\left[\langle f, v_\omega\rangle\right], \quad \forall v \in \mathscr{Y},$$

$$\mathcal{C} : \widehat{\mathscr{X}} \times \widehat{\mathscr{X}} \to \mathbb{R} \quad \text{such that} \quad \mathcal{C}\left((y,u),(w,v)\right) := \mathbb{E}\left[\langle \Lambda_{L^2}((1+\gamma)y_\omega - \gamma\mathbb{E}\left[y_\omega\right]), w_\omega\rangle\right] + \beta\langle \Lambda_Y u, v\rangle, \tag{36}$$

$$\mathcal{B} : \widehat{\mathscr{X}} \times \mathscr{Y} \to \mathbb{R} \quad \text{such that} \quad \mathcal{B}\left((y,u),p\right)) := \mathbb{E}\left[\langle \mathcal{A}_\omega y_\omega, p_\omega\rangle - \langle \Lambda_Y u, p_\omega\rangle\right], \tag{37}$$

The optimality conditions can be formulated as:

$$\begin{aligned}\text{Find } (\underline{x},p) \in \widehat{\mathscr{X}} \times \mathscr{Y} \text{ such that } \mathcal{C}(\underline{x},\underline{r}) + \mathcal{B}(\underline{r},p) &= \langle \mathcal{F},\underline{r}\rangle, \quad \forall \underline{r} = (w,v) \in \widehat{\mathscr{X}},\\ \mathcal{B}(\underline{x},q) &= \langle \mathcal{G},q\rangle, \quad \forall q \in \mathscr{Y},\end{aligned} \tag{38}$$

where $\mathcal{F} \in \widehat{\mathscr{X}}' : \langle \mathcal{F},\underline{r}\rangle = \mathbb{E}\left[\langle \Lambda_{L^2} y_d, w_\omega\rangle\right], \forall \underline{r} = (w,v) \in \widehat{\mathscr{X}}$, and $\mathcal{G} \in \mathscr{Y}' : \langle \mathcal{G},q\rangle = \mathbb{E}\left[\langle f, q_\omega\rangle\right], \forall q \in \mathscr{Y}$. The bilinear form $\mathcal{C}(\cdot,\cdot)$ is symmetric, as a direct generalization of Lemma 4 shows.

To obtain (β,γ)-independent continuity constants α and C, we have to consider a slightly modified formulation of (38). Let us define the subspace in $\mathscr{Y}$ of functions with zero average, $G := \{v \in \mathscr{Y} : \mathbb{E}\left[v(\cdot,\omega)\right] = 0\}$ and its polar space $G^0 := \left\{\mathcal{F} \in \mathscr{Y}' : \mathcal{F}(v) = 0,\ \forall v \in G.\right\}$. We can prove the following Lemma.

Lemma 9 (Isomorphism between Y' and G^0)**.** *$\mathcal{I}$ is an isomorphism between Y' and G^0.*

Proof. To prove that $\mathcal{I}$ is injective, we show that for any $\widetilde{u},\widetilde{w} \in Y'$, $\mathcal{I}\widetilde{u} = \mathcal{I}\widetilde{w}$ in $\mathscr{Y}'$ implies $\widetilde{u} = \widetilde{w}$. A direct calculation leads to

$$\langle \mathcal{I}\widetilde{u} - \mathcal{I}\widetilde{w}, v\rangle = \mathbb{E}\left[\langle \widetilde{u} - \widetilde{w}, v_\omega\rangle\right] = 0, \quad \forall v \in \mathscr{Y}. \tag{39}$$

Consider now the sets $\Gamma^n := \left\{\omega \in \Omega : \max\left\{\kappa_{\max}(\omega), 1/\kappa_{\min}(\omega)\right\} < n\right\}$. The sets Γ^n are measurable, and $|\Gamma^n| > 0$ for a sufficiently large n. Taking $v(x,\omega) = \mathbb{1}_{\Gamma^n}(\omega)\phi(x)$, where $\phi \in Y$ is arbitrary and n is large enough, we have that $v \in \mathscr{Y}$ and (39) implies $\widetilde{u} = \widetilde{w}$ in Y'.

For the surjectivity, first note that $\mathrm{Im}\mathcal{I} \subset G^0$ since

$$\langle \mathcal{I}\widetilde{u}, v\rangle = \mathbb{E}\left[\langle \widetilde{u}, v_\omega\rangle\right] = \langle \widetilde{u}, \mathbb{E}\left[v_\omega\right]\rangle = 0, \quad \forall v \in G,$$

where one can exchange the duality pair between Y and Y' and the expectation operator due to the property of the Bochner integral[44, E. 11]. We now prove that $G^0 \subset \mathrm{Im}\mathcal{I}$. Take any $F \in G^0 \subset \mathscr{Y}'$. Due to Riesz theorem, there exists a $\widetilde{f} \in \mathscr{Y}$ such that $F(v) = a(\widetilde{f}, v)$, $\forall v \in \mathscr{Y}$. Restricting to $v \in G$,

$$0 = F(v) = a(\widetilde{f}, v) = \mathbb{E}\left[\langle \mathcal{A}_\omega \widetilde{f}_\omega, v_\omega\rangle\right]. \tag{40}$$

Consider now the set $\Gamma^\infty := \cup_{n\in\mathbb{N}}\Gamma^n$ which has full measure, i.e. $\mathbb{P}(\Gamma^\infty) = 1$ [2] and the restricted sigma algebra $\mathcal{M} := \{E \cap \Gamma^\infty : E \in \mathcal{F}\}$ on Γ^∞. Let us define $v = \psi(\omega)\phi(x)$ where $\phi(x) \in Y$, and $\psi(\omega) = \mathbb{1}_E(\omega) - \overline{\mathbb{1}_E}$ for an arbitrary $E \in \mathcal{M}$,

[2] If $\mathbb{P}\left((\Gamma^\infty)^c\right) > 0$ then either $\frac{1}{\kappa_{\min}(\omega)}$ or $\kappa_{\max}(\omega)$ would not lie in $L^1(\Omega)$, contradicting Assumption 1.

with $\overline{\mathbb{1}_E} := \mathbb{E}\left[\mathbb{1}_E(\omega)\right]$, so that $v \in G \subset \mathscr{Y}$. Replacing the expression of v into (40), we obtain

$$\mathbb{E}\left[\mathbb{1}_E\langle \mathcal{A}_\omega \widetilde{f}_\omega, \phi\rangle\right] = \mathbb{E}\left[\overline{\mathbb{1}_E}\langle \mathcal{A}_\omega \widetilde{f}_\omega, \phi\rangle\right] = \overline{\mathbb{1}_E}\langle \mathbb{E}\left[\mathcal{A}_\omega \widetilde{f}_\omega\right], \phi\rangle,$$

and denoting $f := \mathbb{E}\left[\mathcal{A}_\omega \widetilde{f}_\omega\right]$ we have,

$$\mathbb{E}\left[\mathbb{1}_E\langle \mathcal{A}_\omega \widetilde{f}_\omega - f, \phi\rangle\right] = \int_E \langle \mathcal{A}_\omega \widetilde{f}_\omega - f, \phi\rangle = 0, \quad \forall E \in \mathcal{M}, \forall \phi \in Y. \tag{41}$$

Due to the arbitrariness of E and the full measure of Γ^∞, (41) implies $\langle \mathcal{A}_\omega \widetilde{f}_\omega - f, \phi\rangle = 0$ $\mathbb{P}$-a.s., $\forall \phi \in Y$, hence $\mathcal{A}_\omega \widetilde{f}_\omega = f \in Y'$ $\mathbb{P}$-a.s.
Thus, we conclude

$$\mathcal{F}(v) = a(\widetilde{f}, v) = \langle \mathcal{A} f, v\rangle = \mathbb{E}\left[\langle \mathcal{A}_\omega \widetilde{f}_\omega, v_\omega\rangle\right] = \mathbb{E}\left[\langle f, v_\omega\rangle\right] = \mathcal{I} f(v), \quad \forall v \in \mathscr{Y},$$

that is, for every $\mathcal{F} \in G^0$, there exists a $f \in Y'$ such that $\mathcal{F} = \mathcal{I} f$. □

Considering the state equation, we remark that

$$a(y, v) = \langle \mathcal{I}\left(\Lambda_U u + f\right), v\rangle = 0 \quad \forall v \in G,$$

that is, if y is a solution to the state equation, then y is a-orthogonal to G, i.e. $y \in G^\perp := \{y \in \mathscr{Y} : a(y, v) = 0,\ \forall v \in G\}$. In other words, whatever control function u we choose, we cannot obtain a generic state $y \in \mathscr{Y}$, but the state solution is constrained to lie in the subspace $G^\perp$ of $\mathscr{Y}$.

Remark 1. A similar constraint on the state variable has been observed in Elvetun et al.[32] for deterministic OCP with a control function acting only on a subdomain $\widetilde{D} \subset D$. The parallelism between a robust OCPUU and a deterministic OCP with local control lies in the observation that, in both OCPs, one cannot generate the whole dual of the state space using only elements of the control space. For robust OCPUU one has $\mathrm{Im}\,\mathcal{I} \subsetneq \mathscr{Y}'$, see Lemma 9, and similarly for a deterministic OCP with local control one cannot generate $(H^1(D))'$ using only elements of $(H^1(\widetilde{D}))'$[32]. From the algebraic point of view, this leads to low-rank perturbed Schur complements, where the rank of the perturbation is equal to the size of the finite element discretization of the control space (see (17) and Ref. [33]).

To get (β, γ) robust continuity constants at the continuous level, it is essential to use these properties of the continuous formulation of the saddle point system. We thus consider the OCP (9) with the state space equal to $G^\perp$. As the residual of the state equation $\mathcal{A}y - \mathcal{I}(\Lambda_Y u + f) \in G^0 = (G^\perp)'$, the adjoint variable p belongs to $G^\perp$ as well. Computing the directional derivatives of the restricted Lagrangian $\widehat{\mathcal{L}}(y, u, p) : G^\perp \times Y \times G^\perp \to \mathbb{R}$ with $\widehat{\mathcal{L}}(y, u, p) := \mathcal{L}(y, u, p)$, the optimality system becomes: find $(y, u, p) \in G^\perp \times Y \times G^\perp$ such that

$$\begin{aligned}
\mathbb{E}\left[\langle \mathcal{A}_\omega p_\omega, v_\omega\rangle\right] + \mathbb{E}\left[\langle \Lambda_{L^2}\left(y_\omega + \gamma(y_\omega - \mathbb{E}\left[y_\omega\right])\right), v_\omega\rangle\right] &= \mathbb{E}\left[\langle \Lambda_{L^2} y_d, v_\omega\rangle\right], \ \forall v \in G^\perp, \\
\langle \beta \Lambda_Y u - \Lambda_Y \mathbb{E}\left[p_w\right], v\rangle &= 0, \qquad \forall v \in Y, \\
\mathbb{E}\left[\langle \mathcal{A}_\omega y_\omega, v_\omega\rangle\right] - \mathbb{E}\left[\langle \Lambda_Y u, v_\omega\rangle\right] &= \mathbb{E}\left[\langle f, v_\omega\rangle\right], \qquad \forall v \in G^\perp.
\end{aligned} \tag{42}$$

Defining the space $\mathcal{X} := G^\perp \times Y$ and using the bilinear and linear forms defined above, the optimality conditions read: Find $(\underline{x}, p) \in \mathcal{X} \times G^\perp$ such that

$$\begin{aligned}
C(\underline{x}, \underline{v}) + \mathcal{B}(\underline{v}, p) &= \langle \mathcal{F}, \underline{v}\rangle, \quad \forall \underline{v} \in \mathcal{X}, \\
\mathcal{B}(\underline{x}, q) &= \langle \mathcal{G}, q\rangle, \quad \forall q \in G^\perp.
\end{aligned} \tag{43}$$

We now prove an important result stating that $G^\perp$ is homeomorphic to Y. We introduce the operator $\mathcal{E} : Y \to G^\perp$ defined as $\mathcal{E} y = \mathcal{A}^{-1}\mathcal{I}\Lambda_Y y$, and its inverse $\mathcal{E}^{-1} : G^\perp \to Y$ such that $\mathcal{E}^{-1} v = \Lambda_Y^{-1}\mathcal{I}_{G^0}^{-1}\mathcal{A} v$, where $\mathcal{I}_{G^0}^{-1}$ is the inverse of the map $\mathcal{I} : Y' \to G^0$.

Theorem 4. Let us consider $G^\perp$ equipped with the norm $\|\cdot\|_{\mathcal{A}}^2 := \langle \mathcal{A}\cdot, \cdot\rangle$ and Y equipped with the norm $\|\cdot\|_Y$. The map $\mathcal{E}^{-1}$ is a homeomorphism between $G^\perp$ and Y, and further it holds

$$\frac{1}{\sqrt{\mathbb{E}\left[\frac{1}{\kappa_{\min}(\omega)}\right]}}\|y\|_{\mathcal{A}} \le \|\mathcal{E}^{-1} y\|_Y \le \sqrt{\mathbb{E}\left[\kappa_{\max}(\omega)\right]}\|y\|_{\mathcal{A}}.$$

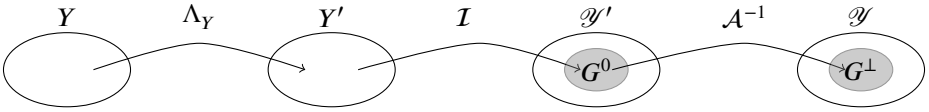

FIGURE 1 Graphical representation of the maps between the different functional spaces.

Proof. Observe that for every $y \in G^\perp$, $\mathcal{E}^{-1}y$ is well defined because $\mathcal{A}y \in G^0$ and thus it can be written as $\mathcal{A}y = \mathcal{I}\widetilde{L}$ for a unique $\widetilde{L} \in Y'$ due to Lemma 9. Finally the Riesz's representation isomorphism on Y returns the Riesz representative L, so that $L = \mathcal{E}^{-1}y$. Moreover, on the one hand

$$\|y\|_{\mathcal{A}}^2 = \langle \mathcal{A}y, y\rangle = \mathbb{E}\left[\langle \mathcal{I}_{G^0}^{-1}\mathcal{A}y, y_\omega\rangle\right] \le \|\Lambda_Y^{-1}\mathcal{I}_{G^0}^{-1}\mathcal{A}y\|_Y \mathbb{E}\left[\|y_\omega\|_Y\right] \le \|\mathcal{E}^{-1}y\|_Y \mathbb{E}\left[\frac{1}{\sqrt{\kappa_{\min}(\omega)}}\|y_\omega\|_{\mathcal{A}_\omega}\right]$$
$$\le \|\mathcal{E}^{-1}y\|_Y \left(\mathbb{E}\left[\frac{1}{\kappa_{\min}(\omega)}\right]\right)^{\frac{1}{2}} \left(\mathbb{E}\left[\|y_\omega\|_{\mathcal{A}_\omega}^2\right]\right)^{\frac{1}{2}} = \|\mathcal{E}^{-1}y\|_Y \sqrt{\mathbb{E}\left[\frac{1}{\kappa_{\min}(\omega)}\right]}\|y\|_{\mathcal{A}},$$

which implies $\|y\|_{\mathcal{A}} \le \sqrt{\mathbb{E}\left[\frac{1}{\kappa_{\min}(\omega)}\right]}\|\mathcal{E}^{-1}y\|_Y$. On the other hand,

$$\|\mathcal{E}^{-1}y\|_Y^2 = (\Lambda_Y^{-1}\mathcal{I}_{G^0}^{-1}\mathcal{A}y, \mathcal{E}^{-1}y)_Y = \langle \mathcal{I}_{G^0}^{-1}\mathcal{A}y, \mathcal{E}^{-1}y\rangle = \mathbb{E}\left[\langle \mathcal{I}_{G^0}^{-1}\mathcal{A}y, \mathcal{E}^{-1}y\rangle\right] = \mathbb{E}\left[\langle \mathcal{A}_\omega y_\omega, \mathcal{E}^{-1}y\rangle\right]$$
$$\le \mathbb{E}\left[\|y_\omega\|_{\mathcal{A}_\omega}\|\mathcal{E}^{-1}y\|_{\mathcal{A}_\omega}\right] \le \|\mathcal{E}^{-1}y\|_Y \mathbb{E}\left[\|y_\omega\|_{\mathcal{A}_\omega}\sqrt{\kappa_{\max}(\omega)}\right] \le \|\mathcal{E}^{-1}y\|_Y \|y\|_{\mathcal{A}}\sqrt{\mathbb{E}\left[\kappa_{\max}(\omega)\right]}.$$

which implies $\|\mathcal{E}^{-1}y\|_Y \le \sqrt{\mathbb{E}\left[\kappa_{\max}(\omega)\right]}\|y\|_{\mathcal{A}}$. □

Fig. 1 provides a useful graphical overview of the relations between the spaces Y, Y', G^0 and $G^\perp$. Due to Theorem (4) we can parametrize the space $G^\perp \subset \mathscr{Y}$, since any $y \in G^\perp$ is in a one-to-one correspondence with an element of Y through the operator $\mathcal{E}$. This property is essential to prove a (β, γ)-independent inf-sup condition.
Let us now consider the following functional setting,

$$\mathcal{Y} := \left(G^\perp, (\cdot,\cdot)_{\mathcal{Y}}\right), \quad U = \left(Y, (\cdot,\cdot)_U\right), \quad \mathcal{X} = \left(\mathcal{Y}\times U, (\cdot,\cdot)_{\mathcal{X}}\right), \quad \mathcal{P} := \left(G^\perp, (\cdot,\cdot)_{\mathcal{P}}\right),$$

where the scalar products define the weighted-norms

$$\begin{aligned}
\|y\|_{\mathcal{Y}}^2 &:= (y,y)_{\mathcal{Y}} = \mathbb{E}\left[(y_\omega, y_\omega)_{L^2} + \gamma\left(y_\omega - \mathbb{E}\left[y_\omega\right], y_\omega - \mathbb{E}\left[y_\omega\right]\right)_{L^2}\right] + \beta\mathbb{E}\left[\langle \mathcal{A}_\omega y_\omega, y_\omega\rangle\right] = \\
&= (y,y)_{L^2,\gamma} + \beta(y,y)_{\mathcal{A}}, \\
\|u\|_U^2 &:= (u,u)_U = \beta(u,u)_Y, \\
\|(y,u)\|_{\mathcal{X}}^2 &:= ((y,u),(y,u))_{\mathcal{X}} = (y,y)_{\mathcal{Y}} + (u,u)_U, \\
\|p\|_{\mathcal{P}}^2 &:= \frac{1}{\beta}\mathbb{E}\left[\langle \mathcal{A}_\omega p_\omega, p_\omega\rangle\right] = \frac{1}{\beta}(p,p)_{\mathcal{A}}.
\end{aligned} \tag{44}$$

For the state variable y we introduce the scalar product $(\cdot,\cdot)_{L^2,\gamma}$ which consists of two parts: the first one is the simple $L^2(\Omega, L^2(D))$ norm. The second part proportional to γ consists in expectation of the $L^2(D)$ norm of the difference between y_ω from its mean value.

Remark 2. We remark that the energy norm and the $L^2(\Omega, Y)$ norm are not equivalent, unless $\kappa(x,\omega) \in L^\infty(\Omega, L^\infty(D))$. In the latter case, one could show the well-posedness of the saddle point system working exclusively with the energy norm (obtaining, though, β-dependent constants). In constrast, for a not uniformly bounded $\kappa(x,\omega)$, one would fail to bound the bilinear form $C(\cdot,\cdot)$ with only the energy norm, and thus one would have to rely on the framework of Gittelson et al.[42,43], and introduce an energy norm with respect to an auxiliary measure to be able to bound the $L^2(\Omega, Y)$ norm with the new modified energy norm. In this manuscript, we are interested to study β-robust preconditioners, which are derived taking a weighted combination of both the $L^2(\Omega, Y)$ norm and the energy norm, see, e.g., Zulehner[19], Mardal et al.[20], Schöberl et al.[63] for deterministic OCP, and thus we can avoid the framework of Refs.[42,43], since we do not need any relation between the two norms, as the next Theorem shows.

Theorem 5 (Well-posedness of (43))**.**

1. The bilinear form $\mathcal{C}$ is bounded: $\mathcal{C}(\underline{x},\underline{v}) \le \|\underline{x}\|_{\mathcal{X}}\|\underline{v}\|_{\mathcal{X}}, \quad \forall \underline{x},\underline{v} \in \mathcal{X}$.
2. The bilinear form $\mathcal{C}$ is coercive on the Kernel of $\mathcal{B}$: $\mathcal{C}(\underline{x},\underline{x}) \ge C_1\|\underline{x}\|^2_{\mathcal{X}} \quad \forall \underline{x} \in \mathrm{Ker}\mathcal{B}$, where $C_1 := \min\left\{\frac{1}{2}, \frac{1}{2\mathbb{E}\left[\frac{1}{\kappa_{\min}(\omega)}\right]}\right\}$.
3. The bilinear form $\mathcal{B}$ is bounded: $\sup_{0\neq\underline{x}\in\mathcal{X}} \frac{\mathcal{B}(\underline{x},q)}{\|\underline{x}\|_{\mathcal{X}}} \le C_2\|q\|_{\mathcal{P}} \quad \forall q \in \mathcal{P}$, where $C_2 = \max\left\{1, \sqrt{\mathbb{E}\left[\frac{1}{\kappa_{\min}(\omega)}\right]}\right\}$.
4. The bilinear form $\mathcal{B}$ satisfies the inf-sup condition: $\sup_{0\neq\underline{x}\in\mathcal{X}} \frac{\mathcal{B}(\underline{x},\underline{q})}{\|\underline{x}\|_{\mathcal{X}}} \ge C_3\|\underline{q}\|_{\mathcal{P}}, \quad \forall \underline{q} \in \mathcal{P}$, where $C_3 = \frac{1}{\sqrt{\mathbb{E}[\kappa_{\max}(\omega)]}}$.

Further, the solution $(\underline{x}, p)$ of (43) satisfies the stability estimate

$$\begin{aligned}\|\underline{x}\|_{\mathcal{X}} &\le \frac{1}{C_1}\|\mathcal{F}\|_{\mathcal{X}'} + \frac{2}{\sqrt{C_1}C_3}\|\mathcal{G}\|_{\mathcal{P}'},\\ \|p\|_{\mathcal{P}} &\le \frac{2}{\sqrt{C_1}C_3}\|\mathcal{F}\|_{\mathcal{X}'} + \frac{1}{C_3^2}\|\mathcal{G}\|_{\mathcal{P}'}.\end{aligned} \tag{45}$$

Proof. Let us first show the continuity of $\mathcal{C}$. Being $\mathcal{C}(\cdot,\cdot)$ symmetric, it is sufficient to show that $\mathcal{C}(\underline{x},\underline{x}) \le \|\underline{x}\|^2_{\mathcal{X}}$ which is trivially true since, for $\underline{x} = (y,u)$,

$$\mathcal{C}(\underline{x},\underline{x}) = (y,y)_{L^2,\gamma} + \beta(u,u)_Y \le (y,y)_{L^2,\gamma} + \beta(y,y)_{\mathcal{A}} + \beta(u,u)_Y = (\underline{x},\underline{x})_{\mathcal{X}} = \|\underline{x}\|^2_{\mathcal{X}}.$$

Next, we focus on the coercivity of $\mathcal{C}$ on $\mathrm{Ker}\mathcal{B}$. If $\underline{x} = (y,u) \in \mathrm{Ker}\mathcal{B}$ then $\langle \mathcal{A}y, q\rangle = \langle \mathcal{I}\Lambda_Y u, q\rangle = \mathbb{E}\left[\langle \Lambda_Y u, q_\omega\rangle\right]$ which, choosing $q = y$, implies

$$(y,y)_{\mathcal{A}} \le \|u\|_Y \mathbb{E}\left[\frac{1}{\sqrt{\kappa_{\min}(\omega)}}\|y_\omega\|_{\mathcal{A}_\omega}\right] \le \|u\|_Y\sqrt{\mathbb{E}\left[\frac{1}{\kappa_{\min}(\omega)}\right]}\|y\|_{\mathcal{A}}.$$

Then,

$$\mathcal{C}((y,u),(y,u)) = (y,y)_{L^2,\gamma} + \beta(u,u)_Y \ge (y,y)_{L^2,\gamma} + \frac{\beta}{2}(u,u)_Y + \frac{\beta}{2\mathbb{E}\left[\frac{1}{\kappa_{\min}(\omega)}\right]}(y,y)_{\mathcal{A}} \ge \min\left\{\frac{1}{2}, \frac{1}{2\mathbb{E}\left[\frac{1}{\kappa_{\min}(\omega)}\right]}\right\}((y,u),(y,u))_{\mathcal{X}}.$$

To show the continuity of $\mathcal{B}$, we consider

$$\sup_{0\neq(y,u)\in\mathcal{X}} \frac{\mathcal{B}^2((y,u),q)}{\|(y,u)\|^2_{\mathcal{X}}} = \sup_{0\neq(y,u)\in\mathcal{X}} \frac{\left((y,q)_{\mathcal{A}} - \langle \mathcal{I}\Lambda_Y u, q\rangle\right)^2}{\|(y,u)\|^2_{\mathcal{X}}} = \sup_{0\neq y\in\mathcal{Y}} \frac{(y,q)^2_{\mathcal{A}}}{\|y\|^2_{\mathcal{Y}}} + \sup_{0\neq u\in U} \frac{\left(\mathbb{E}\left[(u,q_\omega)_Y\right]\right)^2}{\|u\|^2_U},$$

where the last equality follows from Zulehner[19, Lemma 2.1]. The second term simplifies to

$$\sup_{0\neq u\in U} \frac{\left(\mathbb{E}\left[(u,q_\omega)\right]\right)^2}{\|u\|^2_U} = \frac{1}{\beta}\sup_{0\neq u\in Y} \frac{\left((u,\mathbb{E}\left[q_\omega\right])_Y\right)^2}{\|u\|^2_Y} = \frac{1}{\beta}\left\|\mathbb{E}\left[q_\omega\right]\right\|^2_Y. \tag{46}$$

Considering the first term,

$$\sup_{0\neq y\in\mathcal{Y}} \frac{(y,q)^2_{\mathcal{A}}}{\|y\|^2_{\mathcal{Y}}} = \sup_{0\neq y\in\mathcal{Y}} \frac{(y,q)^2_{\mathcal{A}}}{(y,y)_{L^2,\gamma} + \beta(y,y)_{\mathcal{A}}} \le \frac{1}{\beta}\sup_{0\neq y\in\mathcal{Y}} \frac{(y,q)^2_{\mathcal{A}}}{(y,y)_{\mathcal{A}}} = \frac{1}{\beta}(q,q)_{\mathcal{A}}. \tag{47}$$

Putting together (46) and (47), using the Cauchy-Schwarz inequality and equivalence between $\|\cdot\|_Y$ and $\|\cdot\|_{A_\omega}$,

$$\sup_{0\neq(y,u)\in\mathcal{X}} \frac{\mathcal{B}^2((y,u),q)}{\|(y,u)\|^2_{\mathcal{X}}} \le \frac{1}{\beta}\left\|\mathbb{E}\left[q_\omega\right]\right\|^2_Y + \frac{1}{\beta}(q,q)_{\mathcal{A}} \le \frac{1}{\beta}\left(\mathbb{E}\left[\frac{1}{\kappa_{\min}(\omega)}\right]\|q\|^2_{\mathcal{A}} + \|q\|^2_{\mathcal{A}}\right) \le \max\left\{1, \mathbb{E}\left[\frac{1}{\kappa_{\min}(\omega)}\right]\right\}\|q\|^2_{\mathcal{P}}.$$

Finally, we deal with the inf-sup condition. Using again Zulehner[19, Lemma 2.1] and choosing $(y,u) = (0,u) \in \mathcal{X}$, we simply obtain the estimate

$$\sup_{0\neq(y,u)\in\mathcal{X}} \frac{\mathcal{B}^2(x,q)}{\|x\|^2_{\mathcal{X}}} = \sup_{0\neq y\in\mathcal{Y}} \frac{(y,q)^2_{\mathcal{A}}}{\|y\|^2_{\mathcal{Y}}} + \sup_{0\neq u\in U} \frac{\left(\langle \mathcal{I}\Lambda_Y u, q\rangle\right)^2}{\|u\|^2_U} \ge \frac{1}{\beta}\sup_{0\neq u\in Y} \frac{\left(\langle \mathcal{I}\Lambda_Y u, q\rangle\right)^2}{\|u\|^2_Y}.$$

As $q \in G^\perp$, we set $u = \mathcal{E}^{-1}q = \Lambda_Y^{-1}\mathcal{I}_{G^0}^{-1}\mathcal{A}q$ and Theorem 4 guarantees that $\|u\|_Y^2 \le \mathbb{E}\left[\kappa_{\max}(\omega)\right]\|q\|_{\mathcal{A}}^2$, so that

$$\sup_{0\neq(y,u)\in\mathcal{X}} \frac{\mathcal{B}^2(x,q)}{\|x\|_{\mathcal{X}}^2} \ge \frac{1}{\beta}\frac{\|q\|_{\mathcal{A}}^4}{\|u\|_Y^2} \ge \frac{1}{\mathbb{E}\left[\kappa_{\max}(\omega)\right]}\left(\frac{1}{\beta}\|q\|_{\mathcal{A}}^2\right) = \frac{1}{\mathbb{E}\left[\kappa_{\max}(\omega)\right]}\|q\|_{\mathcal{P}}^2.$$

Finally, the stability estimate follows from classical theory of saddle-point problems[64, Theorem 4.2.3]. □

We have now all ingredients to apply the operator preconditioning framework and identify the two constants C and α. Let us introduce the space $\mathcal{V} := \mathcal{X}\times\mathcal{P}$, equipped with norm $\|(\underline{x},p)\|_{\mathcal{X}} = \sqrt{\|\underline{x}\|_{\mathcal{V}}^2 + \|p\|_{\mathcal{P}}^2}$, and the operator $\mathscr{T} : \mathcal{V}\to\mathcal{V}'$ defined as

$$\langle \mathscr{T}(\underline{x},p),(\underline{r},q)\rangle_{\mathcal{V}',\mathcal{V}} := C(\underline{x},\underline{r}) + \mathcal{B}(\underline{r},p) + \mathcal{B}(\underline{x},q).$$

On the one hand, from the continuity of the bilinear forms C and $\mathcal{B}$ (Theorem 5), and using the triangle inequality,

$$\|\mathcal{T}\|_{\mathcal{L}(\mathcal{V},\mathcal{V}')} = \sup_{(\underline{x},p)\in\mathcal{V}}\sup_{(\underline{r},q)\in\mathcal{V}} \frac{\left|C(\underline{x},\underline{r}) + \mathcal{B}(\underline{r},p) + \mathcal{B}(\underline{x},q)\right|}{\|(\underline{x},p)\|_{\mathcal{V}}\|(\underline{r},q)\|_{\mathcal{V}}} \le 1 + 2C_2 =: C. \tag{48}$$

On the other hand, using the stability estimate of Theorem 5,

$$\|\mathcal{T}^{-1}\|_{\mathcal{L}(\mathcal{V}',\mathcal{V})} = \sup_{(\mathcal{F},\mathcal{G})\in\mathcal{V}'} \frac{\|\mathcal{T}^{-1}(\mathcal{F},\mathcal{G})\|_{\mathcal{V}}}{\|(\mathcal{F},\mathcal{G})\|_{\mathcal{V}'}} = \frac{\|(\underline{x},p)\|_{\mathcal{V}}}{\|(\mathcal{F},\mathcal{G})\|_{\mathcal{V}'}} \le \sqrt{2}\max\left\{\sqrt{\frac{1}{C_1}+\frac{4}{C_1C_3^2}}, \sqrt{\frac{1}{C_3^4}+\frac{4}{C_1C_3^2}}\right\} =: \alpha. \tag{49}$$

As the two constants C and α are (β,γ)-independent, we conclude that the condition number $\kappa(\mathcal{R}\mathscr{T})$, where $\mathcal{R}$ is the Riesz isomorphism with respect to the scalar product in $\mathcal{V}$, is bounded uniformly with respect to these two parameters. The condition number still depends on the statistical properties of the random field through the constants C_2, C_3 and C_4 which, however, involve only the first moments of $\frac{1}{\kappa_{\min}(\omega)}$, $\frac{1}{\kappa_{\max}(\omega)}$ and $\kappa_{\max}(\omega)$.

6.1 | Mean and Chebyshev semi-iterative approximations

The optimality system (42) involves the non standard trial and test space $G^\perp$. To implement it efficiently, we can rely on the isomorphism $\mathcal{E}$ between Y and $G^\perp$, so that (42) is equivalent to: find $(y,u,p)\in Y\times Y\times Y$ such that $\forall(v,w,r)\in Y\times Y\times Y$

$$\begin{aligned}
\mathbb{E}\left[\langle\mathcal{A}_\omega(\mathcal{E}p)_\omega,(\mathcal{E}v)_\omega\rangle + \langle\Lambda_{L^2}\big((1+\gamma)(\mathcal{E}y)_\omega - \gamma\mathbb{E}\left[(\mathcal{E}y)_\omega\right]\big),(\mathcal{E}v)_\omega\rangle\right] &= \mathbb{E}\left[\langle\Lambda_{L^2}y_d,(\mathcal{E}v)_\omega\rangle\right],\\
\langle\beta\Lambda_Y u - \Lambda_Y\mathbb{E}\left[(\mathcal{E}p)_\omega\right], w\rangle &= 0,\\
\mathbb{E}\left[\langle\mathcal{A}_\omega(\mathcal{E}y)_\omega,(\mathcal{E}r)_\omega\rangle\right] - \mathbb{E}\left[\langle\Lambda_Y u,(\mathcal{E}r)_\omega\rangle\right] &= \mathbb{E}\left[\langle f,(\mathcal{E}r)_\omega\rangle\right].
\end{aligned} \tag{50}$$

A discretization of (50) leads to the discrete system $S_{\mathrm{OP}}\mathbf{x} = E^\top\mathscr{S}E\mathbf{x} = \mathbf{f}$, where $\mathbf{f} = E^\top\mathbf{b}$, $\mathscr{S}$ and $\mathbf{b}$ are given by (14) [3], while

$$E := \begin{pmatrix} A^{-1}Z\mathbb{1}K & & \\ & I_s & \\ & & A^{-1}Z\mathbb{1}K\end{pmatrix},\ \mathbf{x} = \begin{pmatrix}\mathbf{y}\\ \mathbf{u}\\ \mathbf{p}\end{pmatrix}\in\mathbb{R}^{3N_h},$$

The matrix E is the discretization of the isomorphism $\mathcal{E}$. As a preconditioner we use $P_{\mathrm{OP}} = E^\top RE$, where R is the matrix representing the weighted norms defined in (44),

$$R := \begin{pmatrix} M_\gamma + \beta A & & \\ & \beta K & \\ & & \frac{1}{\beta}A\end{pmatrix}\in\mathbb{R}^{(2N+1)\cdot N_h\times(2N+1)\cdot N_h}.$$

A direct calculation leads to

$$P_{\mathrm{OP}} = \begin{pmatrix} R_1 & & \\ & R_2 & \\ & & R_3\end{pmatrix} := \begin{pmatrix} K\mathbb{1}^\top\left(A^{-1}ZM_\gamma ZA^{-1} + \beta ZA^{-1}Z\right)\mathbb{1}K & & \\ & \beta K & \\ & & \frac{1}{\beta}K\mathbb{1}^\top ZA^{-1}Z\mathbb{1}K.\end{pmatrix}. \tag{51}$$

Similarly to Section 5.2, we can approximate the inverse of P_{OP} using a mean approximation of the blocks R_1 and R_3. Replacing formally the matrix A^{-1} with a matrix of equal size with A_0^{-1} on the diagonal, we obtain the mean preconditioner $P_{\mathrm{OPM}} :=$

[3] Replacing the $L^2(D)$ Riesz operator M_s, with the Y Riesz operator K.

TABLE 3 Intervals for the eigenvalues of the preconditioned systems. The number of collocation points is $N = m^4$ for κ_B and $N = m^3$ for κ_L.

β		10^{-2}	10^{-4}	10^{-6}	10^{-8}
$P_{\text{OP}}^{-1}S_{\text{OP}}$	$\kappa_B(x,\omega)$	[-1.12,-0.55] ∪ [0.57,1.55]	[-0.75,-0.36] ∪ [0.69,1.51]	[-0.37,-0.36] ∪ [0.99,1.37]	[-0.37,-0.37] ∪ [1.00,1.37]
$P_{\text{OPM}}^{-1}S_{\text{OP}}$	$\kappa_B(x,\omega)$	[-1.21,-0.56] ∪ [0.58,1.61]	[-0.80,-0.37] ∪ [0.70,1.58]	[-0.40,-0.36] ∪ [1.00,1.41]	[-0.39,-0.36] ∪ [1.00,1.39]
$P_{\text{OPC}}^{-1}S_{\text{OP}}$	$\kappa_B(x,\omega)$	[-1.12,-0.54] ∪ [0.57,1.55]	[-0.75,-0.36] ∪ [0.69,1.51]	[-0.37,-0.36] ∪ [0.99,1.37]	[-0.37,-0.36] ∪ [1.00,1.37]
$P_{\text{OP}}^{-1}S_{\text{OP}}$	$\kappa_L(x,\omega)$	[-1.29,-0.73] ∪ [0.41,1.91]	[-0.84,-0.68] ∪ [0.81,1.81]	[-0.73,-0.68] ∪ [1.03,1.73]	[-0.73,-0.68] ∪ [1.03,1.73]
$P_{\text{OPM}}^{-1}S_{\text{OP}}$	$\kappa_L(x,\omega)$	[-1.95,-0.97] ∪ [0.54,2.79]	[-1.23,-0.88] ∪ [1.57,4.06]	[-1.01,-0.88] ∪ [1.88,4.24]	[-1.01,-0.88] ∪ [1.88,4.25]
$P_{\text{OPC}}^{-1}S_{\text{OP}}$	$\kappa_L(x,\omega)$	[-1.30,-0.73] ∪ [0.42,1.92]	[-0.83,-0.68] ∪ [0.75,1.80]	[-0.73,-0.68] ∪ [0.92,1.73]	[-0.73,-0.68] ∪ [0.92,1.73]

$N_h = 225$, $m = 3$, $\sigma^2 = 0.5$, $\gamma = 0.1$, $L^2 = 0.5$. $N_{\text{it}} = 2$ for both κ_B and κ_L.

σ^2		0.1	0.5	1	1.5
$P_{\text{OP}}^{-1}S_{\text{OP}}$	$\kappa_B(x,\omega)$	[-0.37,-0.37] ∪ [1.00,1.37]	[-0.37,-0.36] ∪ [1.00,1.37]	[-0.37,-0.34] ∪ [1.00,1.37]	[-0.37,-0.32] ∪ [1.00,1.37]
$P_{\text{OPM}}^{-1}S_{\text{OP}}$	$\kappa_B(x,\omega)$	[-0.37,-0.37] ∪ [1.00,1.37]	[-0.39,-0.36] ∪ [1.00,1.39]	[-0.48,-0.36] ∪ [1.00,2.57]	[-0.72,-0.36] ∪ [1.01,8.49]
$P_{\text{OPC}}^{-1}S_{\text{OP}}$	$\kappa_B(x,\omega)$	[-0.37,-0.37] ∪ [1.00,1.37]	[-0.37,-0.36] ∪ [1.00,1.37]	[-0.37,-0.34] ∪ [0.98,1.37]	[-0.37,-0.32] ∪ [0.94,1.37]
$P_{\text{OP}}^{-1}S_{\text{OP}}$	$\kappa_L(x,\omega)$	[-0.64,-0.63] ∪ [1.01,1.64]	[-0.73,-0.68] ∪ [1.03,1.73]	[-0.86,-0.75] ∪ [1.04,1.86]	[-1.00,-0.83] ∪ [1.06,2.00]
$P_{\text{OPM}}^{-1}S_{\text{OP}}$	$\kappa_L(x,\omega)$	[-0.68,-0.66] ∪ [1.23,1.68]	[-1.01,-0.88] ∪ [1.88,4.65]	[-1.59,-1.23] ∪ [2.23,16.66]	[-2.42,-1.69] ∪ [2.69,59.98]
$P_{\text{OPC}}^{-1}S_{\text{OP}}$	$\kappa_L(x,\omega)$	[-0.64,-0.63] ∪ [1.00,1.64]	[-0.73,-0.68] ∪ [0.99,1.73]	[-0.86,-0.75] ∪ [0.94,1.86]	[-1.00,-0.83] ∪ [0.81,2.00]

$N_h = 225$, $m = 3$, $\beta = 10^{-8}$, $\gamma = 0.1$, $L^2 = 0.5$. N_{it} is equal to $2, 2, 4, 4$ for $\kappa_B(x,\omega)$ and equal to $2, 4, 6, 8$ for $\kappa_L(x,\omega)$.

m		2	3	4	5
$P_{\text{OP}}^{-1}S_{\text{OP}}$	$\kappa_B(x,\omega)$	[-0.37,-0.36] ∪ [0.99,1.37]	[-0.37,-0.36] ∪ [0.99,1.37]	[-0.37,-0.36] ∪ [0.99,1.37]	[-0.37,-0.36] ∪ [0.99,1.37]
$P_{\text{OPM}}^{-1}S_{\text{OP}}$	$\kappa_B(x,\omega)$	[-0.40,-0.36] ∪ [1.00,1.41]	[-0.40,-0.36] ∪ [1.00,1.41]	[-0.40,-0.36] ∪ [1.00,1.41]	[-0.40,-0.36] ∪ [1.00,1.41]
$P_{\text{OPC}}^{-1}S_{\text{OP}}$	$\kappa_B(x,\omega)$	[-0.37,-0.36] ∪ [0.99,1.37]	[-0.37,-0.36] ∪ [0.99,1.37]	[-0.37,-0.36] ∪ [0.99,1.37]	[-0.37,-0.36] ∪ [0.99,1.37]
$P_{\text{OP}}^{-1}S_{\text{OP}}$	$\kappa_L(x,\omega)$	[-0.73,-0.68] ∪ [1.02,1.73]	[-0.73,-0.68] ∪ [1.03,1.73]	[-0.73,-0.68] ∪ [1.03,1.73]	[-0.73,-0.68] ∪ [1.03,1.73]
$P_{\text{OPM}}^{-1}S_{\text{OP}}$	$\kappa_L(x,\omega)$	[-0.99,-0.86] ∪ [1.86,3.68]	[-1.01,-0.88] ∪ [1.88,4.24]	[-1.01,-0.88] ∪ [1.88,4.31]	[-1.01,-0.88] ∪ [1.88,4.32]
$P_{\text{OPC}}^{-1}S_{\text{OP}}$	$\kappa_L(x,\omega)$	[-0.73,-0.68] ∪ [0.94,1.73]	[-0.73,-0.68] ∪ [0.92,1.73]	[-0.73,-0.68] ∪ [0.91,1.73]	[-0.73,-0.68] ∪ [0.91,1.73]

$N_h = 225$, $\sigma^2 = 0.5$, $\beta = 10^{-6}$, $\gamma = 0.1$, $L^2 = 0.5$, $N_{\text{it}} = 2$ for both $\kappa_B(x,\omega)$ and $\kappa_L(x,\omega)$.

$E^\top R_M E$,

$$P_{\text{OPM}}^{-1} = \begin{pmatrix} R_{1,M}^{-1} & & \\ & R_{2,M}^{-1} & \\ & & R_{3,M}^{-1} \end{pmatrix} := \begin{pmatrix} K^{-1}A_0\left(M_s + \beta A_0\right)^{-1}A_0K^{-1} & & \\ & \frac{1}{\beta}K^{-1} & \\ & & \beta K^{-1}A_0K^{-1} \end{pmatrix}. \tag{52}$$

If the variance is large, we use $R_{1,M}^{-1}$ and $R_{3,M}^{-1}$ as preconditioners inside a Chebyshev Semi-Iterative method to invert R_1 and R_3. The two Chebyshev Semi-Iterative method can be executed separately and in parallel. To choose the parameters α, $\underline{\lambda}$ and $\overline{\lambda}$, we rely on the following Lemma, obtained using the same argument of Lemma 8.

Lemma 10. The spectra of $R_{1,M}^{-1}R_1$ and of $R_{3,M}^{-1}R_3$ are real and bounded from below by 1.

Notice that, on the one hand, the solution of $S_{\text{OP}}\mathbf{x} = \mathbf{f}$ using a Krylov method requires the matrix-vector multiplication between S_{OP} and a vector, that involves the computation of the action of the inverse of A on four vectors. The action of the inverse of A must be computed exactly, or up to a very low tolerance. On the other hand, S_{OP} is a matrix of dimension $3N_h \lll (2N+1)N_h$, the latter being the size of $\mathscr{S}$. Thus, a Krylov method is less prone to saturation of memory and instability due to orthogonalization. According to the software, architecture and problem at hand, the pros could be larger than the cons, or viceversa.
Tables 3 report the negative and positive intervals containing the spectrum of the system preconditioned by P_{OP}, P_{OPM} or P_{OPC}. All preconditioners exhibits a β-robust spectrum, and in particular the mean preconditioner P_{OPM}^{-1} performs quite better than the algebraic one S_{LRM} (see Table 2). The dependence of P_{OP} on σ^2 is weak, and similar to that of $\mathcal{P}^{-1}$, analysed in Section 5.1, as Theorem 5 involves the first moments of $1/\kappa_{\min}(\omega)$ and $\kappa_{\max}(\omega)$. Finally Table 3 shows that all preconditioners are robust with respect to the number of collocation points.

7 | NUMERICAL EXPERIMENTS

The aim of this section is to further validate the theoretical results presented in Section 5 and 6, and to compare the preconditioners analysed on a model problem. We consider the domain $D = (0,1)^2$ discretized with a regular mesh of size h, and a finite element approximation using $\mathbb{P}_1$ finite elements. For each preconditioner $\widetilde{P}$, P_{LRM}, P_{LRC}, P_{OPM} and P_{OPC}, we report the number of iterations and computational times in seconds to solve the saddle point system using preconditioned MINRES . Although it is tempting to compare the computational times and number of iterations among all preconditioners, we stress that P_{OPM} and

P_{OPC} compute a different optimal control with respect to $\widetilde{P}$, P_{LRM} and P_{LRC}, as the control belongs to Y and acts on the state equation through the Riesz map of Y (see Elvetun et al.[32] for an instance of application arising in electrocardiography). In all experiments, the matrix A is inverted approximately using the Fortran Algebraic MultiGrid (AMG) library HSL_MI20[65], which is called using the Matlab interface. We specifically used two V-cycles with one iteration of the damped Jacobi smoother with parameter $\theta = \frac{8}{9}$ and 5 levels. All other parameters are left to default values. The inverse of C is computed approximately, inverting the mass matrix M with 25 iterations of the Chebyshev semi-iterative method using as preconditioner the diagonal of M itself. The damping parameters α as well as $\underline{\lambda}$ and $\overline{\lambda}$ are estimated once and for all using the mass matrix M_s. The application of A^{-1} and of M^{-1} onto a vector is performed in parallel, using the Matlab Parallel Computing Toolbox. Further, we compute once for all the LU decomposition of $A_0 + \frac{1}{\sqrt{\beta}} M_s$, $\beta A_0 + M_s$ and K, which is feasible as their size corresponds to a single PDE discretization. Clearly, one could further approximate them using AMG, if β is not too small, or using other iterative methods. Finally, when using P_{OPM} and P_{OPC}, we compute the exact action of A^{-1} using eight iterations of the conjugate gradient method preconditioned by AMG, which are enough to have a (unpreconditioned) residual of approximately 10^{-11}. MINRES is stopped when the relative (unpreconditioned) residual is smaller than $\text{Tol} = 10^{-6}$. The simulations have been performed on a workstation equipped with an Intel® Core™ i9-10900X and 32 GB of RAM. For reproducibility, data and codes are available at Nobile et al.[66].

7.1 | Bounded random field with Stochastic Collocation

We consider the bounded random field defined in (29) and use a full tensorized Gauss-Legendre quadrature formula with 5 points for each random variable $\xi_j(\omega)$, $j = 1, \dots, 4$, ($N = 5^4 = 625$). The mesh size is $h = 2^{-5}$ and $N_h = 961$. The global system has approximately 1.2 million degrees of freedom. The target state is $y_d = \sin(\pi x)\sin(\pi y)$.

First we consider $u \in L^2(D)$. The results in Table 4 confirm that $\widetilde{P}$ is extremely efficient when β is sufficiently large, but its performance deteriorates when $\beta \to 0$ as Theorem 2 predicts. P_{LRM} performs well unless for extremely small values of β (e.g. $\beta \approx 10^{-8}$), as remarked in Table 2. P_{LRC} recovers robustness, at the price of additional Chebyshev semi-iterations, and leads to constant numbers of iterations and computational times as $\beta \to 0$.

Next, we show how the preconditioners behave as σ^2 increases. We set $\beta = 10^{-2}$ for $\widetilde{P}$, while $\beta = 10^{-6}$ for P_{LRM} and P_{LRC}, that is, we set β according to the regime where we would use the preconditioners in practice. $\widetilde{P}$ exhibits a very weak dependence on σ^2. This is reflected both by the estimates of Theorem 2 and by Table 1. Recall that $\kappa_B(x, \omega) \geq 1$ for a.e. ω, so that $\widehat{\mathbb{E}}\left[\frac{1}{\kappa^2_{\min}(\omega)}\right]$ is bounded as σ^2 grows. P_{LRM} becomes inefficient when σ^2 grows, since the mean matrix A_0^{-1} is a crude approximation of $\mathbb{1}^{\mathsf{T}} Z A^{-1} Z \mathbb{1}$ that does not take into account the variability of the stiffness matrices[27]. The addition of the Chebyshev semi-iteration helps to reduce the computational time and number of iterations, but does not remove completely the dependence over σ^2.

TABLE 4 Number of iterations and computational time in seconds to reach a relative residual smaller than 10^{-6}.

β	10^{-2}	10^{-4}	10^{-6}	10^{-8}
$\widetilde{P}^{-1}\mathscr{S}$	29 (20.3)	37 (25.0)	109 (72.3)	734 (479.3)
$P_{\text{LRM}}^{-1}\mathscr{S}$	31 (42.7)	33 (44.6)	37 (49.7)	169 (221.4)
$P_{\text{LRC}}^{-1}\mathscr{S}$	31 (89.5)	33 (94.5)	31 (88.2)	31 (87.8)

$\sigma^2 = 0.5$, $\gamma = 10^{-1}$, $N_{\text{it}} = 2$.

	β \ σ^2	0.1	0.5	1	1.5
$\widetilde{P}^{-1}\mathscr{S}$	10^{-2}	29 (20.6)	29 (19.4)	31 (20.8)	33 (21.9)
$P_{\text{LRM}}^{-1}\mathscr{S}$	10^{-6}	27 (36.8)	37 (48.9)	87 (112.6)	198 (254.5)
$P_{\text{LRC}}^{-1}\mathscr{S}$	10^{-6}	27 (76.4)	31 (86.5)	35 (132.9)	39 (148.6)

$\gamma = 10^{-1}$. $N_{\text{it}} = 2$ for $\sigma^2 \in \{0.1, 0.5\}$ and $N_{\text{it}} = 4$ for $\sigma^2 \in \{1, 1.5\}$.

Next, we consider a control $u \in H^1(D)$ and the operator preconditioning approach. Table 5 shows that both P_{OPM} and P_{OPC} are very robust with respect to β. Interestingly, P_{OPM} performs well also for $\beta \approx 10^{-8}$ in contrast with P_{LRM}. P_{OPC} still exhibits a

TABLE 5 Number of iterations and computational time in seconds to reach a relative residual smaller than 10^{-6}.

β	10^{-2}	10^{-4}	10^{-6}	10^{-8}
$P_{\text{OPM}}^{-1}\mathscr{S}$	27 (50.8)	43 (76.8)	19 (35.2)	19 (35.4)
$P_{\text{OPC}}^{-1}\mathscr{S}$	28 (140.4)	46 (226.8)	28 (139.8)	28 (134.8)

$\sigma^2 = 0.5$, $\gamma = 10^{-1}$, $N_{\text{it}} = 2$.

σ^2	0.1	0.5	1	1.5
$P_{\text{OPM}}^{-1}\mathscr{S}$	12 (23.5)	19 (35.2)	37 (65.6)	92 (159.5)
$P_{\text{OPC}}^{-1}\mathscr{S}$	27 (134.0)	31 (138.6)	35 (216.9)	39 (236.7)

$\beta = 10^{-6}$ and $\gamma = 10^{-1}$. $N_{\text{it}} = 2$ for $\sigma^2 \in \{0.1, 0.5\}$ and $N_{\text{it}} = 4$ for $\sigma^2 \in \{1, 1.5\}$.

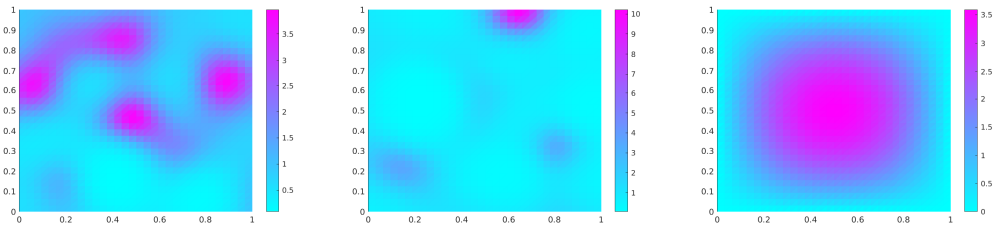

FIGURE 2 The left and center panels show two random realizations of $\kappa_L(x,\omega)$. The right panel shows the optimal control to reach the target state $y_d = \sin(\pi y)\sin(\pi x)$. Parameters: $L^2 = 0.025$, $\sigma^2 = 0.5$, $\beta = 10^{-2}$, $\gamma = 0.1$ and $N = 10^4$.

σ^2 dependence as expected, since the estimates of Theorem 5 involve the first moments of $\frac{1}{\kappa_{\min}(\omega)}$ and $\kappa_{\max}(\omega)$. Notice that, even though P_{OPC} is more robust than P_{OPM} in terms of number of iterations for increasing σ^2, the additional cost due to the inner semi-Chebyshev method leads actually to higher computational times.

To conclude, for the bounded random field considered and for a control in $L^2(D)$, $\widetilde{P}$ is the most efficient preconditioner for a large β, and it is also robust again increasing values of σ^2. P_{LRM} is an efficient alternative if $\beta \geq 10^{-6}$. For either extremely small values of β (e.g., 10^{-8}) or for large values of σ^2, P_{OPC} is instead to be preferred. Notice that P_{OPC} is robust both in terms of number of iterations and computational times with respect to β (for a fixed σ^2).

For a control in $H^1(D)$, P_{OPM} outperforms P_{OPC} in all tests in terms of computational times.

7.2 | Log-normal field with Monte Carlo sampling

In this subsection, we consider the log-normal field $\kappa_L(x,\omega)$ defined in (28) with covariance function $Cov_g(x,y) = \sigma^2 \exp\left(-\frac{\|x-y\|_2^2}{L^2}\right)$. We consider a relatively small correlation length, setting $L^2 = 0.025$. Fig. 2 shows two random realizations of $\kappa_L(x,\omega)$. To keep 99% of the variance, we retain $M = 37$ components in the Karhunen-Loève expansion (28), so that SCM on tensor grids is not feasible due to the curse of dimensionality. We thus rely on a standard Monte Carlo with $N = 10^4$ samples The saddle point system involves approximately 19.2 millions degrees of freedom, and we first consider a control $u \in L^2(D)$. Table 6 reports the number of iterations and computational times in seconds for different values of β and σ^2. Notice that the performance of both $\widetilde{P}$ and P_{LRM} deteriorates quickly as $\beta \to 0$. In constrat, P_{LRC} exhibits a weak dependence on β, but still remains quite efficient for the broad range of values. The performance of all preconditioners deteriorates when σ^2 increases. We remark that $\sigma^2 = 1.5$ is quite a challenging setting: in our experiments we had $\max_{1\leq i\leq N} \frac{\kappa_{\max}(\omega_i)}{\kappa_{\min}(\omega_i)} = 1.08e4$, that is the random diffusion field can vary up to four order of magnitude inside the domain (the expected variation is $\widehat{\mathbb{E}}\left[\frac{\kappa_{\max}(\omega)}{\kappa_{\min}(\omega)}\right] = 396.39$).

Finally, we look for a $u \in H^1(D)$. Tables 7 further confirm that both P_{OPM} and P_{OPC} lead to a $\beta-$robust convergence. The latter is again not σ^2-robust as the theory predicts, but the increase of the number of iterations is modest. Nevertheless, an increasing number of inner iterations is nedded as the mean approximations $R_{1,M}$ and $R_{3,M}$ lose their efficacy as preconditioners inside the Chebyshev semi-iterative method, and this results in a significant increase of computational times.

TABLE 6 Number of iterations and computational time in seconds to reach a relative residual smaller than 10^{-6}.

β	10^{-2}	10^{-4}	10^{-6}	10^{-8}
$\widetilde{P}^{-1}\mathscr{S}$	37 (390.0)	55 (564.7)	265 (2699.6)	>800 (>8249.9)
$P_{\text{LRM}}^{-1}\mathscr{S}$	37 (897.4)	53 (1282.1)	422 (10170.6)	>800 (>19583.1)
$P_{\text{LRC}}^{-1}\mathscr{S}$	37 (2540.1)	37 (2577.2)	41 (2887.30)	49 (3526.7)

$\sigma^2 = 0.5$, $\gamma = 10^{-1}$, $N = 10^4$, $N_{\text{it}} = 4$.

	β \ σ^2	0.1	0.5	1	1.5
$\widetilde{P}^{-1}\mathscr{S}$	10^{-2}	31 (335.8)	37 (468.4)	43 (550.5)	49 (576.0)
$P_{\text{LRM}}^{-1}\mathscr{S}$	10^{-6}	53 (1372.3)	442 (13408.8)	796 (23978.0)	760 (22121.2)
$P_{\text{LRC}}^{-1}\mathscr{S}$	10^{-6}	31 (1739.1)	41 (3724.3)	59 (8267.9)	79 (11635.6)

N_{it} equal to $\{2, 4, 8, 10\}$ for σ^2 equal respectively to $\{0.1, 0.5, 1, 1.5\}$; $\gamma = 0.1$.

TABLE 7 Number of iterations and computational time in seconds to reach a relative residual smaller than 10^{-6}.

β	10^{-2}	10^{-4}	10^{-6}	10^{-8}
$P_{\text{OPM}}^{-1}\mathscr{S}$	39 (1190.8)	44 (1340.1)	31 (961.2)	32 (990.0)
$P_{\text{OPC}}^{-1}\mathscr{S}$	37 (4937.6)	38 (5074.6)	29 (3896.6)	31 (4178.0)

$\sigma^2 = 0.5$, $\gamma = 10^{-1}$, $N_{\text{it}} = 4$.

σ^2	0.1	0.5	1	1.5
$P_{\text{OPM}}^{-1}\mathscr{S}$	16 (518.2)	31 (980.5)	68 (2103.3)	145 (4534.6)
$P_{\text{OPC}}^{-1}\mathscr{S}$	26 (2420.8)	29 (3922.4)	31 (6869.8)	39 (10438.2)

N_{it} equal to $\{2, 4, 8, 10\}$ for σ^2 equal respectively to $\{0.1, 0.5, 1, 1.5\}$; $\beta = 10^{-6}$ and $\gamma = 10^{-1}$.

To summarize, for a control in $L^2(D)$, $\widetilde{P}$ is again the most efficient preconditioner for β large enough. The regime for β small is more challenging than in subsection 7.1. P_{LRC} is always to be preferred over P_{LRM} except for very small values of σ^2 (i.e. $\sigma^2 = 0.1$). Nevertheless, we remark that the computational times of P_{LRC} show a stronger dependence on σ^2 compared to subsection 7.1 since the number of semi-Chebyshev iterations grows faster as σ^2 increases. For a control in $H^1(D)$, P_{OPM} outperforms again P_{OPC} in all tests.

Notice that, according to the numerical tests, the mean approximation is more effective in the operator preconditioning approach than in the matching Schur complement approximation. We believe this is due to the persistent (weak) dependence of the exact matching Schur approximation on β (see Table 2), which seems to amplify the approximation error of $\mathbb{1}^\top Z A^{-1} Z \mathbb{1}$ with A_0^{-1} as $\beta \to 0$.

The development of improved preconditioners which capture better the effective spectrum of $\mathbb{1}^\top Z A^{-1} Z \mathbb{1}$, are expected to reduce the number of inner iterations needed, or even to replace directly the mean approximation A_0 into (32) and (52), leading to improved S_{LRM} and P_{OPM}, and thus reducing the overall computational time. A possibility would be to subsample A_ω^{-1} and to replace the subsamples with cheap approximation using, e.g., sparse approximate inverse method[67].

8 | CONCLUSION

In this manuscript, we studied preconditioners for the large saddle point systems which arise in the context of quadratic robust OCPUU. Our theoretical analysis casts light on the dependence of these preconditioners on the regularization parameter β, on the variance σ^2 of the random field, and on the level of discretization in the probability space. For large values of β, the coupled saddle point system can be efficiently solved by preconditioning separately and in parallel all the state and adjoint equations. For small values of β, robustness can be recovered using two different preconditioners which require the additional solution of a linear system (whose size is equal to a single PDE discretization) which couples all the equations and involves the sum of the inverses of the stiffness matrices. We solved such reduced system using a mean approximation or a preconditioned Chebyshev

semi-iterative method. Our theoretical analysis characterizes the dependence of the preconditioners on the variance of the random field through either the first or second moment of $1/\kappa_{\min}(\omega)$ or $\kappa_{\max}(\omega)$. The weak dependence for physically relevant ranges of σ^2 is confirmed by our numerical experiments in terms of number of iterations, but not necessarily in terms of computational times, as one needs to increase the number of inner Chebyshev semi-iterations for large values of σ^2. Hence, the combination of small values of β and large values of σ^2 is still challenging, and the development of tailored preconditioners for the reduced system involving the sum of the inverses of the stiffness matrices is expected to close the gap between the theoretical results and practical implementations.

9 | CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

References

1. Lions JL. Optimal control of systems governed by partial differential equations. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer-Verlag; 1971.

2. Hinze M, Pinnau R, Ulbrich M, and Ulbrich S. Optimization with PDE constraints. vol. 23. Springer Science & Business Media; 2008.

3. Tröltzsch F. Optimal control of partial differential equations: theory, methods, and applications. Graduate studies in mathematics. American Mathematical Society; 2010.

4. Shapiro A, Dentcheva D, and Ruszczyński A. Lectures on stochastic programming: modeling and theory. SIAM; 2014.

5. Kouri DP, and Surowiec TM. Existence and optimality conditions for risk-averse PDE-constrained optimization. SIAM/ASA Journal on Uncertainty Quantification. 2018;**6**(2):787–815.

6. Kouri DP, and Surowiec TM. Risk-averse PDE-constrained optimization using the conditional Value-At-Risk. SIAM Journal on Optimization. 2016;**26**(1):365–396.

7. Martin M, Krumscheid S, and Nobile F. Complexity analysis of stochastic gradient methods for PDE-constrained optimal control problems with uncertain parameters. ESAIM: Mathematical Modelling and Numerical Analysis. 2021;**55**(4):1599–1633.

8. Martin M, and Nobile F. PDE-Constrained optimal control problems with uncertain parameters using SAGA. SIAM/ASA Journal on Uncertainty Quantification. 2021;**9**(3):979–1012.

9. Geiersbach C, and Wollner W. A stochastic gradient method with mesh refinement for PDE-constrained optimization under uncertainty. SIAM Journal on Scientific Computing. 2020;**42**(5):A2750–A2772.

10. Ayoul-Guilmard Q, Ganesh S, Nobile F, Rossi R, and Soriano C. D6.3 Report on stochastic optimisation for simple problems. Open Access Repository of the ExaQUte project: Deliverables. 2021;Available from: URLhttps://www.scipedia.com/public/Ayoul-Guilmard_et_al_2021f.

11. Guth PA, Kaarnioja V, Kuo F, Schillings C, and Sloan IH. A Quasi-Monte Carlo method for optimal control under Uncertainty. SIAM/ASA Journal on Uncertainty Quantification. 2021;**9**(2):354–383.

12. Van Barel A, and Vandewalle S. Robust optimization of PDEs with random coefficients using a multilevel Monte Carlo method. SIAM/ASA Journal on Uncertainty Quantification. 2019;**7**(1):174–202.

13. Kouri DP, Heinkenschloss M, Ridzal D, and van Bloemen Waanders BG. A Trust-Region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty. SIAM Journal on Scientific Computing. 2013;**35**(4):A1847–A1879.

14. Schillings C, and Schwab C. Sparse, adaptive Smolyak quadratures for Bayesian inverse problems. Inverse Problems. 2013;**29**(6):065011.

15. Kouri DP, and Ridzal D. In: Inexact Trust-Region methods for PDE-constrained optimization. New York, NY: Springer New York; 2018. p. 83–121.

16. Rees T, Stoll M, and Wathen A. All-at-once preconditioning in PDE-constrained optimization. Kybernetika. 2010;**46**(2):341–360.

17. Rees T, Dollar HS, and Wathen A. Optimal solvers for PDE-constrained optimization. SIAM Journal on Scientific Computing. 2010;**32**(1):271–298.

18. Pearson JW, and Wathen A. A new approximation of the Schur complement in preconditioners for PDE-constrained optimization. Numerical Linear Algebra with Applications. 2012;**19**(5):816–829.

19. Zulehner W. Nonstandard norms and robust estimates for saddle point problems. SIAM Journal on Matrix Analysis and Applications. 2011;**32**(2):536–560.

20. Mardal KA, and Winther R. Preconditioning discretizations of systems of partial differential equations. Numerical Linear Algebra with Applications. 2011;**18**(1):1–40.

21. Borzì A, and von Winckel G. Multigrid methods and sparse-grid collocation techniques for parabolic optimal control problems with random coefficients. SIAM Journal on Scientific Computing. 2009;**31**(3):2172–2192.

22. Borzì A. Multigrid and sparse-grid schemes for elliptic control problems with random coefficients. Computing and visualization in science. 2010;**13**(4):153–160.

23. Ghanem RG, and Kruger RM. Numerical solution of spectral stochastic finite element systems. Computer Methods in Applied Mechanics and Engineering. 1996;**129**(3):289–303.

24. Pellissetti MF, and Ghanem RG. Iterative solution of systems of linear equations arising in the context of stochastic finite elements. Advances in Engineering Software. 2000;**31**(8):607–616.

25. Benner P, Onwunta A, and Stoll M. Block-diagonal preconditioning for optimal control problems constrained by PDEs with uncertain inputs. SIAM Journal on Matrix Analysis and Applications. 2016;**37**(2):491–518.

26. Rosseel E, and Wells GN. Optimal control with stochastic PDE constraints and uncertain controls. Computer Methods in Applied Mechanics and Engineering. 2012;**213**:152–167.

27. Powell CE, and Elman HC. Block-diagonal preconditioning for spectral stochastic finite-element systems. IMA Journal of Numerical Analysis. 2009;**29**(2):350–375.

28. Kouri DP. A multilevel stochastic collocation algorithm for optimization of PDEs with uncertain coefficients. SIAM/ASA Journal on Uncertainty Quantification. 2014;**2**(1):55–81.

29. Van Barel A, and Vandewalle S. MG/OPT and MLMC for robust optimization of PDEs. arXiv preprint arXiv:200601231. 2020;.

30. Zahr MJ, Carlberg KT, and Kouri DP. An efficient, globally convergent method for optimization under uncertainty using adaptive model reduction and sparse grids. SIAM/ASA Journal on Uncertainty Quantification. 2019;**7**(3):877–912.

31. Smith RC. Uncertainty quantification: theory, implementation, and applications. Computational Science and Engineering. SIAM; 2013.

32. Elvetun OL, and Nielsen BF. PDE-constrained optimization with local control and boundary observations: robust preconditioners. SIAM Journal on Scientific Computing. 2016;**38**(6):A3461–A3491.

33. Heidel G, and Wathen A. Preconditioning for boundary control problems in incompressible fluid dynamics. Numerical Linear Algebra with Applications. 2019;**26**(1):e2218.

34. Ullmann E. A Kronecker product preconditioner for stochastic Galerkin finite element discretizations. SIAM Journal on Scientific Computing. 2010;**32**(2):923–946.

35. Murphy MF, Golub GH, and Wathen A. A note on preconditioning for indefinite linear systems. SIAM Journal on Scientific Computing. 2000;**21**(6):1969–1972.

36. Kouri DP, Ridzal D, and Tuminaro R. KKT preconditioners for PDE-constrained optimization with the Helmholtz equation. SIAM Journal on Scientific Computing. 2020;**0**(0):S225–S248.

37. Liu J, and Pearson JW. Parameter-robust preconditioning for the optimal control of the wave equation. Numerical Algorithms. 2020;**83**(3):1171–1203.

38. Pearson JW, and Wathen A. Matching Schur complement approximations for certain saddle-point systems. In: Contemporary Computational Mathematics-A Celebration of the 80th Birthday of Ian Sloan. Springer; 2018. p. 1001–1016.

39. Malek J, and Strakos Z. Preconditioning and the conjugate gradient method in the context of solving PDEs. SIAM Spotlights. SIAM, Society for Industrial and Applied Mathematics; 2014.

40. Kirby RC. From functional analysis to iterative methods. SIAM Review. 2010;**52**(2):269–293.

41. Khan A, Powell CE, and Silvester DJ. Robust preconditioning for stochastic Galerkin formulations of parameter-dependent nearly incompressible elasticity equations. SIAM Journal on Scientific Computing. 2019;**41**(1):A402–A421.

42. Gittelson CJ. Stochastic Galerkin discretization of the log-normal isotropic diffusion problem. Mathematical Models and Methods in Applied Sciences. 2010;**20**(02):237–263.

43. Schwab C, and Gittelson CJ. Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs. Acta Numerica. 2011;**20**:291–467.

44. Cohn DL. Measure theory: second edition. Birkhäuser Advanced Texts Basler Lehrbücher. Springer New York; 2013.

45. Charrier J. Strong and weak error estimates for elliptic partial differential equations with random coefficients. SIAM Journal on Numerical Analysis. 2012;**50**(1):216–246.

46. Cheng AHD, and Bear J. Modeling Groundwater Flow and Contaminant Transport. Springer Publishing Company, Incorporated; 2016.

47. McLaughlin D, and Townley LR. A reassessment of the groundwater inverse problem. Water Resources Research. 1996;**32**(5):1131–1161.

48. Charrier J, Scheichl R, and Teckentrup AL. Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods. SIAM Journal on Numerical Analysis. 2013;**51**(1):322–352.

49. Bonizzoni F, and Nobile F. Perturbation analysis for the Darcy problem with log-normal permeability. SIAM/ASA Journal on Uncertainty Quantification. 2014;**2**(1):223–244.

50. Charrier J. Analyse numérique d'équations aux dérivées aléatoires, applications à l'hydrogéologie (Theses). École normale supérieure de Cachan - ENS Cachan; 2011. Available from: https://tel.archives-ouvertes.fr/tel-00625092.

51. Lord GJ, Powell CE, and Shardlow T. An introduction to computational stochastic PDEs. Cambridge Texts in Applied Mathematics. Cambridge University Press; 2014.

52. Ciarlet PG. Linear and nonlinear functional analysis with applications. Applied mathematics. Philadelphia, PA: SIAM; 2013.

53. Martínez-Frutos J, and Esparza F. Optimal control of PDEs under uncertainty: an introduction with application to optimal shape design of structures. Springer; 2018.

54. Ayoul-Guilmard Q, Ganesh S, Nobile F, Rossi R, and Soriano C. D6.2: Report on the calculation of stochastic sensitivities. Open Access Repository of the ExaQUte project: Deliverables. 2021;Available from: https://www.scipedia.com/public/table_Soriano_2019d.

55. Elman H, Silvester D, and Wathen A. Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics. Oxford University Press; 2014.

56. Quarteroni A, and Valli A. Numerical approximation of partial differential equations. Springer Series in Computational Mathematics. Springer Berlin Heidelberg; 2009.

57. Ern A, and Guermond JL. Theory and practice of finite elements. vol. 159. Springer; 2004.

58. Pearson JW, and Gondzio J. Fast interior point solution of quadratic programming problems arising from PDE-constrained optimization. Numerische Mathematik. 2017;**137**(4):959–999.

59. Gordon AD, and Powell CE. On solving stochastic collocation systems with algebraic multigrid. IMA Journal of Numerical Analysis. 2012;**32**(3):1051–1070.

60. Wathen AJ, and Rees T. Chebyshev semi-iteration in preconditioning. ETNA. 2008;.

61. Golub GH, and Van Loan CF. Matrix computations. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press; 2013.

62. Whittle P. A multivariate generalization of Chebyshev's inequality. Quart J Math Oxford Ser. 1958;**2**:232–240.

63. Schöberl J, and Zulehner W. Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimization problems. SIAM Journal on Matrix Analysis and Applications. 2007;**29**(3):752–773.

64. Boffi D, Brezzi F, and Fortin M. Mixed Finite Element Methods and Applications. Springer Series in Computational Mathematics. Springer Berlin Heidelberg; 2013.

65. Boyle J, Mihajlović M, and Scott J. HSL_MI20: An efficient AMG preconditioner for finite element problems in 3D. International Journal for Numerical Methods in Engineering;**82**(1):64–98.

66. Nobile F, and Vanzan T. Preconditioners for robust optimal control problems under uncertainty - numerical tests;. Available from: https://zenodo.org/record/7040795.

67. Grote MJ, and Huckle T. Parallel preconditioning with sparse approximate inverses. SIAM Journal on Scientific Computing. 1997;**18**(3):838–853.

□