SIAM J. CONTROL OPTIM.
Vol. 59, No. 2, pp. 1161–1187

RISK-AVERSE CONTROL OF FRACTIONAL DIFFUSION WITH UNCERTAIN EXPONENT*

HARBIR ANTIL†, DREW P. KOURI‡, AND JOHANNES PFEFFERER§

Abstract. In this paper, we introduce and analyze a new class of optimal control problems constrained by elliptic equations with uncertain fractional exponents. We utilize risk measures to formulate the resulting optimization problem. We develop a functional analytic framework, study the existence of solution, and rigorously derive the first-order optimality conditions. Additionally, we employ a sample-based approximation for the uncertain exponent and the finite element method to discretize in space. We prove the rate of convergence for the optimal risk neutral controls when using quadrature approximation for the uncertain exponent and conclude with illustrative examples.

Key words. optimal control, risk measures, fractional diffusion, uncertain fractional exponent, finite element method, error estimates

AMS subject classifications. 35J75, 65D05, 26A33, 49J20, 49M25, 65M12, 65M15, 65M60

DOI. 10.1137/20M1324958

1. Introduction.

In the recent article [58], the authors demonstrate a direct qualitative correlation between geophysical electromagnetic data and numerical simulations of the fractional Helmholtz equation. This application, as well as many others (see, e.g., [13, 43]), motivate our study of the fractional diffusion equation

$$\mathcal{L}^s \mathsf{u} = \mathsf{z} \text{ in } \Omega, \qquad \mathsf{u} = 0 \text{ on } \partial\Omega, \tag{1.1}$$

where Ω is an open and bounded domain in $\mathbb{R}^N$ $(N \geq 1)$, with Lipschitz boundary $\partial\Omega$. The operator $\mathcal{L}^s$, $s \in (0, 1)$, is a fractional power of the second-order, symmetric, and uniformly elliptic differential operator $\mathcal{L}$, supplemented with homogeneous Dirichlet boundary conditions. For example, $\mathcal{L}$ represents the second-order differential operator $\mathcal{L}w = -\mathrm{div}_x(A\nabla_x w) + cw$, where $0 \leq c \in L^\infty(\Omega)$ and $A \in C^{0,1}(\Omega, \mathsf{GL}(N, \mathbb{R}))$ is symmetric and positive definite on Ω. Here, $\mathsf{GL}(N, \mathbb{R})$ denotes the general linear group consisting of $N \times N$ invertible matrices on $\mathbb{R}$ endowed with the operation of matrix multiplication.

*Received by the editors March 13, 2020; accepted for publication (in revised form) January 8, 2021; published electronically March 29, 2021. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.
https://doi.org/10.1137/20M1324958
Funding: The first author's research was supported in part by NSF grants DMS-1818772 and DMS-1913004, and by the Air Force Office of Scientific Research under award FA9550-19-1-0036. The second author's research was supported by DARPA grant SNL 01450709 and by the Air Force Office of Scientific Research under award F4FGA09135G001. The third author's research was supported in part by NSF grant DMS-1521590.
†Department of Mathematical Sciences and The Center for Mathematics and Artificial Intelligence, George Mason University, Fairfax, VA 22030 USA (hantil@gmu.edu).
‡Optimization and Uncertainty Quantification, Sandia National Laboratories, Albuquerque, NM 87185-1320 USA (dpkouri@sandia.gov).
§Chair of Optimal Control, Center of Mathematical Sciences, Technical University of Munich, 85748 Garching by Munich, Munich, Germany (pfefferer@ma.tum.de).

Identification of the right-hand side, z, is a classical problem in optimal control and inverse problems. One motivating application is the distributed control of the fractional Helmholtz equation, which as shown in [58] can accurately model certain geophysical phenomena. Although (1.1) is simpler than the fractional Helmholtz equation, we consider the optimal control of (1.1) a natural first step. However, one difficulty when considering the fractional model (1.1) is that the exponent s is often unknown. Recently, there have been systematic studies to determine s for models of type (1.1) (see, e.g., [4, 51], as well as [7]), where s is allowed to be spatially varying. These works consider deterministic versions of (1.1), whereas the authors in [31] consider stochastic z.

In this work, we model the fractional power s of $\mathcal{L}$ as a random parameter and consider the risk-averse optimal control problem

$$\min_{\mathsf{z}\in\mathsf{Z}_{\mathrm{ad}}} \mathcal{R}(J(\cdot,\mathsf{S}(\mathsf{z}),\mathsf{z})), \tag{1.2a}$$

where $\mathsf{S}(z) = \mathsf{u}$ solves

$$\mathcal{L}^s\mathsf{u} = \mathsf{z} \text{ in } \Omega, \quad \mathsf{u} = 0 \text{ on } \partial\Omega \text{ a.s.} \tag{1.2b}$$

Here, the controls z are deterministic and Z_{ad} is a nonempty, closed, and convex set of admissible controls. Moreover, u in (1.2b) is a random field and we denote the explicit dependence of u on the uncertain fractional power s by $\mathsf{u}(s)$. For fixed $\mathsf{u}(\cdot)$ and z, the objective function $s \mapsto J(s,\mathsf{u}(s),\mathsf{z})$ is a random variable and $\mathcal{R}$ is a function that maps random variables into the extended real numbers. We refer to (1.2) as the *uncertain fractional optimal control problem.*

There are a number of reasonable choices for the functional $\mathcal{R}$. For example, if $J(\cdot,\mathsf{u}(\cdot),\mathsf{z})$ represents a reliability (e.g., failure) metric or cost associated with (1.1), then it is often reasonable to minimize the probability that $J(\cdot,\mathsf{u}(\cdot),\mathsf{z})$ exceeds a specified threshold. Another option, and the focus of this paper, is to choose $\mathcal{R}$ to be a *measure of risk.* Risk measures numerically quantify the overall "hazard" associated with the random variable $J(\cdot,\mathsf{u}(\cdot),\mathsf{z})$ [48]. For example, one could define $\mathcal{R}$ to be the expected value plus a measure of deviation of $J(\cdot,\mathsf{u}(\cdot),\mathsf{z})$, or one could define $\mathcal{R}$ to be the average of the $(1-\beta)\times 100\%$ largest scenarios of $J(\cdot,\mathsf{u}(\cdot),\mathsf{z})$ for some fixed $0<\beta<1$. See [50] for examples of common risk measures. In general, there is no single method for choosing $\mathcal{R}$. However, $\mathcal{R}$ should encode the required conservativeness or risk preference of the user.

The uncertain fractional optimal control problem presents numerous theoretical and numerical challenges. First, the regularity of solutions to (1.1) depends on the fractional exponent s. Since s is random, the random field solution $\mathsf{u}(\cdot)$ of (1.1) has varying regularity. Therefore, the classical techniques for analyzing (1.1) and (1.2) may not apply. Second, in general the numerical solution of optimization problems governed by PDEs with uncertain inputs is computationally expensive, since one not only must discretize the governing PDE in space but also must approximate the dependence of the PDE solution on the uncertain parameter s. Such problems have recently received much attention [38]. For example, in the recent works [36, 37], the authors developed an optimization algorithm that adaptively refines quadrature approximations of the expectation of the objective function to efficiently solve PDE-constrained optimization problems with uncertain inputs. Similarly, in [35], the author presents an optimization algorithm that utilizes multilevel quadrature approximations of the control problem to reduce computational cost. On the other hand, the authors in [18, 24, 25, 26, 60] discuss efficient optimization approaches that employ reduced-order models of the governing physics. Finally, the authors in [17] demonstrate the use of multigrid to solve optimal control problems with parametric uncertainties.

Aside from the computational difficulties associated with uncertainty, additional complexity arises due to the nonlocal operator $\mathcal{L}^s$. To handle $\mathcal{L}^s$, one can use a spectral Fourier approach [2], the Balakrishnan formula [59, p. 260] (see also [16]), or the so-called Stinga–Torrea extension [52]; see also [20] for the Caffarelli–Silvestre extension. In this work, we employ the Stinga–Torrea extension. However, our results hold in case of the first two approaches as well. A similar strategy was used in [3], where the authors studied a linear quadratic optimal control problem constrained by (1.1) with fixed (deterministic) s. Given the exponent $s \in (0,1)$, a desired state $\mathsf{u}_d \in L^2(\Omega)$, and $J(\mathsf{u},\mathsf{z}) = \frac{1}{2}\|\mathsf{u}-\mathsf{u}_d\|^2_{L^2(\Omega)} + \frac{\mu}{2}\|\mathsf{z}\|^2_{L^2(\Omega)}$, the authors in [3] investigated the so-called *fractional optimal control problem*: minimize $J(\mathsf{u},\mathsf{z})$ subject to the state equation (1.1) and the control constraints $\mathsf{a} \le \mathsf{z} \le \mathsf{b}$ for given $\mathsf{a} \le \mathsf{b}$. For completeness, we refer the reader to [10] for the control of the fractional semilinear PDEs with both integral and spectral fractional diffusion operators and [8, 9] for the control of fractional p-Laplacian with control in the coefficient. We mention that our approach, in principle, can be directly applied to other definitions, integral and regional, of fractional Laplacian [56, 57], provided the solution to the corresponding PDEs fulfills the s-related properties discussed in this paper.

Our principle contributions in this paper span both the theoretical and the numerical treatments of (1.2). In the following list, we summarize the aforementioned difficulties associated with analyzing and solving (1.2), and we outline our contributions.

Variable spatial regularity. The spatial regularity of the random field solution $\mathsf{u}(\cdot)$ depends explicitly on the random fractional power s. Consequently, the classical techniques used to analyze (1.1) and (1.2) do not directly apply.

Existence and optimality conditions. We prove that solutions to (1.2) exist in Theorem 3.1 using the direct method of the calculus of variations. Although this technique is standard, the details for (1.2) rely heavily on the properties of the risk measure; see section 3.1. In addition, we derive the first-order necessary optimality conditions in Theorem 3.2.

Extension and truncation estimates. The numerical solution of (1.2) is daunting due to the inherent complexity of (1.1) combined with the need to accurately resolve the stochastic variability. We employ the Stinga–Torrea extension to equivalently rewrite (1.1) and then approximate it via truncation. The key challenge with this approach is the derivation of the precise s-dependence in the energy estimates of the solution to the truncated problem. We show, for the risk-neutral case, i.e., $\mathcal{R}$ is the expected value, that the error for the optimal controls, when using this truncation, decays exponentially with respect to the truncation parameter.

$L^2(\Omega)$-error estimate. The finite-element error estimates in [3, 47] require $s \in [\varepsilon, 1-\varepsilon]$, with $\varepsilon \in (0,1)$, as well as sufficient regularity of the problem datum (i.e., Sobolev regularity rather than $L^2(\Omega)$). In Theorem 6.3, we provide a finite-element error estimate for $L^2(\Omega)$ problem datum.

Discrete optimal control problem. We introduce a numerical scheme for (1.2) that employs the truncated extension of (1.1), finite elements for the spatial discretization, and interpolation with respect to s. We prove the convergence of this method for the risk-neutral case. However, due to the s-dependent regularity of the adjoint state, we are unable to establish higher regularity of the optimal controls (see Remark 6.4) and therefore cannot, in general, prove error estimates.

The outline of this paper is as follows. In section 2, we introduce notation and study the well-posedness and differentiability of u with respect to s. Section 3 is devoted to the setup of our risk-averse optimal control problem. Under fairly general assumptions, we show the existence of solutions and derive the first-order optimality conditions. In section 4, we realize the nonlocal operator $\mathcal{L}^s$ using the Stinga–Torrea extension. The Stinga–Torrea extension produces a PDE defined on a semi-infinite cylinder. We discuss a truncation strategy in section 5 to handle this semi-infinite cylinder. In particular, we prove that the constants associated with the truncation error bounds are independent of s. To our knowledge, these estimates are new even for the forward problem. In section 6, we present a discrete scheme for general risk-measures, and we also discuss error estimates in case the risk measure is given by the expected value. For these error estimates, we need to assume that $s \in [\varepsilon, 1-\varepsilon]$, with $\varepsilon \in (0,1)$, which is an artifact of current numerical schemes. We conclude with numerical results in section 7.

2. The uncertain fractional equation. We begin by setting our notation (cf. [5]). Let Ω be an open, bounded, and connected domain in $\mathbb{R}^N$, $N \geq 1$, with Lipschitz boundary $\partial\Omega$. Let $\mathcal{L}$ be the realization in $L^2(\Omega)$ of the elliptic operator

$$\mathcal{L}w := -\mathrm{div}_x(A\nabla_x w) + cw$$

with zero Dirichlet boundary conditions. Here, we assume $0 \leq c \in L^\infty(\Omega)$ and $A \in C^{0,1}(\Omega, \mathsf{GL}(N,\mathbb{R}))$ is symmetric and positive definite. It is well known that $\mathcal{L}$ has a compact resolvent and its eigenvalues form a nondecreasing sequence $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_k \leq \cdots$ satisfying $\lim_{k\to\infty} \lambda_k = \infty$ [41]. We denote by $\{\varphi_k\} \subset H_0^1(\Omega)$ the orthonormal eigenfunctions associated with $\{\lambda_k\}$.

For any $s \geq 0$, we define the fractional-order Sobolev space

$$\mathbb{H}^s(\Omega) := \left\{ u = \sum_{k=1}^\infty u_k \varphi_k \in L^2(\Omega) \ : \ \|u\|^2_{\mathbb{H}^s(\Omega)} := \sum_{k=1}^\infty \lambda_k^s u_k^2 < \infty \right\},$$

where the coefficients u_k are defined as

$$u_k := (u, \varphi_k)_{L^2(\Omega)} = \int_\Omega u\varphi_k \,\mathrm{d}x.$$

It is well known that

$$\mathbb{H}^s(\Omega) = \begin{cases} H^s(\Omega) = H_0^s(\Omega) & \text{if } 0 < s < \frac{1}{2}, \\ H_{00}^{\frac{1}{2}}(\Omega) & \text{if } s = \frac{1}{2}, \\ H_0^s(\Omega) & \text{if } \frac{1}{2} < s < 1, \end{cases} \tag{2.1}$$

where $H^s(\Omega)$ denotes the classical fractional-order Sobolev space, $H_{00}^{\frac{1}{2}}(\Omega)$ is the so-called Lions–Magenes space, and $H_0^s(\Omega)$ denotes the fractional-order Sobolev space of functions with zero trace when $s > \frac{1}{2}$. We denote the dual space of $\mathbb{H}^s(\Omega)$ by $\mathbb{H}^{-s}(\Omega)$. We recall the following definition of the nonlocal operator $\mathcal{L}^s$.

Definition 2.1. *The Dirichlet fractional operator is defined on $C_0^\infty(\Omega)$ by*

$$\mathcal{L}^s u := \sum_{k=1}^\infty \lambda_k^s u_k \varphi_k, \quad \textit{with} \quad u_k = (u, \varphi_k)_{L^2(\Omega)}.$$

Using density arguments, we extend the operator $\mathcal{L}^s$ to an operator mapping $\mathbb{H}^s(\Omega)$ into $\mathbb{H}^{-s}(\Omega)$.

We model s as a random variable. Here and throughout, we abuse notation and let s denote a random variable and its realizations. Let $(\Sigma, \mathcal{F}, \mathbb{P})$ denote a complete probability space where Σ is a set of outcomes, $\mathcal{F} \subseteq 2^\Sigma$ is a σ-algebra of events, and $\mathbb{P} : \Sigma \to [0,1]$ is a probability measure. We assume that s is a random variable mapping Σ into $\Xi := (0,1)$ with probability law $P = \mathbb{P} \circ s^{-1}$. We further assume that s is Borel measurable, i.e., $s^{-1}(A) \in \mathcal{B}$ for all $A \in \mathcal{F}$, where $\mathcal{B} \subseteq 2^\Xi$ denotes the Borel σ-algebra on Ξ. This allows us to analyze our optimization problem in the complete probability space $(\Xi, \mathcal{B}, P)$. We denote the expected value of a random variable $X : \Xi \to \mathbb{R}$ by $\mathbb{E}[X] = \int_\Xi X(s)\mathrm{d}P(s)$, and we denote "$P$-almost everywhere" by a.s. (i.e., "almost surely"). In this probabilistic setting, we require the fractional diffusion equation (1.1) to hold for all $s \in \Xi$. The solution to (1.1), u, is a function that maps Ξ into $L^2(\Omega)$ with the added regularity that $\mathsf{u}(s) \in \mathbb{H}^s(\Omega)$ for all $s \in \Xi$. Throughout, we denote the control-to-state map for (1.1) by $\mathsf{z} \mapsto \mathsf{S}(\mathsf{z})$. For a fixed z, $\mathsf{S}(\mathsf{z})$ is parametrized by the fractional power s. We denote this dependence by $s \mapsto [\mathsf{S}(\mathsf{z})](s)$. Given a real Banach space $(V, \|\cdot\|_V)$, we denote the Bochner space for $p \in [1,\infty)$ and $p = \infty$, respectively, by

$$L^p(\Xi,\mathcal{B},P;V) := \{\zeta : \Xi \to V \,|\, \zeta \text{ is strongly } \mathcal{B}\text{-measurable, } \mathbb{E}[\|\zeta\|_V^p] < \infty\},$$
$$L^\infty(\Xi,\mathcal{B},P;V) := \left\{\zeta : \Xi \to V \,|\, \zeta \text{ is strongly } \mathcal{B}\text{-measurable, } \operatorname*{ess\,sup}_{s\in\Xi} \|\zeta(s)\|_V < \infty\right\}.$$

When $V = \mathbb{R}$, we obtain the usual Lebesgue space, denoted by $L^p(\Xi,\mathcal{B},P)$. We now study the existence of solution to (1.1) and the differentiability of $[\mathsf{S}(\mathsf{z})](\cdot)$.

PROPOSITION 2.2. *For fixed* $\mathsf{z} \in L^2(\Omega)$*, there exists a unique solution to* (1.1) *given by*

$$[\mathsf{S}(\mathsf{z})](s) = \sum_{k=1}^\infty \lambda_k^{-s} \mathsf{z}_k \varphi_k, \qquad \mathsf{z}_k := (\mathsf{z}, \varphi_k)_{L^2(\Omega)}, \tag{2.2}$$

with $\mathsf{S}(\mathsf{z}) \in L^\infty(\Xi,\mathcal{B},P;L^2(\Omega))$*. Moreover,* $s \mapsto [\mathsf{S}(\mathsf{z})](s)$ *is infinitely often Fréchet differentiable as a function from* Ξ *into* $L^2(\Omega)$ *for any* $n \in \mathbb{N}$ *and the nth-order derivative is given by*

$$\frac{\partial^n}{\partial s^n}[\mathsf{S}(\mathsf{z})](s) = \sum_{k=1}^\infty (-1)^n \log(\lambda_k)^n \lambda_k^{-s} \mathsf{z}_k \varphi_k, \tag{2.3}$$

where log *denotes the natural logarithm.*

Proof. Theorem 2.5 of [21] (see also [6, Prop. 2.8]) shows that $[\mathsf{S}(\mathsf{z})](s)$ in (2.2) is the unique solution to (1.1) for fixed $s \in \Xi$. Motivated by [51], the three-times Fréchet differentiability was shown in [4, Thm. 3.1]. The higher-order differentiability follows by similar arguments. Now, since $s \mapsto [\mathsf{S}(\mathsf{z})](s)$ is Fréchet differentiable as a function from Ξ into $L^2(\Omega)$, it is continuous on Ξ and therefore, for all $\mathsf{u}^* \in L^2(\Omega)$, the mapping $s \mapsto \langle \mathsf{u}^*, [\mathsf{S}(\mathsf{z})](s)\rangle_{L^2(\Omega)}$ is also continuous. Hence, $s \mapsto [\mathsf{S}(\mathsf{z})](s)$ is weakly $\mathcal{B}$-measurable. Moreover, since $L^2(\Omega)$ is separable, Theorem 3.5.2 in [33] guarantees that $s \mapsto [\mathsf{S}(\mathsf{z})](s)$ is strongly $\mathcal{B}$-measurable. Finally, we recall from [4] that $\|[\mathsf{S}(\mathsf{z})](s)\|_{L^2(\Omega)} \le C\|\mathsf{z}\|_{L^2(\Omega)}$, where the constant $C > 0$ is independent of s whence $\mathsf{S}(\mathsf{z}) \in L^\infty(\Xi,\mathcal{B},P;L^2(\Omega))$. ∎

For fixed $s \in \Xi$, it is easy to see that the state $[\mathsf{S}(\mathsf{z})](s)$ enjoys the following higher spatial regularity.

LEMMA 2.3. *Suppose* $[\mathsf{S}(\mathsf{z})](s)$ *solves* (1.1) *for fixed* $\mathsf{z} \in L^2(\Omega)$ *and* $s \in \Xi$. *Then*

$$\|[\mathsf{S}(\mathsf{z})](s)\|_{\mathbb{H}^{2s}(\Omega)} = \|\mathsf{z}\|_{L^2(\Omega)}.$$

Therefore, the random variable $s \mapsto \|[\mathsf{S}(\mathsf{z})](s)\|_{\mathbb{H}^{2s}(\Omega)}$ *is finite and constant.*

Proof. Since $\mathsf{u}_k = \lambda_k^{-s}\mathsf{z}_k$, we arrive at the asserted result by using the definition of $\mathbb{H}^{2s}$-norm. ∎

3. The optimal control problem. As demonstrated in Proposition 2.2, the state $\mathsf{S}(\mathsf{z}) : \Xi \to L^2(\Omega)$ is n-times continuously Fréchet differentiable for fixed $\mathsf{z} \in L^2(\Omega)$ and any $n \in \mathbb{N}$, and $\mathsf{S}(\mathsf{z}) \in L^\infty(\Xi, \mathcal{B}, P; L^2(\Omega))$. Substituting $\mathsf{S}(\mathsf{z})$ into the objective function J in (1.2) results in the uncertain reduced objective function $J(\cdot, [\mathsf{S}(\mathsf{z})](\cdot), \mathsf{z})$. For the remainder of this paper, we assume J has the specific form

$$J(s, \mathsf{u}, \mathsf{z}) = f(s, \mathsf{u}) + g(\mathsf{z}),$$

where $f : \Xi \times L^2(\Omega) \to \mathbb{R}$ and $g : L^2(\Omega) \to \mathbb{R}$. We further assume that there exist $p, q \in [1, \infty]$ such that $f(\cdot, [\mathsf{S}(\mathsf{z})](\cdot)) \in L^p(\Xi, \mathcal{B}, P)$ for any $\mathsf{u} \in L^q(\Xi, \mathcal{B}, P; L^2(\Omega))$. Such a scenario is typical for optimal control problems (cf. [54]), and in the subsequent results, we provide explicit assumptions on f that ensure this condition holds. To formulate the optimal control problem, we choose a functional $\mathcal{R} : L^p(\Xi, \mathcal{B}, P) \to (-\infty, \infty]$ and solve

$$\min_{\mathsf{z} \in \mathsf{Z}_{\mathrm{ad}}} \mathcal{R}(f(\cdot, [\mathsf{S}(\mathsf{z})](\cdot))) + g(\mathsf{z}). \tag{3.1}$$

Common choices of $\mathcal{R}$ include risk measures, worst-case functionals, and probabilistic functions (see [50] and the references within). Such $\mathcal{R}$ often are not differentiable, adding complication to the analysis and numerical solution of (3.1). In this work, we focus solely on risk measures.

3.1. Risk measures. Minimizing the average objective function is often not sufficiently conservative for real applications. To address this issue, risk measures were developed to numerically quantify hazards associated with uncertain outcomes. There are many ways to model risk preference. In this work, we discuss two classical approaches: (i) mean-deviation models and (ii) disutility models. In the mean-deviation approach, we define $\mathcal{R}$ as

$$\mathcal{R}(X) = \mathbb{E}[X] + \mathcal{D}(X),$$

where $\mathcal{D} : L^p(\Xi, \mathcal{B}, P) \to [0, \infty]$ is a deviation measure and quantifies how nonconstant a random variable is. For a discussion of generalized deviation measures, see [49]. In the utility approach, we first choose a utility function $\mathcal{U} : L^p(\Xi, \mathcal{B}, P) \to \overline{\mathbb{R}}$ that quantifies our expected utility for the random outcomes $X \in L^p(\Xi, \mathcal{B}, P)$. We then define the associated disutility (or regret) as $\mathcal{V}(X) = -\mathcal{U}(-X)$ and the risk as

$$\mathcal{R}(X) = \inf_{t \in \mathbb{R}} \{t + \mathcal{V}(X - t)\}.$$

This risk model is a generalization of the "optimized certainty equivalent" risk measures [15]. These two risk models are linked through the risk quadrangle which provides fundamental relationships connecting measure of risk, deviation, regret, and error [48]. Independent of how $\mathcal{R}$ is constructed, the authors in [11] postulated the following four useful properties for modeling risk: For X, $X' \in L^p(\Xi, \mathcal{B}, P)$ and $t \in \mathbb{R}$, we have the following:

(R1) Subadditivity: $\mathcal{R}(X + X') \le \mathcal{R}(X) + \mathcal{R}(X')$.
(R2) Monotonicity: If $X \le X'$ a.s., then $\mathcal{R}(X) \le \mathcal{R}(X')$.
(R3) Translation equivariance: $\mathcal{R}(X + t) = \mathcal{R}(X) + t$.
(R4) Positive homogeneity: If $t \ge 0$, then $\mathcal{R}(tX) = t\mathcal{R}(X)$.

Clearly, if $\mathcal{R}$ satisfies (R4), then (R4) holds if and only if $\mathcal{R}$ is convex. Thus, (R1) is often replaced by convexity:

(R1′) Convexity: $\mathcal{R}(tX + (1-t)X') \le t\mathcal{R}(X) + (1-t)\mathcal{R}(X')$ for all $t \in [0, 1]$.

If $\mathcal{R}$ satisfies (R1)–(R4), then it is called *coherent*. On the other hand, if $\mathcal{R}$ satisfies (R1′), (R2), and (R3), then it is called a *convex risk measure* [30]. Under appropriate assumptions on $\mathcal{D}$ and $\mathcal{V}$, the associated risk measures are convex, even coherent. Two popular coherent risk measures are the mean-plus-semideviation,

$$\mathcal{R}(X) = \mathbb{E}[X] + c\mathbb{E}[\max\{0, X - \mathbb{E}[X]\}^p]^{\frac{1}{p}}, \quad 0 \le c \le 1,$$

and the conditional value-at-risk,

$$\mathcal{R}(X) = \frac{1}{1-\beta}\int_\beta^1 q_X(\alpha)\,\mathrm{d}\alpha = \inf_{t\in\mathbb{R}}\left\{t + \frac{1}{1-\beta}\mathbb{E}[\max\{0, X - t\}]\right\}, \quad 0 \le \beta \le 1,$$

where $q_X(\alpha)$ denotes the α-quantile of X. Another popular convex risk measure is the entropic risk,

$$\mathcal{R}(X) = \beta^{-1}\log\mathbb{E}[\exp(\beta X)], \quad \beta > 0. \tag{3.2}$$

The entropic risk measure is not coherent since it does not satisfy (R4).

3.2. Existence and optimality conditions. In the following theorem, we provide conditions on f that ensure

$$\widehat{f}(\mathsf{z}) := f(\cdot, [\mathsf{S}(\mathsf{z})](\cdot)) \in L^p(\Xi, \mathcal{B}, P) \quad \forall\, \mathsf{u} \in L^q(\Xi, \mathcal{B}, P; L^2(\Omega))$$

and, under these assumptions, prove the existence of solutions to (3.1). We then prove first-order necessary optimality conditions for (3.1).

Theorem 3.1 (existence of optimal control). *Suppose* $\mathsf{Z}_{\mathrm{ad}} \subset L^2(\Omega)$ *is nonempty, closed, convex, and bounded;* $g : L^2(\Omega) \to \mathbb{R}$ *is convex and lower semicontinuous; and* $f : \Xi \times L^2(\Omega) \to \mathbb{R}$ *satisfies* (i) $f(\cdot, u)$ *is* $\mathcal{B}$*-measureable for all* $u \in L^2(\Omega)$*,* (ii) $f(s, \cdot)$ *is continuous for almost all* $s \in \Xi$*, and* (iii) *there exist* $p, q \in [1, \infty]$ *such that one of the following growth conditions holds:*

1. $p < \infty$*,* $q < \infty$*, and there exist* $a \in L^p(\Xi, \mathcal{B}, P)$ *with* $a \ge 0$ *a.s. and* $C > 0$ *such that*

$$|f(s, \mathsf{u})| \le a(s) + C\|\mathsf{u}\|^{q/p}_{L^2(\Omega)} \quad \forall\, \mathsf{u} \in L^2(\Omega) \quad a.s.; \tag{3.3}$$

2. $p = q = \infty$*, and for all* $C > 0$ *there exists* $\gamma = \gamma(C) \ge 0$ *such that*

$$|f(s, \mathsf{u})| \le \gamma \quad \forall\, \|\mathsf{u}\|_{L^2(\Omega)} \le C \quad a.s. \tag{3.4}$$

Finally, suppose $\mathcal{R} : L^p(\Xi, \mathcal{B}, P) \to (-\infty, \infty]$ *is proper, convex, and lower semicontinuous if* $p < \infty$ *or weakly* lower semicontinuous if* $p = \infty$ *and satisfies the monotonicity property* (R2)*. Then* (3.1) *has a solution.*

Proof. We first show that $[\mathsf{S}(\cdot)](s)$ is a compact operator from $L^2(\Omega)$ into $L^2(\Omega)$ for arbitrary fixed $s \in \Xi$. Since $[\mathsf{S}(\cdot)](s)$ is a continuous linear operator from $L^2(\Omega)$ into $\mathbb{H}^s(\Omega)$, we have that if $\mathsf{z}'_n \rightharpoonup \mathsf{z}'$ in $L^2(\Omega)$, then $[\mathsf{S}(\mathsf{z}'_n)](s) \rightharpoonup [\mathsf{S}(\mathsf{z}')](s)$ in $\mathbb{H}^s(\Omega)$. The compact embedding of $\mathbb{H}^s(\Omega)$ into $L^2(\Omega)$ [28, Thm. 7.1] then ensures that $[\mathsf{S}(\mathsf{z}'_n)](s) \to [\mathsf{S}(\mathsf{z}')](s)$ in $L^2(\Omega)$. In particular, $[\mathsf{S}(\cdot)](s)$ is a compact operator from $L^2(\Omega)$ into $L^2(\Omega)$ for fixed $s \in \Xi$ (cf. [27, Prop. 3.3(b)]). The continuity assumption of $f(s,\cdot)$ and the compactness of $[\mathsf{S}(\cdot)](s)$ ensure that if $\mathsf{z}_n \rightharpoonup \mathsf{z}$ in $L^2(\Omega)$, then $\mathsf{S}(\mathsf{z}_n) \to \mathsf{S}(\mathsf{z})$ in $L^2(\Omega)$ a.s. and $F_n(s) := [\widehat{f}(\mathsf{z}_n)](s) \to [\widehat{f}(\mathsf{z})](s) =: F(s)$ a.s.

Now suppose $p < \infty$, $q < \infty$, and the growth condition (3.3) holds. Then the desired result follows from Proposition 3.12 in [39]. On the other hand, if $p = q = \infty$ and the uniform boundedness condition (3.4) holds, then for $C = \sup_n \|\mathsf{z}_n\|_{L^2(\Omega)}$ (which is finite since z_n weakly converges), there exists $\gamma = \gamma(C) \geq 0$ such that $|F_n(s)| \leq \gamma$ a.s. Now, let $\vartheta \in L^1(\Xi, \mathcal{B}, P)$ be arbitrary. Then

$$|\vartheta F_n| \leq |\vartheta|\gamma \quad \text{a.s.}$$

and $|\vartheta|\gamma \in L^1(\Xi, \mathcal{B}, P)$. Therefore, the Lebesgue dominated convergence theorem ensures that

$$\lim_{n\to\infty} \mathbb{E}[\vartheta F_n] = \mathbb{E}[\vartheta F].$$

That is, $F_n \rightharpoonup^* F$ in $L^\infty(\Xi, \mathcal{B}, P)$. Since $\mathcal{R}$ is weakly* lower semicontinuous, we have

$$\liminf_{n\to\infty} \mathcal{R}(F_n) \geq \mathcal{R}(F),$$

and hence $z \mapsto \mathcal{R}(\widehat{f}(\mathsf{z}))$ is weakly lower semicontinuous. Since $g : L^2(\Omega) \to \mathbb{R}$ is assumed to be convex and lower semicontinuous, it is weakly lower semicontinuous. The existence of a minimizer then follows from the direct method in the calculus of variations [12, Thm. 3.2.1]. ∎

THEOREM 3.2 (first-order optimality conditions). *Let the assumptions of Theorem* 3.1 *hold, and assume that $f(s,\cdot)$ is continuously Fréchet differentiable for almost all $s \in \Xi$ and $\nabla_{\mathsf{u}} f(\cdot, \mathsf{u})$ is strongly measurable for all $\mathsf{u} \in L^2(\Omega)$. Moreover, assume that there exists $\alpha > 0$ and $K \in L^r(\Xi, \mathcal{B}, P)$ with*

$$r = \begin{cases} pq/(q-(1+\alpha)p) & \text{if } q > (1+\alpha)p, \\ \infty & \text{if } q = (1+\alpha)p \end{cases}$$

such that

$$|\nabla_{\mathsf{u}} f(s,\mathsf{u}) - \nabla_{\mathsf{u}} f(s,\mathsf{u}')| \leq K(s)\|\mathsf{u}-\mathsf{u}'\|^\alpha_{L^2(\Omega)} \quad \text{a.s.}$$

Finally, assume that $\mathcal{R}$ is finite on $L^p(\Xi, \mathcal{B}, P)$. If $\bar{\mathsf{z}} \in \mathsf{Z}_{\mathrm{ad}}$ minimizes $\widehat{f}(\cdot) + g(\cdot)$ over Z_{ad}, then the first-order necessary conditions,

(3.5)
$$\sup_{\theta \in \partial\mathcal{R}(f(\cdot,[\mathsf{S}(\bar{\mathsf{z}})](\cdot)))} \mathbb{E}\left[\theta \int_\Omega [\mathsf{P}(\bar{\mathsf{z}})](\cdot)(\mathsf{z}'-\bar{\mathsf{z}})\,\mathrm{d}x\right] + \int_\Omega \nabla g(\bar{\mathsf{z}})(\mathsf{z}'-\bar{\mathsf{z}})\,\mathrm{d}x \geq 0 \quad \forall\, \mathsf{z}' \in \mathsf{Z}_{\mathrm{ad}},$$

hold where for fixed $s \in \Xi$ and $\mathsf{z} \in L^2(\Omega)$, $\mathsf{p} = [\mathsf{P}(\mathsf{z})](s) \in \mathbb{H}^s(\Omega)$ solves the adjoint equation

$$\mathcal{L}^s \mathsf{p} = \nabla_{\mathsf{u}} f(s, [\mathsf{S}(\mathsf{z})](s)) \text{ in } \Omega, \qquad \mathsf{p} = 0 \text{ on } \partial\Omega$$

and $\partial\mathcal{R}(X)$ is the usual convex subdifferential of $\mathcal{R}$ at X.

Proof. If $1 \le p,\, q < \infty$, then Theorem 3.11 in [39] ensures that $\mathsf{u} \mapsto f(\cdot, \mathsf{u}(\cdot))$ is Fréchet differentiable. On the other hand, if $p = q = \infty$, then $r = \infty$ and we can take K to be constant. In this case, for any $\mathsf{h} \in L^\infty(\Xi, \mathcal{B}, P; L^2(\Omega))$, we have that

$$\begin{aligned} &|f(s, (\mathsf{u}+\mathsf{h})(s)) - f(s, \mathsf{u}(s)) - \nabla_{\mathsf{u}} f(s, \mathsf{u}(s))\mathsf{h}(s)| \\ &\quad \le \int_0^1 |(\nabla_{\mathsf{u}} f(s, (\mathsf{u}+t\mathsf{h})(s)) - \nabla_{\mathsf{u}} f(s, \mathsf{u}(s)))\mathsf{h}(s)|\,\mathrm{d}t \le \frac{1}{2} K \|\mathsf{h}(s)\|_{L^2(\Omega)}^{1+\alpha} \quad \text{a.s.} \end{aligned}$$

Passing to the essential supremum on both sides and noting that

$$\operatorname*{ess\,sup}_{s\in\Xi} \|\mathsf{h}(s)\|_{L^2(\Omega)}^{1+\alpha} = \left(\operatorname*{ess\,sup}_{s\in\Xi} \|\mathsf{h}(s)\|_{L^2(\Omega)} \right)^{1+\alpha}$$

ensures that $\mathsf{u} \mapsto f(\cdot, \mathsf{u}(\cdot))$ is Fréchet differentiable. Since $\mathsf{z} \mapsto \mathsf{S}(\mathsf{z})$ is a continuous linear mapping and $\mathsf{u} \mapsto f(\cdot, \mathsf{u}(\cdot))$ is Fréchet differentiable, $\mathsf{z} \mapsto \widehat{f}(\mathsf{z})$ is Fréchet differentiable from $L^2(\Omega)$ into $L^p(\Xi, \mathcal{B}, P)$. The first-order conditions then follow directly from Corollary 3.14 in [39]. ∎

Remark 3.3 (expected value and entropic risk). Suppose $\mathcal{R} \equiv \mathbb{E}$; then the optimality conditions (3.5) simplify to

$$\int_\Omega (\mathbb{E}\left[[\mathsf{P}(\bar{\mathsf{z}})](\cdot)\right] + \nabla g(\bar{\mathsf{z}}))(\mathsf{z}' - \bar{\mathsf{z}})\,\mathrm{d}x \ge 0 \quad \forall\, \mathsf{z}' \in \mathsf{Z}_{\mathrm{ad}};$$

cf. [33, Thm. 3.7.12]. The above variational inequality is equivalent to the following projection formula:

$$\bar{\mathsf{z}} = \mathbb{P}_{\mathsf{Z}_{\mathrm{ad}}}\left(\bar{\mathsf{z}} - \gamma(\mathbb{E}[[\mathsf{P}(\bar{\mathsf{z}})](\cdot)] + \nabla g(\bar{\mathsf{z}}))\right) \quad \forall\, \gamma > 0, \tag{3.6}$$

where $\mathbb{P}_{\mathsf{Z}_{\mathrm{ad}}} : L^2(\Omega) \to L^2(\Omega)$ is the projection onto the convex set Z_{ad}; cf. [12, Thm. 3.3.5].

Similarly, suppose $p = q = \infty$ and $\mathcal{R}$ is the entropic risk measure (3.2). Then the optimality conditions (3.5) simplify to

$$\int_\Omega \left(\mathbb{E}\left[\frac{\exp(\beta f(\cdot, [\mathsf{S}(\bar{\mathsf{z}})](\cdot)))}{\mathbb{E}[\exp(\beta f(\cdot, [\mathsf{S}(\bar{\mathsf{z}})](\cdot)))]} [\mathsf{P}(\bar{\mathsf{z}})](\cdot) \right] + \nabla g(\bar{\mathsf{z}}) \right) (\mathsf{z}' - \bar{\mathsf{z}})\,\mathrm{d}x \ge 0 \quad \forall\, \mathsf{z}' \in \mathsf{Z}_{\mathrm{ad}}.$$

Again, this variational inequality is equivalent to the following projection formula:

$$\bar{\mathsf{z}} = \mathbb{P}_{\mathsf{Z}_{\mathrm{ad}}} \left(\bar{\mathsf{z}} - \gamma \left(\mathbb{E}\left[\frac{\exp(\beta f(\cdot, [\mathsf{S}(\bar{\mathsf{z}})](\cdot)))}{\mathbb{E}[\exp(\beta f(\cdot, [\mathsf{S}(\bar{\mathsf{z}})](\cdot)))]} [\mathsf{P}(\bar{\mathsf{z}})](\cdot) \right] + \nabla g(\bar{\mathsf{z}}) \right) \right) \quad \forall \gamma > 0.$$

4. Extended optimal control problem. It is well known that problem (1.1) can equivalently be posed on a semi-infinite cylinder. This approach was originally due to Molchanov and Ostrovskii [45] and was rediscovered by Caffarelli and Silvestre [20] for $\mathbb{R}^N$. Stinga and Torrea exploited the ideas of Caffarelli and Silvestre to define the fractional Dirichlet Laplacian on bounded open sets [52] (see also [19, 22]). We mention that for the existence and uniqueness of solutions to the problem on this semi-infinite cylinder it is sufficient to consider an open set with a Lipschitz continuous boundary; see [21, Thm. 2.5] for details. We operate under the same setup in the present section, and we follow the notation of [6].

Let $\mathcal{C}$ be the aforementioned semi-infinite cylinder with base Ω, i.e., $\mathcal{C} = \Omega \times (0, \infty)$, and denote its lateral boundary by $\partial_L \mathcal{C} := \partial\Omega \times [0, \infty)$. We denote the truncated cylinder by $\mathcal{C}_\tau := \Omega \times (0, \tau)$ for $\tau > 0$. Similar to the lateral boundary $\partial_L \mathcal{C}$, we set $\partial_L \mathcal{C}_\tau := (\partial\Omega \times [0, \tau]) \cup (\Omega \times \{\tau\})$. As a result, the semi-infinite cylinder and its truncated version are objects defined in $\mathbb{R}^{N+1}$. Throughout the remainder of the paper, y denotes the extended variable, such that a vector $x' \in \mathbb{R}^{N+1}$ admits the representation $x' = (x_1, \ldots, x_N, x_{N+1}) = (x, x_{N+1}) = (x, y)$ with $x_i \in \mathbb{R}$ for $i = 1, \ldots, N+1$, $x \in \mathbb{R}^N$ and $y \in \mathbb{R}$.

The extension problem requires certain weighted Sobolev spaces for its solvability due to the degenerate/singular nature of the operator. The weight function is y^α, $\alpha \in (-1, 1)$; see [46], [55, sect. 2.1], [40], and [32, Thm. 1] for a more sophisticated discussion of such spaces. To this end, let $\mathcal{D} \subset \mathbb{R}^N \times [0, \infty)$ be an open set, such as $\mathcal{C}$ or $\mathcal{C}_\tau$; then we define the weighted space $L^2(y^\alpha, \mathcal{D})$ as the space of all measurable functions defined on $\mathcal{D}$ with finite norm $\|w\|_{L^2(y^\alpha,\mathcal{D})} := \|y^{\alpha/2} w\|_{L^2(\mathcal{D})}$. Similarly, using a standard multi-index notation, the space $H^1(y^\alpha, \mathcal{D})$ denotes the space of all measurable functions w on $\mathcal{D}$ whose weak derivatives $D^\delta w$ exist for $|\delta| = 1$ and fulfill

$$\|w\|_{H^1(y^\alpha,\mathcal{D})} := \left(\sum_{|\delta|\le 1} \|D^\delta w\|^2_{L^2(y^\alpha,\mathcal{D})} \right)^{1/2} < \infty.$$

To study the extended problems, we also need the space

$$\mathring{H}^1_L(y^\alpha, \mathcal{C}) := \{ w \in H^1(y^\alpha, \mathcal{C}) \,|\, w = 0 \text{ on } \partial_L \mathcal{C} \}.$$

The space $\mathring{H}^1_L(y^\alpha, \mathcal{C}_\tau)$ is defined in an analogous manner.

The extended problem reads as follows: Given $\mathsf{z} \in L^2(\Omega)$ and fixed $s \in \Xi$, find $\mathscr{U} \in \mathring{H}^1_L(y^{\alpha_s}, \mathcal{C})$ such that

$$\int_{\mathcal{C}} y^{\alpha_s} \left(\mathbf{A}(x,y) \nabla \mathscr{U} \cdot \nabla \Phi + c(x) \mathscr{U} \Phi \right) \,\mathrm{d}x \,\mathrm{d}y = d_s \int_\Omega \mathsf{z} \Phi|_{\Omega \times \{0\}} \,\mathrm{d}x \tag{4.1a}$$

for all $\Phi \in \mathring{H}^1_L(y^{\alpha_s}, \mathcal{C})$ with

$$\alpha_s := 1 - 2s \quad \text{and} \quad d_s := 2^{\alpha_s} \frac{\Gamma(1-s)}{\Gamma(s)}. \tag{4.1b}$$

That is, the function $\mathscr{U} \in \mathring{H}^1_L(y^{\alpha_s}, \mathcal{C})$ is a weak solution of the following problem:

$$\begin{cases} -\mathrm{div}(y^{\alpha_s} \mathbf{A}(x,y) \nabla \mathscr{U}) + y^{\alpha_s} c(x) \mathscr{U} = 0 & \text{in } \mathcal{C}, \\ \frac{\partial \mathscr{U}}{\partial \nu^{\alpha_s}} = d_s \mathsf{z} & \text{on } \Omega \times \{0\}, \end{cases} \tag{4.2}$$

where we have set

$$\frac{\partial \mathscr{U}}{\partial \nu^{\alpha_s}}(x, 0) = -\lim_{y \to 0} y^{\alpha_s} \mathscr{U}_y(x,y) = -\lim_{y\to 0} y^{\alpha_s} \frac{\partial \mathscr{U}(x,y)}{\partial y}$$

and $\mathbf{A}(x,y) = \mathrm{diag}\{A(x), 1\}$. Throughout, we denote the control-to-state map for (4.2) by $\mathsf{z} \mapsto \mathscr{S}(\mathsf{z})$. For a fixed z, $\mathscr{S}(\mathsf{z})$ is parametrized by the fractional power s. We denote this dependence by $s \mapsto [\mathscr{S}(\mathsf{z})](s)$. Furthermore, we recall that if $[\mathscr{S}(\mathsf{z})](s) \in \mathring{H}^1_L(y^{\alpha_s}, \mathcal{C})$ solves (4.2), then we obtain the solution to (1.1) as $[\mathsf{S}(\mathsf{z})](s) = [\mathscr{S}(\mathsf{z})](s)|_{\Omega \times \{0\}}$.

Using this extension, we arrive at the extended optimal control problem

$$\min_{\mathsf{z} \in \mathsf{Z}_{\mathrm{ad}}} \; \mathcal{R}(f(\cdot, [\mathscr{S}(\mathsf{z})](\cdot)|_{\Omega \times \{0\}})) + g(\mathsf{z}). \tag{4.3}$$

Notice that (4.3) is equivalent to (3.1). Under the assumptions of Theorems 3.1 and 3.2 if $\bar{\mathsf{z}}$ solves (4.3), then the first-order necessary conditions,

$$\sup_{\theta\in\partial\mathcal{R}(f(\cdot,[\mathscr{S}(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}}))} \mathbb{E}\left[\theta\int_\Omega [\mathscr{P}(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}}(\mathsf{z}'-\bar{\mathsf{z}})\,\mathrm{d}x\right] + \int_\Omega \nabla g(\bar{\mathsf{z}})(\mathsf{z}'-\bar{\mathsf{z}})\,\mathrm{d}x \geq 0 \tag{4.4}$$

for all $\mathsf{z}'\in\mathsf{Z}_{\mathrm{ad}}$, hold, where for a fixed $s\in\Xi$ and $\mathsf{z}\in L^2(\Omega)$, $\mathscr{P}=[\mathscr{P}(\mathsf{z})](s)\in \mathring{H}^1_L(y^\alpha,\mathcal{C})$ solves the weak form of the adjoint equation, i.e., for all $\Phi\in\mathring{H}^1_L(y^{\alpha_s},\mathcal{C})$,

$$\begin{aligned}\int_{\mathcal{C}} y^{\alpha_s}\left(\mathbf{A}(x,y)\nabla\mathscr{P}\cdot\nabla\Phi + c(x)\mathscr{P}\Phi\right)\,\mathrm{d}x\,\mathrm{d}y \\ = d_s\int_\Omega \nabla_{\mathsf{u}} f(s,[\mathscr{S}(\mathsf{z})](s)|_{\Omega\times\{0\}})\Phi|_{\Omega\times\{0\}}\,\mathrm{d}x.\end{aligned}$$

Remark 4.1. If $\mathcal{R}\equiv\mathbb{E}$ or $\mathcal{R}$ is the entropic risk measure from Remark 3.3 , then (4.4) reduces to

$$\int_\Omega(\mathbb{E}\left[[\mathscr{P}(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}}\right] + \nabla g(\bar{\mathsf{z}}))(\mathsf{z}'-\bar{\mathsf{z}})\,\mathrm{d}x \geq 0 \tag{4.5}$$

for all $\mathsf{z}'\in\mathsf{Z}_{\mathrm{ad}}$ and

$$\int_\Omega\left(\mathbb{E}\left[\frac{\exp(\beta f(\cdot,[\mathscr{S}(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}}))}{\mathbb{E}[\exp(\beta f(\cdot,[\mathscr{S}(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}}))]}[\mathscr{P}(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}}\right] + \nabla g(\bar{\mathsf{z}})\right)(\mathsf{z}'-\bar{\mathsf{z}})\,\mathrm{d}x \geq 0$$

for all $\mathsf{z}'\in\mathsf{Z}_{\mathrm{ad}}$, respectively.

5. The truncated optimal control problem.

Even though the state equation (4.1) is "local," it is posed on the infinite cylinder $\mathcal{C}=\Omega\times(0,\infty)$. Therefore, it cannot be approximated with finite-element-like techniques. Following [47] we truncate $\mathcal{C}$ to a bounded cylinder $\mathcal{C}_\tau=\Omega\times(0,\tau)$. For a fixed $s\in\Xi$, let v denote the solution to the truncated extended PDE: Given $\zeta\in L^2(\Omega)$, find $v\in\mathring{H}^1_L(y^{\alpha_s},\mathcal{C}_\tau)$ such that

$$\int_{\mathcal{C}_\tau} y^{\alpha_s}\left(\mathbf{A}(x,y)\nabla v\cdot\nabla\Phi + c(x)v\Phi\right)\,\mathrm{d}x\mathrm{d}y = d_s\int_\Omega \zeta\Phi|_{\Omega\times\{0\}}\,\mathrm{d}x \tag{5.1}$$

for all $\Phi\in\mathring{H}^1_L(y^{\alpha_s},\mathcal{C}_\tau)$. We denote the control-to-state map for (5.1) by $\zeta\mapsto\mathscr{S}_\tau(\zeta)$. For a fixed ζ, $\mathscr{S}_\tau(\zeta)$ is parametrized by the fractional power s. We denote this dependence by $s\mapsto[\mathscr{S}_\tau(\zeta)](s)$. Using this truncation, we arrive at the *truncated optimal control problem*

$$\min_{\zeta\in\mathsf{Z}_{\mathrm{ad}}} \mathcal{R}(f(\cdot,[\mathscr{S}_\tau(\zeta)](\cdot)|_{\Omega\times\{0\}})) + g(\zeta). \tag{5.2}$$

5.1. Truncation error bounds.

In this subsection, we derive error bounds for the optimal controls of (3.1) and (5.2). We first recall that for fixed $s\in\Xi$ the solution to (4.1) can be expanded as (see, e.g., [22])

$$\mathscr{U}(x') = \sum_{k=1}^\infty u_k\varphi_k(x)\psi_k(y)$$

for $x'=(x,y)\in\mathcal{C}$, where

$$\psi_k(y) := \frac{2^{1-s}}{\Gamma(s)}(\sqrt{\lambda_k}y)^s K_s(\sqrt{\lambda_k}y)$$

and K_s is the modified Bessel function of the second kind.

The truncation error bounds from [47] do not directly extend to our probabilistic setting since the constants in [47, Prop. 3.1] (which depend on s) are neglected. When s is variable, one must ensure that these constants remain bounded for all s under consideration. The following technical lemma is critical for extending the truncation error bound in [47] to our probabilistic setting.

LEMMA 5.1. *For all $s \in \Xi$, we have the following bound:*

$$|y^{\alpha_s}\psi_k(y)\psi_k'(y)| \le C\frac{c_s}{c_{1-s}}\lambda_k^s(\sqrt{\lambda_k}y)^{\left|s-\frac{1}{2}\right|}e^{-2\sqrt{\lambda_k}y} \quad \textit{for } y \ge \frac{\sqrt{\lambda_1}}{\sqrt{\lambda_k}} \tag{5.3}$$

with a constant $C := C(\lambda_1) > 0$ only depending on λ_1, and $c_r = 2^{1-r}/\Gamma(r)$, $r \in \{s, 1-s\}$.

Proof. Our proof follows closely the arguments in [47, sect. 2.5] and [42, Appx.]. However, we work out the exact dependence on the order s of the fractional operator in the constants. As a preliminary result, we start with bounding the term $c_t z^{1/2}K_t(z)$ for $t \in (0,1)$ and $z \ge z_0 > 0$ with $c_t = 2^{1-t}/\Gamma(t)$. Let $t_0 = \min\{t, \frac{1}{2}\}$. From [44, Thm. 5] we obtain that $z^{t_0}e^zK_t(z)$ is a decreasing function for $z > 0$. Moreover, according to [42, Lem. A.2] we have that $c_t z^t K_t(z)$ is positive and monotone decreasing. We also notice that $c_t z^t K_t(z) \sim 1$ as z tends to zero; see [1, sect. 9.6.9]. By this we deduce

$$\begin{aligned}
0 < c_t z^{1/2}K_t(z) &= e^{-z}z^{1/2-t_0}c_t z^{t_0}e^zK_t(z) \le e^{-z}z^{1/2-t_0}c_t z_0^{t_0}e^{z_0}K_t(z_0)\\
&= e^{-z}z^{1/2-t_0}e^{z_0}z_0^{t_0-t}c_t z_0^t K_t(z_0) \le e^{-z}z^{1/2-t_0}e^{z_0}z_0^{t_0-t}\\
&\le C(z_0)e^{-z}z^{1/2-t_0}
\end{aligned} \tag{5.4}$$

with a constant $C(z_0) > 0$ only depending on z_0 for $t \in (0,1)$. Next, by [47, eq. 2.29], we have

$$\psi_k'(y) = -c_s\sqrt{\lambda_k}(\sqrt{\lambda_k}y)^sK_{1-s}(\sqrt{\lambda_k}y),$$

and hence by means of (5.4) with $z = \sqrt{\lambda_k}y$ and $z_0 = \sqrt{\lambda_1}$, we obtain

$$\begin{aligned}
|y^{\alpha_s}\psi_k(y)\psi_k'(y)| &= y^{\alpha_s}c_s(\sqrt{\lambda_k}y)^sK_s(\sqrt{\lambda_k}y)c_s\sqrt{\lambda_k}(\sqrt{\lambda_k}y)^sK_{1-s}(\sqrt{\lambda_k}y)\\
&= \frac{c_s}{c_{1-s}}\lambda_k^s c_s(\sqrt{\lambda_k}y)^{1/2}K_s(\sqrt{\lambda_k}y)c_{1-s}(\sqrt{\lambda_k}y)^{1/2}K_{1-s}(\sqrt{\lambda_k}y)\\
&\le C(\lambda_1)\frac{c_s}{c_{1-s}}\lambda_k^s(\sqrt{\lambda_k}y)^{1-\min\{s,\frac{1}{2}\}-\min\{1-s,\frac{1}{2}\}}e^{-2\sqrt{\lambda_k}y}\\
&= C(\lambda_1)\frac{c_s}{c_{1-s}}\lambda_k^s(\sqrt{\lambda_k}y)^{\left|s-\frac{1}{2}\right|}e^{-2\sqrt{\lambda_k}y}
\end{aligned}$$

for $y \ge \sqrt{\lambda_1}/\sqrt{\lambda_k}$ with a constant $C(\lambda_1)$ only depending on λ_1. ∎

We emphasize that the constant in estimate (5.3) only depends on λ_1, i.e., the first eigenvalue of $\mathcal{L}$. The dependence on λ_1 can be easily carried out by writing $C(z_0)$ explicitly in (5.4). However, for the current discussion the precise expression of this constant is irrelevant. The key points are that C is independent of s and that we have exponential decay with respect to y. We further emphasize that one might be tempted to let $\alpha_s \to 1$, which means $s \to 0$. This is problematic, as the weight y^{α_s} may not fulfill the Muckenhoupt property (cf. [7, Prop. 2.1]), which is a sufficient condition for the density of smooth functions in the weighted Sobolev space

$\mathring{H}^1_L(y^\alpha, \mathcal{C})$. Furthermore, if we let $s = 1$, then (1.1) is a standard elliptic equation. However, in this case, it is unclear what the extension means. We refer the reader to [7] for a further discussion on this topic where s is a function of the spatial variable $x \in \Omega$.

Using Lemma 5.1, we arrive at the following truncation bounds for the state $\mathscr{S}$.

PROPOSITION 5.2. *If, for a given* $s \in (0,1)$, $[\mathscr{S}(\mathsf{z})](s) \in \mathring{H}^1_L(y^{\alpha_s}, \mathcal{C})$ *solves* (4.1), *then for every* $\tau \geq 1$ *and* $s \in \Xi$, *we have*

$$\|\nabla[\mathscr{S}(\mathsf{z})](s)\|_{L^2(y^{\alpha_s},\Omega\times(\tau,\infty))} \leq C\left(\frac{c_s}{c_{1-s}}\right)^{\frac{1}{2}} e^{-\sqrt{\lambda_1}\tau/2}\|\mathsf{z}\|_{\mathbb{H}^{-s}(\Omega)} \tag{5.5}$$

with a constant $C := C(\lambda_1) > 0$ *only depending on* λ_1, *and* $c_r = 2^{1-r}/\Gamma(r)$, $r \in \{s, 1-s\}$.

Proof. The proof of this truncation bound follows directly from the proof of Proposition 3.1 in [47] using Lemma 5.1. Indeed, from the proof of Proposition 3.1 in [47], Lemma 5.1, and Proposition 2.2, we obtain that

$$\begin{aligned}
&\|\nabla[\mathscr{S}(\mathsf{z})](s)\|^2_{L^2(y^{\alpha_s},\Omega\times(\tau,\infty))} \\
&\quad = \sum_{k=1}^{\infty}\left(\int_\Omega [\mathsf{S}(\mathsf{z})](s)\varphi_k \,\mathrm{d}x\right)^2 y^{\alpha_s}\psi_k(y)\psi_k'(y)\Big|_\tau^\infty \\
&\quad \leq C(\lambda_1)\frac{c_s}{c_{1-s}}\sum_{k=1}^{\infty}\lambda_k^{-s}\left(\int_\Omega \mathsf{z}\varphi_k\,\mathrm{d}x\right)^2 e^{-\sqrt{\lambda_k}\tau}(\sqrt{\lambda_k}\tau)^{|s-\frac{1}{2}|}e^{-\sqrt{\lambda_k}\tau} \\
&\quad \leq C(\lambda_1)\frac{c_s}{c_{1-s}}\sum_{k=1}^{\infty}\lambda_k^{-s}\left(\int_\Omega \mathsf{z}\varphi_k\,\mathrm{d}x\right)^2 e^{-\sqrt{\lambda_k}\tau} \\
&\quad = C(\lambda_1)\frac{c_s}{c_{1-s}}e^{-\sqrt{\lambda_1}\tau}\|\mathsf{z}\|^2_{\mathbb{H}^{-s}(\Omega)},
\end{aligned}$$

where we used $s \in (0,1)$, and several times that the sequence $(\lambda_k)_{k\in\mathbb{N}}$ is nondecreasing. ∎

PROPOSITION 5.3 (exponential convergence). *Let* $s \in (0,1)$. *Moreover, let* $[\mathscr{S}(\mathsf{z})](s) \in \mathring{H}^1_L(y^{\alpha_s}, \mathcal{C})$ *solve* (4.1), *and let* $[\mathscr{S}_\tau(\mathsf{z})](s) \in \mathring{H}^1_L(y^{\alpha_s}, \mathcal{C}_\tau)$ *solve* (5.1) *(extended by zero to* $\mathcal{C}$*). Then, for any* $\tau \geq 1$ *and* $s \in \Xi$, *we have*

$$\|\nabla\left([\mathscr{S}(\mathsf{z})](s) - [\mathscr{S}_\tau(\mathsf{z})](s)\right)\|_{L^2(y^{\alpha_s},\mathcal{C})} \leq C\left(\frac{c_s}{c_{1-s}}\right)^{\frac{1}{2}} e^{-\sqrt{\lambda_1}\tau/4}\|\mathsf{z}\|_{\mathbb{H}^{-s}(\Omega)}, \tag{5.6}$$

$$\|[\mathscr{S}(\mathsf{z})](s)|_{\Omega\times\{0\}} - [\mathscr{S}_\tau(\mathsf{z})](s)|_{\Omega\times\{0\}}\|_{\mathbb{H}^s(\Omega)} \leq Ce^{-\sqrt{\lambda_1}\tau/4}\|\mathsf{z}\|_{\mathbb{H}^{-s}(\Omega)}, \tag{5.7}$$

where $C := C(\Omega, \lambda_1)$ *is a positive constant only depending on* Ω *and* λ_1, *and* $c_r = 2^{1-r}/\Gamma(r)$, $r \in \{s, 1-s\}$.

Proof. To show (5.6), we follow the proof of [47, Thm. 3.5] using Lemma 5.1 and Proposition 5.2 instead of [47, eq. (2.32)] and [47, Prop. 3.1], respectively. According to the trace theorem which follows from [22, Prop. 2.1], we have

$$\|V|_{\Omega\times\{0\}}\|_{\mathbb{H}^s(\Omega)} \leq \left(\frac{c_{1-s}}{c_s}\right)^{\frac{1}{2}}\|y^{\frac{\alpha_s}{2}}\nabla V\|_{L^2(\mathcal{C})} \quad \forall V \in \mathring{H}^1_L(y^{\alpha_s}, \mathcal{C}), \tag{5.8}$$

so that (5.7) follows from (5.6). ∎

Again under the assumptions of Theorems 3.1 and 3.2, if $\bar{\zeta}$ solves (5.2), then the following first-order optimality conditions hold:
(5.9)

$$\sup_{\theta\in\partial\mathcal{R}(f(\cdot,[\mathscr{S}_\tau(\bar{\zeta})](\cdot)))} \mathbb{E}\left[\theta\int_\Omega [\mathscr{P}_\tau(\bar{\zeta})](\cdot)|_{\Omega\times\{0\}}(\zeta'-\bar{\zeta})\,\mathrm{d}x\right] + \int_\Omega \nabla g(\bar{\zeta})(\zeta'-\bar{\zeta})\,\mathrm{d}x \geq 0$$

for all $\zeta' \in \mathsf{Z}_{\mathrm{ad}}$, where for fixed $s \in \Xi$ and $\zeta \in L^2(\Omega)$, $\mathscr{P}_\tau = [\mathscr{P}_\tau(\zeta)](s) \in \mathring{H}^1_L(y^{\alpha_s}, \mathcal{C}_\tau)$ solves the adjoint equation for all $\Phi \in \mathring{H}^1_L(y^{\alpha_s}, \mathcal{C}_\tau)$:

$$\begin{aligned}\int_{\mathcal{C}_\tau} y^{\alpha_s}\left(\mathbf{A}(x,y)\nabla\mathscr{P}_\tau\cdot\nabla\Phi + c(x)\mathscr{P}_\tau\Phi\right)\,\mathrm{d}x\mathrm{d}y \\ = d_s\int_\Omega \nabla_u f(s,[\mathscr{S}_\tau(\zeta)](s)|_{\Omega\times 0})\Phi|_{\Omega\times\{0\}}\,\mathrm{d}x.\end{aligned}$$

Remark 5.4. If $\mathcal{R}\equiv\mathbb{E}$ or $\mathcal{R}$ is the entropic risk measure from Remark 3.3, then (5.9) reduces to

$$\int_\Omega (\mathbb{E}\left[[\mathscr{P}_\tau(\bar{\zeta})](\cdot)|_{\Omega\times\{0\}}\right] + \nabla g(\bar{\zeta}))(\zeta'-\bar{\zeta})\,\mathrm{d}x \geq 0 \quad \forall\,\zeta'\in\mathsf{Z}_{\mathrm{ad}} \tag{5.10}$$

and

$$\int_\Omega\left(\mathbb{E}\left[\frac{\exp(\beta f(\cdot,[\mathscr{S}_\tau(\bar{\zeta})](\cdot)|_{\Omega\times\{0\}}))}{\mathbb{E}[\exp(\beta f(\cdot,[\mathscr{S}_\tau(\bar{\zeta})](\cdot)|_{\Omega\times\{0\}}))]}[\mathscr{P}_\tau(\bar{\zeta})](\cdot)|_{\Omega\times\{0\}}\right] + \nabla g(\bar{\zeta})\right)(\zeta'-\bar{\zeta})\,\mathrm{d}x \geq 0$$

for all $\zeta' \in \mathsf{Z}_{\mathrm{ad}}$, respectively.

We next state a result for the exponential convergence for the control and state for a special case.

THEOREM 5.5. *Let $\mathcal{R}\equiv\mathbb{E}$, $f(\cdot,u)=\frac{1}{2}\|u-\mathsf{u}_d(\cdot)\|^2_{L^2(\Omega)}$, where $\mathsf{u}_d:\Xi\to L^2(\Omega)$ is continuous with $\|\mathsf{u}_d(s)\|_{L^2(\Omega)}\leq D$ for all $s\in\Xi$, and $g(z)=\frac{\mu}{2}\|z\|^2_{L^2(\Omega)}$ with $\mu>0$. If $\bar{\mathsf{z}}$ solves the uncertain fractional optimal control problem and $\bar{\zeta}$ solves the truncated optimal control problem, then for every $\tau\geq 1$ we obtain that*

$$\|\bar{\zeta}-\bar{\mathsf{z}}\|_{L^2(\Omega)} \leq Ce^{-\sqrt{\lambda_1}\tau/4}\left(\|\bar{\mathsf{z}}\|_{L^2(\Omega)}+D\right) \tag{5.11}$$

and, for all $s\in\Xi$, we have that

$$\|[\mathscr{S}(\bar{\mathsf{z}})-\mathscr{S}_\tau(\bar{\zeta})](s)|_{\Omega\times\{0\}}\|_{L^2(\Omega)} \leq Ce^{-\sqrt{\lambda_1}\tau/4}\left(\|\bar{\mathsf{z}}\|_{L^2(\Omega)}+D\right), \tag{5.12}$$

where the constant $C>0$ is independent of s.

Proof. We start by setting $\mathsf{z}'=\bar{\zeta}$ and $\zeta'=\bar{\mathsf{z}}$ in (4.5) and (5.10), respectively. For simplicity, we will skip the notation $\cdot|_{\Omega\times\{0\}}$ and we will suppress the dependence of $\mathscr{S}, S$ on s when it is clear from the context. After adding the resulting inequalities and recalling that $\mathscr{P}(\bar{\mathsf{z}})=\mathscr{S}(\mathscr{S}(\bar{\mathsf{z}})-\mathsf{u}_\mathrm{d})$ and $\mathscr{P}_\tau(\bar{\zeta})=\mathscr{S}_\tau(\mathscr{S}_\tau(\bar{\zeta})-\mathsf{u}_\mathrm{d})$, we obtain that

$$\begin{aligned}&\mu\|\bar{\zeta}-\bar{\mathsf{z}}\|^2_{L^2(\Omega)}\\ &\quad\leq(\mathbb{E}[\mathscr{S}(\mathscr{S}(\bar{\mathsf{z}})-\mathsf{u}_\mathrm{d})-\mathscr{S}_\tau(\mathscr{S}_\tau(\bar{\zeta})-\mathsf{u}_\mathrm{d})],\bar{\zeta}-\bar{\mathsf{z}})_{L^2(\Omega)}\\ &\quad=\mathbb{E}[((\mathscr{S}-S)(\mathscr{S}(\bar{\mathsf{z}})-u_d),\bar{\zeta}-\bar{\mathsf{z}})_{L^2(\Omega)}+(\mathscr{S}(\bar{\mathsf{z}})-\mathscr{S}_\tau(\bar{\mathsf{z}}),\mathscr{S}_\tau(\bar{\zeta}-\bar{\mathsf{z}}))_{L^2(\Omega)}\\ &\qquad+(\mathscr{S}_\tau(\bar{\mathsf{z}}-\bar{\zeta}),\mathscr{S}_\tau(\bar{\zeta}-\bar{\mathsf{z}}))_{L^2(\Omega)}],\end{aligned}$$

where we used twice that S is a self-adjoint operator, so that the Cauchy–Schwarz inequality and Young's inequality imply

$$\begin{aligned}\|\bar{\zeta} - \bar{\mathsf{z}}\|^2_{L^2(\Omega)} &+ \mathbb{E}[\|\mathscr{S}_\tau(\bar{\mathsf{z}} - \bar{\zeta})\|^2_{L^2(\Omega)}] \\ &\leq C\mathbb{E}[\|(\mathscr{S} - S)(\mathscr{S}(\bar{\mathsf{z}}) - u_d)\|^2_{L^2(\Omega)} + \|\mathscr{S}(\bar{\mathsf{z}}) - \mathscr{S}_\tau(\bar{\mathsf{z}})\|^2_{L^2(\Omega)}].\end{aligned}$$

Next, from the definition of the $\mathbb{H}^s(\Omega)$-norm, we notice that

$$\|v\|_{L^2(\Omega)} \leq \lambda_1^{-s/2}\|v\|_{\mathbb{H}^s(\Omega)} \leq C\|v\|_{\mathbb{H}^s(\Omega)} \quad \forall v \in \mathbb{H}^s(\Omega), \tag{5.13}$$

where we have used that $\lambda_1 > 0$, and that the sequence $(\lambda_k)_{k\in\mathbb{N}}$ is nondecreasing, and therefore the constant C remains uniformly bounded when s approaches 0 or 1. As a consequence, (5.11) follows from the above inequalities and (5.7). The proof for (5.12) immediately follows after using (5.11) and (5.7). ☐

6. Discrete problem and error estimates. To get an approximation of $\mathscr{S}$ and $\mathscr{P}$, we apply the discretization from [42, 47] and [3], i.e. the truncated problem is discretized by a finite element method. We follow some of the notation from [6].

Due to the singular behavior of $\mathscr{U}$ towards the boundary Ω, we will use anistropically refined meshes. We define these meshes as follows: Let $\mathscr{T}_\Omega = \{K\}$ be a conforming and quasi-uniform triangulation of Ω, where $K \in \mathbb{R}^N$ is an element that is isoparametrically equivalent either to the unit cube or to the unit simplex in $\mathbb{R}^N$. We assume $\#\mathscr{T}_\Omega \sim M^N$. Thus, the element size $h_{\mathscr{T}_\Omega}$ fulfills $h_{\mathscr{T}_\Omega} \sim M^{-1}$. The collection of all these meshes is denoted by $\mathbb{T}_\Omega$. Furthermore, let $\mathcal{I}_\tau = \{I\}$ be a graded mesh of the interval $[0, \tau]$ in the sense that $[0, \tau] = \bigcup_{k=0}^{M-1}[\tau_k, \tau_{k+1}]$ with

$$\tau_k = \left(\frac{k}{M}\right)^\gamma \tau, \quad k = 0, \ldots, M, \quad \gamma > \frac{1}{s} > 1.$$

Now the triangulations $\mathscr{T}_\tau$ of the cylinder $\mathcal{C}_\tau$ are constructed as tensor product triangulations by means of $\mathscr{T}_\Omega$ and $\mathcal{I}_\tau$. The definitions of both imply $\#\mathscr{T}_\tau \sim M^{N+1}$. Finally, the collection of all those anisotropic meshes $\mathscr{T}_\tau$ is denoted by $\mathbb{T}$.

We denote the finite element spaces defined on the previously introduced meshes. For every $\mathscr{T}_\tau \in \mathbb{T}$, the finite element spaces $\mathbb{V}(\mathscr{T}_\tau)$ are defined by

$$\mathbb{V}(\mathscr{T}_\tau) := \{\Phi \in C^0(\overline{\mathcal{C}_\tau}) : \Phi|_T \in \mathcal{P}_1(K) \oplus \mathbb{P}_1(I) \; \forall\, T = K \times I \in \mathscr{T}_\tau,\; \Phi|_{\partial_L \mathcal{C}_\tau} = 0\}.$$

In case that K in the previous definition is a simplex, then $\mathcal{P}_1(K) = \mathbb{P}_1(K)$, the set of polynomials of degree less than or equal to 1. If K is a cube, then $\mathcal{P}_1(K)$ equals $\mathbb{Q}_1(K)$, the set of polynomials of degree at most 1 in each variable. The discretization of the truncated problem is then as follows: Given $Z \in L^2(\Omega)$ and $s \in \Xi$, find $U_h \in \mathbb{V}(\mathscr{T}_\tau)$ such that

$$\int_{\mathcal{C}_\tau} y^{\alpha_s} \left(\mathbf{A}(x, y)\nabla U_h \cdot \nabla\Phi + c(x)U_h\Phi\right) \,\mathrm{d}x\mathrm{d}y = d_s \int_\Omega Z\Phi \,\mathrm{d}x \quad \forall \Phi \in \mathbb{V}(\mathscr{T}_\tau). \tag{6.1}$$

We denote the discrete control-to-state map for (6.1) by $Z \mapsto \mathsf{S}_h(Z)$, and for a fixed Z, $\mathsf{S}_h(Z)$ is parametrized by the fractional power s. We have denoted this dependence by $s \mapsto [\mathsf{S}_h(Z)](s)$. The semidiscrete optimization problem is then given by

$$\min_{Z\in\mathbb{Z}_{\mathrm{ad}}(\mathscr{T}_\Omega)} \mathcal{R}(f(\cdot, [\mathsf{S}_h(Z)](\cdot)|_{\Omega\times\{0\}})) + g(Z), \tag{6.2}$$

where $\mathbb{Z}_{\rm ad}(\mathscr{T}_\Omega) = \mathsf{Z}_{\rm ad} \cap \mathbb{P}_\ell$ denotes the discrete admissible set of controls, and $\mathbb{P}_\ell$ denotes the set of piecewise polynomials of degree less than or equal to ℓ (note that for the error analysis in a special case in section 6.2 we will set $\ell = 0$). Existence of solutions to (6.2) follows under the assumptions of Theorem 3.1. Next, under the assumptions of Theorem 3.2, the generic optimality conditions for the optimal control $\bar{Z}$ can be derived similarly to (5.9):

(6.3)
$$\sup_{\theta\in\partial\mathcal{R}(f(\cdot,[\mathsf{S}_h(\bar{Z})](\cdot)))} \mathbb{E}\left[\theta\int_\Omega [\mathsf{P}_h(\bar{Z})](\cdot)|_{\Omega\times\{0\}}(Z'-\bar{Z})\,\mathrm{d}x\right] + \int_\Omega \nabla g(\bar{Z})(Z'-\bar{Z})\,\mathrm{d}x \ge 0$$

for all $Z' \in \mathbb{Z}_{\rm ad}(\mathscr{T}_\Omega)$, where for fixed $s \in \Xi$ and $Z \in L^2(\Omega)$, the discrete adjoint $\mathsf{P}_h = [\mathsf{P}_h(Z)](s) \in \mathbb{V}(\mathscr{T}_\tau)$ satisfies for all $\Phi \in \mathbb{V}(\mathscr{T}_\tau)$,

$$\begin{aligned}\int_{\mathcal{C}_\tau} y^{\alpha_s}\left(\mathbf{A}(x,y)\nabla\mathsf{P}_h\cdot\nabla\Phi + c(x)\mathsf{P}_h\Phi\right)\,\mathrm{d}x\mathrm{d}y \\ = d_s\int_\Omega \nabla_u f(s,[\mathsf{S}_h(\bar{Z})](s)|_{\Omega\times\{0\}})\Phi|_{\Omega\times\{0\}}\,\mathrm{d}x.\end{aligned} \tag{6.4}$$

In order to design a generic numerical scheme and to understand the structure of (6.3), for the remainder of this paper, we will assume that the risk measure $\mathcal{R}$ has the following form:

$$\mathcal{R}(X) = \inf_{t\in\mathbb{R}}\{t + \mathbb{E}[\wp(X-t)]\}, \tag{6.5}$$

where $\wp : \mathbb{R} \to \mathbb{R}$ is convex and increasing. This class of risk measures is called the optimized certainty equivalents [14, 15]. For example, if $\wp(x) = \beta^{-1}(\exp(\beta x) - 1)$, then $\mathcal{R}$ is the entropic risk. Similarly, if $\wp(x) = x$, then we recover $\mathcal{R} \equiv \mathbb{E}$.

The key advantage of the above choice for $\mathcal{R}$ in (6.5) is that we can approximate $\mathcal{R}$ by approximating the expectation $\mathbb{E}$ by $\mathbb{E}_Q$ as

$$\mathcal{R}_Q(X) = \inf_{t\in\mathbb{R}}\left\{t + \mathbb{E}_Q[\wp(X-t)]\right\},$$

where $\mathbb{E}_Q[\xi] = \sum_{k=1}^Q \omega_k \xi(s_k)$. Here $\omega_k > 0$ are the probabilities (quadrature weights) associated with the samples (quadrature points) s_k. The fully discretized optimization problem is then

$$\min_{Z\in\mathbb{Z}_{\rm ad}(\mathscr{T}_\Omega)} \mathcal{R}_Q(f(\cdot,[\mathsf{S}_h(Z)](\cdot)|_{\Omega\times\{0\}})) + g(Z). \tag{6.6}$$

Note that $\mathcal{R}_Q$ is a convex functional. The first-order necessary optimality conditions are as follows: if $\bar{Z} \in \mathbb{Z}_{\rm ad}(\mathscr{T}_\Omega)$ solves (6.6), there exist $\theta \in \partial\mathcal{R}_Q(f(\cdot,[\mathsf{S}_h(\bar{Z})](\cdot)|_{\Omega\times\{0\}}))$ such that

(6.7)
$$\int_\Omega\left(\sum_{k=1}^Q \omega_k\left[\theta(s_k)[\mathsf{P}_h(\bar{Z})](s_k)|_{\Omega\times\{0\}}\right] + \nabla g(\bar{Z})\right)(Z'-\bar{Z})\,\mathrm{d}x \ge 0 \quad \forall\, Z' \in \mathbb{Z}_{\rm ad}(\mathscr{T}_\Omega).$$

Remark 6.1. Notice that in view of Remark 3.3, if $\mathcal{R} \equiv \mathbb{E}$, then $\theta(s_k)$ in (6.7) is $\theta(s_k) = 1$ for $k = 1,\ldots,Q$. On the other hand, for the entropic risk measure,

$$\theta(s_k) = \frac{\exp(\beta f(s_k,[\mathsf{S}_h(\bar{Z})](s_k)|_{\Omega\times\{0\}}))}{\mathbb{E}_Q[\exp(\beta f(\cdot,[\mathsf{S}_h(\bar{Z})](\cdot)|_{\Omega\times\{0\}}))]} \quad \text{for} \quad k = 1,\ldots,Q.$$

Before we discuss the error estimates for the case when $\mathcal{R} = \mathbb{E}$, we provide an $\mathbb{H}^s(\Omega)$-estimate for the state variable with less regular data.

6.1. $\mathbb{H}^s(\Omega)$-error estimate for nonsmooth data. Here we derive the $\mathbb{H}^s(\Omega)$ error estimate for the state equation when z is only in $L^2(\Omega)$. Let us recall the corresponding estimate for the case that $\mathsf{z} \in \mathbb{H}^{1-s}(\Omega)$. From here on we will assume that Ω is convex polyhedral.

THEOREM 6.2 (see [42, Thm. 4.9]). *For $s \in [\varepsilon, 1-\varepsilon]$ with $\varepsilon \in (0,1)$, let $[\mathsf{S}(\mathsf{z})](s)$ and $[\mathsf{S}_h(\mathsf{z})](s)$ solve the continuous* (1.1) *and the discrete* (6.1) *equations with datum $\mathsf{z} \in \mathbb{H}^{1-s}(\Omega)$, respectively. Then the following estimate holds:*

$$\|[\mathsf{S}(\mathsf{z})](s) - [\mathsf{S}_h(\mathsf{z})](s)|_{\Omega\times\{0\}}\|_{\mathbb{H}^s(\Omega)} \le C|\log(h_{\mathscr{T}_\Omega})|^s h_{\mathscr{T}_\Omega}\|\mathsf{z}\|_{\mathbb{H}^{1-s}(\Omega)},$$

provided $\tau \sim \log(\#\mathscr{T}_\tau)$. The constant C is independent of $h_{\mathscr{T}_\Omega}$ and s but may depend on ε.

Next, we introduce an auxiliary problem which will help us to derive the estimates in the nonsmooth case: Given $\mathsf{z} \in L^2(\Omega)$, we seek $\widetilde{\mathsf{u}}_{h_{\mathscr{T}_\Omega}} \in \mathbb{H}^s(\Omega)$, which solves (weakly)

$$\mathcal{L}^s \widetilde{\mathsf{u}}_{h_{\mathscr{T}_\Omega}} = \Pi_{\mathscr{T}_\Omega}\mathsf{z} \quad \text{in } \Omega, \tag{6.8}$$

where $\Pi_{\mathscr{T}_\Omega}$ denotes the piecewise linear $L^2(\Omega)$-(orthogonal)-projection operator.

THEOREM 6.3. *For $s \in [\varepsilon, 1-\varepsilon]$ with $\varepsilon > 0$, let $[\mathsf{S}(\mathsf{z})](s)$ and $[\mathsf{S}_h(\mathsf{z})](s)$ solve the continuous* (1.1) *and the discrete* (6.1) *equations with datum $\mathsf{z} \in L^2(\Omega)$, respectively. Then the following estimate holds for $\tau \sim \log(\#\mathscr{T}_\tau)$:*

$$\|[\mathsf{S}(\mathsf{z})](s) - [\mathsf{S}_h(\mathsf{z})](s)|_{\Omega\times\{0\}}\|_{\mathbb{H}^s(\Omega)} \le C|\log(h_{\mathscr{T}_\Omega})|^s h^s_{\mathscr{T}_\Omega}\|\mathsf{z}\|_{L^2(\Omega)},$$

where the constant C is independent of $h_{\mathscr{T}_\Omega}$ and s but may depend on ε.

Proof. To keep the notation simple, we will skip the notation $\cdot|_{\Omega\times\{0\}}$, and we will use u and U_h in place of $[\mathsf{S}(z)](s)$ and $[\mathsf{S}_h(\mathsf{z})](s)$ whenever it is clear from the context.

We begin by recalling the definition of $\widetilde{\mathsf{u}}_{h_{\mathscr{T}_\Omega}}$ from (6.8). We have

$$\begin{aligned}\|\mathsf{u} - \widetilde{\mathsf{u}}_{h_{\mathscr{T}_\Omega}}\|^2_{\mathbb{H}^s(\Omega)} &= \langle \mathcal{L}^s(\mathsf{u} - \widetilde{\mathsf{u}}_{h_{\mathscr{T}_\Omega}}), (\mathsf{u} - \widetilde{\mathsf{u}}_{h_{\mathscr{T}_\Omega}})\rangle_{\mathbb{H}^{-s}(\Omega),\mathbb{H}^s(\Omega)} \\ &= \int_\Omega (\mathsf{z} - \Pi_{\mathscr{T}_\Omega}\mathsf{z})(\mathsf{u} - \widetilde{\mathsf{u}}_{h_{\mathscr{T}_\Omega}}) \le \|\mathsf{z} - \Pi_{\mathscr{T}_\Omega}\mathsf{z}\|_{\mathbb{H}^{-s}(\Omega)}\|\mathsf{u} - \widetilde{\mathsf{u}}_{h_{\mathscr{T}_\Omega}}\|_{\mathbb{H}^s(\Omega)}.\end{aligned}$$

We will estimate $\|\mathsf{z} - \Pi_{\mathscr{T}_\Omega}\mathsf{z}\|_{\mathbb{H}^{-s}(\Omega)}$ using classical interpolation. First, using [53, Lem. 41.3] we have that

$$[L^2(\Omega), H^{-1}(\Omega)]_s = [L^2(\Omega), H^1_0(\Omega)]^\star_s = \mathbb{H}^s(\Omega)^\star = \mathbb{H}^{-s}(\Omega),$$

where $\star$ denotes the dual space. Now from [29, eq. 1.115] and [29, Prop. 1.133] we have that

$$\|\mathsf{z} - \Pi_{\mathscr{T}_\Omega}\mathsf{z}\|_{L^2(\Omega)} \le C\|\mathsf{z}\|_{L^2(\Omega)} \quad \text{and} \quad \|\mathsf{z} - \Pi_{\mathscr{T}_\Omega}\mathsf{z}\|_{H^{-1}(\Omega)} \le Ch_{\mathscr{T}_\Omega}\|\mathsf{z}\|_{L^2(\Omega)}.$$

Using the above mentioned interpolation, we obtain that

$$\|\mathsf{z} - \Pi_{\mathscr{T}_\Omega}\mathsf{z}\|_{\mathbb{H}^{-s}(\Omega)} \le Ch^s_{\mathscr{T}_\Omega}\|\mathsf{z}\|_{L^2(\Omega)}.$$

As a result,

$$\|\mathsf{u} - \widetilde{\mathsf{u}}_{h_{\mathscr{T}_\Omega}}\|_{\mathbb{H}^s(\Omega)} \le \|\mathsf{z} - \Pi_{\mathscr{T}_\Omega}\mathsf{z}\|_{\mathbb{H}^{-s}(\Omega)} \le Ch^s_{\mathscr{T}_\Omega}\|\mathsf{z}\|_{L^2(\Omega)}.$$

Here the constant C is independent of $h_{\mathscr{T}_\Omega}$ and s.

Next, let $\widetilde{U}_h$ be the solution to the discrete truncated problem (6.1) with datum $\Pi_{\mathscr{T}_\Omega}\mathsf{z}$. Then using Theorem 6.2 we have

$$\|\widetilde{\mathsf{u}}_{h_{\mathscr{T}_\Omega}} - \widetilde{U}_h\|_{\mathbb{H}^s(\Omega)} \le C|\log(h_{\mathscr{T}_\Omega})|^s h_{\mathscr{T}_\Omega}\|\Pi_{\mathscr{T}_\Omega}\mathsf{z}\|_{\mathbb{H}^{1-s}(\Omega)} \le C|\log(h_{\mathscr{T}_\Omega})|^s h^s_{\mathscr{T}_\Omega}\|\mathsf{z}\|_{L^2(\Omega)},$$

where the last inequality is due to the inverse estimate (after using an interpolation argument). Indeed, from the classical inverse estimates, combined with the stability of $\Pi_{\mathscr{T}_\Omega}$ in $L^2(\Omega)$, we obtain

$$\|\Pi_{\mathscr{T}_\Omega}\mathsf{z}\|_{H^1_0(\Omega)} \le Ch^{-1}_{\mathscr{T}_\Omega}\|\mathsf{z}\|_{L^2(\Omega)} \quad \text{and} \quad \|\Pi_{\mathscr{T}_\Omega}\mathsf{z}\|_{L^2(\Omega)} \le C\|\mathsf{z}\|_{L^2(\Omega)},$$

which after recalling that $[L^2(\Omega), H^1_0(\Omega)]_{1-s} = \mathbb{H}^{1-s}(\Omega)$ yields $\|\Pi_{\mathscr{T}_\Omega}\mathsf{z}\|_{\mathbb{H}^{1-s}(\Omega)} \le Ch^{s-1}_{\mathscr{T}_\Omega}\|\mathsf{z}\|_{L^2(\Omega)}$ as asserted.

From the definition of U_h and $\widetilde{U}_h$ we have that

$$\|\widetilde{U}_h - U_h\|_{\mathbb{H}^s(\Omega)} \le C\|\mathsf{z} - \Pi_{\mathscr{T}_\Omega}\mathsf{z}\|_{\mathbb{H}^{-s}(\Omega)} \le Ch^s_{\mathscr{T}_\Omega}\|\mathsf{z}\|_{L^2(\Omega)},$$

where we have used the estimate for $\|\mathsf{z} - \Pi_{\mathscr{T}_\Omega}\mathsf{z}\|_{\mathbb{H}^{-s}(\Omega)}$ from the above interpolation argument, from where

$$\begin{aligned}\|\mathsf{u} - U_h\|_{\mathbb{H}^s(\Omega)} &\le \|\mathsf{u} - \widetilde{\mathsf{u}}_{h_{\mathscr{T}_\Omega}}\|_{\mathbb{H}^s(\Omega)} + \|\widetilde{\mathsf{u}}_{h_{\mathscr{T}_\Omega}} - \widetilde{U}_h\|_{\mathbb{H}^s(\Omega)} + \|\widetilde{U}_h - U_h\|_{\mathbb{H}^s(\Omega)} \\ &\le C|\log h_{\mathscr{T}_\Omega}|^s h^s_{\mathscr{T}_\Omega}\|\mathsf{z}\|_{L^2(\Omega)}.\end{aligned}$$

This completes the proof. ∎

6.2. Error estimates for a special case. Next, we derive the discretization error estimates for the setting in Theorem 5.5. We assume that Z_{ad} is defined as

$$\mathsf{Z}_{\text{ad}} := \{z \in L^2(\Omega) \mid \mathsf{a} \le z \le \mathsf{b} \text{ a.e. in } \Omega\}, \tag{6.9}$$

where $\mathsf{a}, \mathsf{b} \in \mathbb{R}$, with $\mathsf{a} < \mathsf{b}$, are given. Notice that one can easily consider more generic control bounds than the constants a and b. For instance, when $\mathsf{a}, \mathsf{b} \in L^2(\Omega)$, we can use the average of a, b on each element K to approximate the control constraints. Since $\nabla g(\bar{\mathsf{z}}) = \mu\bar{\mathsf{z}}$, the projection formula in (3.6) after setting $\gamma = 1/\mu$ becomes

$$\bar{\mathsf{z}} = \mathbb{P}_{\mathsf{Z}_{\text{ad}}}\left(-\frac{1}{\mu}\mathbb{E}\left[[\mathsf{P}(\bar{\mathsf{z}})](\cdot)\right]\right). \tag{6.10}$$

Before we proceed further, we need to understand the regularity of the optimal solution. In the deterministic setting, one can use a boot-strap argument to improve the regularity for $\bar{\mathsf{z}}$ using the regularity of the optimal adjoint state and then improve the regularity of the optimal state and the optimal control. However, owing to the relation (6.10) this is not as easy as it may appear. The key issue is that the expectation is being carried out with respect to the random variable s which in turn is the exponent for the Sobolev space $\mathbb{H}^s(\Omega)$. The latter determines the spatial regularity for the optimal adjoint state. With the help of a simple example, we next illustrate that in general it is not possible to improve the regularity of $\bar{\mathsf{z}}$.

Remark 6.4 (regularity of the control).

(i) Let $t \geq 0$, and suppose u_d is independent of $s \in \Xi$ and is sufficiently smooth. Then, since the eigenfunctions φ_k form an orthogonal basis of $L^2(\Omega)$ and $H_0^1(\Omega)$, we have that

$$\begin{aligned}\|\mathbb{E}\left[[\mathsf{P}(\bar{\mathsf{z}})](\cdot)\right]\|_{\mathbb{H}^t(\Omega)}^2 &= \sum_{k=1}^{\infty} \lambda_k^t (\mathbb{E}\left[[\mathsf{P}(\bar{\mathsf{z}})](\cdot)\right], \varphi_k)^2 \\ &= \sum_{k=1}^{\infty} \lambda_k^t \left(\mathbb{E}\left[\sum_{j=1}^{\infty} \lambda_j^{-\cdot} \left(\lambda_j^{-\cdot} (\bar{\mathsf{z}}, \varphi_j)_{L^2(\Omega)} - (\mathsf{u}_d, \varphi_j)_{L^2(\Omega)} \right) \varphi_j \right], \varphi_k \right)^2 \\ &= \sum_{k=1}^{\infty} \lambda_k^t \left(\mathbb{E}\left[\lambda_k^{-\cdot} \left(\lambda_k^{-\cdot} (\bar{\mathsf{z}}, \varphi_k)_{L^2(\Omega)} - (\mathsf{u}_d, \varphi_k)_{L^2(\Omega)} \right) \right] \right)^2,\end{aligned}$$

where we have assumed that we can switch the order of summation and integration. For uniformly distribution s, the quantities, $\mathbb{E}[\lambda_k^{-2\cdot}]$ and $\mathbb{E}[\lambda_k^{-\cdot}]$, are

$$\mathbb{E}[\lambda_k^{-2\cdot}] = \int_0^1 \lambda_k^{-2s}\, \mathrm{d}s = \frac{1}{2\log(\lambda_k)} \left(1 - \frac{1}{\lambda_k^2}\right), \quad \mathbb{E}[\lambda_k^{-\cdot}] = \frac{1}{\log(\lambda_k)} \left(1 - \frac{1}{\lambda_k}\right).$$

Thus, $\mathbb{E}\left[[\mathsf{P}(\bar{\mathsf{z}})](\cdot)\right]$ is expected to have logarithmic regularity.

(ii) If $\mathsf{u}_{\mathrm{d}} : \Xi \to H^1(\Omega)$ is continuous, then $s \mapsto [\mathsf{P}(\bar{\mathsf{z}})](s) \in L^\infty(\Xi, \mathcal{B}, P; \mathbb{H}^\beta(\Omega))$, where $\beta = \min\{4s, 1+s\}$. In addition, if $P((0,\varepsilon)) = 0$ with $\varepsilon > 0$, then we deduce $\bar{\mathsf{z}} \in H^1(\Omega)$ by means of a bootstrapping argument using (6.10), [34, Thm. A.1], and

$$\|\mathbb{E}\left[[\mathsf{P}(\bar{\mathsf{z}})](\cdot)\right]\|_{\mathbb{H}^t(\Omega)} \leq \mathbb{E}[\|[\mathsf{P}(\bar{\mathsf{z}})](\cdot)\|_{\mathbb{H}^t(\Omega)}]$$

with $t \geq 0$ which is due to Theorem 3.7.6 in [33]. Of course, for $t \geq 1/2$ we have to take care of compatibility conditions for a and b, which hold if we assume $\mathsf{a} \leq 0 \leq \mathsf{b}$. Notice that in general we cannot expect higher regularity for $\bar{\mathsf{z}}$ as in the deterministic case [3].

For the optimal control, we use a piecewise constant discretization, i.e., the discrete controls belong to

$$\mathbb{Z}_{\mathrm{ad}}(\mathscr{T}_\Omega) := \mathsf{Z}_{\mathrm{ad}} \cap \mathbb{P}_0(\mathscr{T}_\Omega), \tag{6.11}$$

where by $\mathbb{P}_0(\mathscr{T}_\Omega)$ we denote the space of piecewise constant functions. In addition, we define the piecewise constant $L^2(\Omega)$-projection $\Pi^0_{\mathscr{T}_\Omega} : L^2(\Omega) \to \mathbb{P}_0(\mathscr{T}_\Omega)$ as

$$(\mathsf{z} - \Pi^0_{\mathscr{T}_\Omega}\mathsf{z}, Z)_{L^2(\Omega)} = 0 \quad \forall Z \in \mathbb{Z}_{\mathrm{ad}}(\mathscr{T}_\Omega). \tag{6.12}$$

Notice that $\|\Pi^0_{\mathscr{T}_\Omega}\mathsf{z}\|_{L^2(\Omega)} \leq c\|\mathsf{z}\|_{L^2(\Omega)}$ and for $\mathsf{z} \in H^1(\Omega)$, we have $\|\mathsf{z} - \Pi^0_{\mathscr{T}_\Omega}\mathsf{z}\|_{L^2(\Omega)} \leq c h_{\mathscr{T}_\Omega}\|\mathsf{z}\|_{H^1(\Omega)}$, where the constant c is independent of $h_{\mathscr{T}_\Omega}$ and z on both occasions. From (6.12) it follows that $\Pi^0_{\mathscr{T}_\Omega}\mathsf{z}|_K = |K|^{-1}\int_K \mathsf{z}\,\mathrm{d}x$. Moreover, due to a and b being constants, we conclude that $\Pi^0_{\mathscr{T}_\Omega}\mathsf{z} \in \mathbb{Z}_{\mathrm{ad}}(\mathscr{T}_\Omega)$ and as a result, $\Pi^0_{\mathscr{T}_\Omega} : L^2(\Omega) \to \mathbb{Z}_{\mathrm{ad}}(\mathscr{T}_\Omega)$ is well defined.

In view of Remark 6.4, we state a general error estimate for the optimal control which only requires $L^2(\Omega)$ regularity of $\bar{\mathsf{z}}$ and $\mathsf{u}_{\mathrm{d}}(s)$. A few special cases will be discussed in what follows.

THEOREM 6.5. *Let the problem setting and assumptions of Theorem* 5.5 *hold. Moreover, let* $\bar{\mathsf{z}}$ *solve* (4.3) *with* $\mathcal{R} = \mathbb{E}$*, and let* $\bar{Z}$ *solve* (6.6) *with* $\mathbb{Z}_{\rm ad}(\mathscr{T}_\Omega)$ *as in* (6.11) *and* $\mathcal{R}_Q = \mathbb{E}_Q$*. Then*

$$\|\bar{\mathsf{z}} - \bar{Z}\|_{L^2(\Omega)} \le \mathcal{E}_{\rm fem} + \mathcal{E}_{\rm quad} \tag{6.13}$$

and

$$\mathbb{E}_Q[\|[\mathsf{S}(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}} - [\mathsf{S}_h(\bar{Z})](\cdot)|_{\Omega\times\{0\}}\|_{\mathbb{H}^\cdot(\Omega)}] \le \mathcal{E}_{\rm fem} + \mathcal{E}_{\rm quad}, \tag{6.14}$$

where

$$\begin{aligned}
\mathcal{E}_{\rm fem} := C\big(&\mathbb{E}_Q\big[\|[\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{S}_h(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}}\|_{L^2(\Omega)}\\
&+ \|[(\mathscr{S} - \mathsf{S}_h)(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_d)](\cdot)|_{\Omega\times\{0\}}\|_{L^2(\Omega)}\\
&+ \|[\mathscr{P}(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}} - \Pi^0_{\mathscr{T}_\Omega}[\mathscr{P}(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}}\|_{L^2(\Omega)}\big] + \|\bar{\mathsf{z}} - \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}}\|_{L^2(\Omega)}\big),\\
\mathcal{E}_{\rm quad} := \big\|&\mathbb{E}\left[[\mathscr{P}(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}}\right] - \mathbb{E}_Q\left[[\mathscr{P}(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}}\right]\big\|_{L^2(\Omega)}
\end{aligned}$$

with a positive constant C *independent of* $h_{\mathcal{T}_\Omega}$ *and* s*.*

Remark 6.6 (rate of convergence). The estimators $\mathcal{E}_{\rm fem}$ and $\mathcal{E}_{\rm quad}$ denote the spatial and quadrature approximation errors, respectively. The precise estimate on $\mathcal{E}_{\rm quad}$ depends on the specific quadrature rule, but in view of the regularity result in Proposition 2.2 this error can be easily controlled, for instance, using Gauss quadrature. On the other hand, we can directly use Theorem 6.3 to estimate all the terms in $\mathcal{E}_{\rm fem}$, except the last two, which represent the L^2-projection approximation errors. For $s \in \Xi$, we have $[\mathscr{P}(\bar{\mathsf{z}})](s)|_{\Omega\times\{0\}} \in \mathbb{H}^s(\Omega)$ and therefore $\|[\mathscr{P}(\bar{\mathsf{z}})](s)|_{\Omega\times\{0\}} - \Pi^0_{\mathscr{T}_\Omega}[\mathscr{P}(\bar{\mathsf{z}})](s)|_{\Omega\times\{0\}}\|_{L^2(\Omega)} \le Ch^s\|[\mathscr{P}(\bar{\mathsf{z}})](s)|_{\Omega\times\{0\}}\|_{\mathbb{H}^s(\Omega)}$. Since we have not established the regularity of $\bar{\mathsf{z}}$ (cf. Remark 6.4), we can only assert convergence (without rates) of $\|\bar{\mathsf{z}} - \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}}\|_{L^2(\Omega)}$ in general.

Proof. For simplicity, we will skip the notation $\cdot|_{\Omega\times\{0\}}$ and we will suppress the dependence of $\mathscr{S}, \mathscr{S}_\tau, \mathsf{S}_h, \mathsf{u}_d$ on s when it is clear from the context.

Since $\mathbb{Z}_{\rm ad}(\mathscr{T}_\Omega) \subset \mathsf{Z}_{\rm ad}$, by setting $\mathsf{z}' = \bar{Z}$ in the variational inequality (4.5) and using that $g(\bar{\mathsf{z}}) = \mu\bar{\mathsf{z}}$, we obtain that

$$(\mathbb{E}[\mathscr{P}(\bar{\mathsf{z}})] + \mu\bar{\mathsf{z}}, \bar{Z} - \bar{\mathsf{z}})_{L^2(\Omega)} \ge 0.$$

Next, setting $Z' = \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} \in \mathbb{Z}_{\rm ad}(\mathscr{T}_\Omega)$ in the corresponding variational inequality for the discrete problem (6.6), we obtain that

$$\Big(\mathbb{E}_Q[\mathsf{P}_h(\bar{Z})] + \mu\bar{Z}, \bar{\mathsf{z}} - \bar{Z}\Big)_{L^2(\Omega)} + \Big(\mathbb{E}_Q[\mathsf{P}_h(\bar{Z})] + \mu\bar{Z}, \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}}\Big)_{L^2(\Omega)} \ge 0,$$

where we have added and subtracted $\bar{\mathsf{z}}$. Adding the resulting expressions and rearranging terms, we arrive at

$$\begin{aligned}
&\mu\|\bar{\mathsf{z}} - \bar{Z}\|^2_{L^2(\Omega)}\\
&\le \Big(\mathbb{E}[\mathscr{P}(\bar{\mathsf{z}})] - \mathbb{E}_Q[\mathsf{P}_h(\bar{Z})], \bar{Z} - \bar{\mathsf{z}}\Big)_{L^2(\Omega)} + \Big(\mathbb{E}_Q[\mathsf{P}_h(\bar{Z})] + \mu\bar{Z}, \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}}\Big)_{L^2(\Omega)}\\
&= \Big(\mathbb{E}[\mathscr{P}(\bar{\mathsf{z}})] - \mathbb{E}_Q[\mathscr{P}(\bar{\mathsf{z}})], \bar{Z} - \bar{\mathsf{z}}\Big)_{L^2(\Omega)} + \Big(\mathbb{E}_Q[\mathscr{P}(\bar{\mathsf{z}}) - \mathsf{P}_h(\bar{Z})], \bar{Z} - \bar{\mathsf{z}}\Big)_{L^2(\Omega)}\\
&\quad + \Big(\mathbb{E}_Q[\mathsf{P}_h(\bar{Z})] + \mu\bar{Z}, \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}}\Big)_{L^2(\Omega)} =: \mathrm{I} + \mathrm{II} + \mathrm{III}.
\end{aligned}$$

By applying the Cauchy–Schwarz inequality we get

$$|\mathrm{I}| \leq \|\mathbb{E}[\mathscr{P}(\bar{\mathsf{z}})] - \mathbb{E}_Q[\mathscr{P}(\bar{\mathsf{z}})]\|_{L^2(\Omega)} \|\bar{\mathsf{z}} - \bar{Z}\|_{L^2(\Omega)}.$$

For the estimate of II, we can proceed in a manner similar to the proof of Theorem 5.5. Indeed, as $\mathsf{P}_h(\bar{Z}) = \mathsf{S}_h(\mathsf{S}_h(\bar{Z}) - \mathsf{u}_d)$ and $\mathscr{P}(\bar{\mathsf{z}}) = \mathscr{S}(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_\mathrm{d})$, we obtain by adding and subtracting $\mathsf{S}_h(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_\mathrm{d})$

$$\begin{aligned}
\mathrm{II} &= \mathbb{E}_Q[(\mathscr{S}(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_\mathrm{d}) - \mathsf{S}_h(\mathsf{S}_h(\bar{Z}) - \mathsf{u}_d), \bar{Z} - \bar{\mathsf{z}})_{L^2(\Omega)}] \\
&= \mathbb{E}_Q[((\mathscr{S} - \mathsf{S}_h)(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_\mathrm{d}), \bar{Z} - \bar{\mathsf{z}})_{L^2(\Omega)}] + \mathbb{E}_Q[(\mathsf{S}_h(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{S}_h(\bar{Z})), \bar{Z} - \bar{\mathsf{z}})_{L^2(\Omega)}] \\
&= \mathbb{E}_Q[((\mathscr{S} - \mathsf{S}_h)(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_\mathrm{d}), \bar{Z} - \bar{\mathsf{z}})_{L^2(\Omega)}] + \mathbb{E}_Q[(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{S}_h(\bar{Z}), \mathsf{S}_h(\bar{Z} - \bar{\mathsf{z}}))_{L^2(\Omega)}].
\end{aligned}$$

For the first term, we simply get

$$\mathbb{E}_Q[((\mathscr{S} - \mathsf{S}_h)(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_\mathrm{d}), \bar{Z} - \bar{\mathsf{z}})_{L^2(\Omega)}] \leq \mathbb{E}_Q[\|(\mathscr{S} - \mathsf{S}_h)(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_\mathrm{d})\|_{L^2(\Omega)}] \|\bar{Z} - \bar{\mathsf{z}}\|_{L^2(\Omega)}.$$

In case of the second one, we add and subtract $\mathsf{S}_h(\bar{\mathsf{z}})$ such that

$$\begin{aligned}
\mathbb{E}_Q[(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{S}_h(\bar{Z}), \mathsf{S}_h(\bar{Z} - \bar{\mathsf{z}}))_{L^2(\Omega)}] &\leq \mathbb{E}_Q[(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{S}_h(\bar{\mathsf{z}}), \mathsf{S}_h(\bar{Z} - \bar{\mathsf{z}}))_{L^2(\Omega)}] \\
&\leq \mathbb{E}_Q[\|\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{S}_h(\bar{\mathsf{z}})\|_{L^2(\Omega)} \|\mathsf{S}_h(\bar{Z} - \bar{\mathsf{z}})\|_{L^2(\Omega)}].
\end{aligned}$$

Next, we use that S_h maps stable (in terms of $h_{\mathcal{T}_\Omega}$ and s) from $L^2(\Omega)$ to $L^2(\Omega)$, which can be seen from (5.13), (5.8), and (6.1). Thus, we get

$$\mathbb{E}_Q[(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{S}_h(\bar{Z}), \mathsf{S}_h(\bar{Z} - \bar{\mathsf{z}}))_{L^2(\Omega)}] \leq \mathbb{E}_Q[\|\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{S}_h(\bar{\mathsf{z}})\|_{L^2(\Omega)}] \|\bar{\mathsf{z}} - \bar{Z}\|_{L^2(\Omega)}.$$

Collecting the previous estimates yields

$$|\mathrm{II}| \leq C(\mathbb{E}_Q[\|(\mathscr{S} - \mathsf{S}_h)(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_d)\|_{L^2(\Omega)}] + \mathbb{E}_Q[\|(\mathscr{S} - \mathsf{S}_h)(\bar{\mathsf{z}})\|_{L^2(\Omega)}]) \|\bar{\mathsf{z}} - \bar{Z}\|_{L^2(\Omega)}.$$

It then remains to estimate III. Using the orthogonality of $\Pi^0_{\mathscr{T}_\Omega}$, we have $\mu(\bar{Z}, \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}})_{L^2(\Omega)} = 0$. Since $\mathsf{P}_h(\bar{Z}) = \mathsf{S}_h(\mathsf{S}_h(\bar{Z}) - \mathsf{u}_d)$, we obtain that

$$\begin{aligned}
\mathrm{III} &= (\mathbb{E}_Q[\mathsf{S}_h(\mathsf{S}_h(\bar{Z}) - \mathsf{u}_d)], \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}})_{L^2(\Omega)} \\
&= \mathbb{E}_Q[(\mathsf{S}_h(\mathsf{S}_h(\bar{Z} - \bar{\mathsf{z}})), \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}})_{L^2(\Omega)}] + \mathbb{E}_Q[(\mathsf{S}_h(\mathsf{S}_h(\bar{\mathsf{z}}) - \mathsf{u}_d), \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}})_{L^2(\Omega)}] \\
&=: \mathrm{III}_1 + \mathrm{III}_2,
\end{aligned}$$

where we have added and subtracted $(\mathbb{E}_Q[\mathsf{S}_h(\mathsf{S}_h(\bar{\mathsf{z}}) - \mathsf{u}_d)], \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}})_{L^2(\Omega)}$. To estimate III_1, we use that S_h as a map from $L^2(\Omega)$ to $L^2(\Omega)$ is stable. By this we arrive at

$$|\mathrm{III}_1| \leq C\|\Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}}\|_{L^2(\Omega)} \|\bar{\mathsf{z}} - \bar{Z}\|_{L^2(\Omega)}.$$

Next, we will estimate III_2. By adding and subtracting $(\mathbb{E}_Q[\mathsf{S}_h(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_d)], \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}})_{L^2(\Omega)}$ to III_2, we obtain that

$$\begin{aligned}
|\mathrm{III}_2| &= \Big|\mathbb{E}_Q[(\mathsf{S}_h(\mathsf{S}_h(\bar{\mathsf{z}}) - \mathscr{S}(\bar{\mathsf{z}})), \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}})_{L^2(\Omega)}] \\
&\qquad + \mathbb{E}_Q[(\mathsf{S}_h(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_d), \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}})_{L^2(\Omega)}]\Big| \\
&\leq C\mathbb{E}_Q[\|\mathsf{S}_h(\bar{\mathsf{z}}) - \mathscr{S}(\bar{\mathsf{z}})\|_{L^2(\Omega)}] \|\Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}}\|_{L^2(\Omega)} \\
&\qquad + \Big|\mathbb{E}_Q[((\mathsf{S}_h - \mathscr{S})(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_\mathrm{d}), \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}})_{L^2(\Omega)}]\Big| \\
&\qquad + \Big|\mathbb{E}_Q[(\mathscr{S}(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_\mathrm{d}), \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}})_{L^2(\Omega)}]\Big|,
\end{aligned}$$

where we have used the stability of S_h once again and where we have added and subtracted $(\mathbb{E}_Q[\mathscr{S}(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_d)], \Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}}))_{L^2(\Omega)}$ to the second term. Subsequently, by using $\mathscr{P}(\bar{\mathsf{z}}) = \mathscr{S}(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_\mathrm{d})$, the orthogonality of the $L^2(\Omega)$-projection, and by applying the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}|\mathrm{III}_2| \le C\Big(&\mathbb{E}_Q[\|\mathsf{S}_h(\bar{\mathsf{z}}) - \mathscr{S}(\bar{\mathsf{z}})\|_{L^2(\Omega)}] + \mathbb{E}_Q[\|(\mathsf{S}_h - \mathscr{S})(\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{u}_\mathrm{d})\|_{L^2(\Omega)}] \\ &+ \|\mathbb{E}_Q[\bar{\mathscr{P}} - \Pi^0_{\mathscr{T}_\Omega}\bar{\mathscr{P}}]\|_{L^2(\Omega)}\Big)\|\Pi^0_{\mathscr{T}_\Omega}\bar{\mathsf{z}} - \bar{\mathsf{z}}\|_{L^2(\Omega)}.\end{aligned}$$

Collecting all the estimates, we arrive at (6.13) due to Young's inequality.

It remains to show (6.14). Towards this end, we have

$$\begin{aligned}\mathbb{E}_Q[\|\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{S}_h(\bar{Z})\|_{\mathbb{H}^s(\Omega)}] &\le \mathbb{E}_Q[\|\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{S}_h(\bar{\mathsf{z}})\|_{L^2(\Omega)}] + \mathbb{E}_Q[\|\mathsf{S}_h(\bar{\mathsf{z}} - \bar{Z})\|_{L^2(\Omega)}] \\ &\le \mathbb{E}_Q[\|\mathscr{S}(\bar{\mathsf{z}}) - \mathsf{S}_h(\bar{\mathsf{z}})\|_{L^2(\Omega)}] + \mathbb{E}_Q[\|(\bar{\mathsf{z}} - \bar{Z})\|_{L^2(\Omega)}],\end{aligned}$$

where we have used the discrete stability. Then combined with the previous estimate (6.13) we obtain (6.14). This concludes the proof. ∎

COROLLARY 6.7. *In addition to the assumptions of Theorem* 6.5*, if* $P((0,\varepsilon)) = P((1-\varepsilon,1)) = 0$ *for some* $\varepsilon > 0$*,* $\mathsf{u}_d : \Xi \to H^1(\Omega)$ *is continuous, and* $\mathsf{a} \le 0 \le \mathsf{b}$*, then the following estimate holds:*

(6.15)
$$\begin{aligned}\|\bar{\mathsf{z}} - \bar{Z}\|_{L^2(\Omega)} &+ \mathbb{E}_Q[\|[\mathscr{S}(\bar{\mathsf{z}})](\cdot)|_{\Omega\times\{0\}} - [\mathsf{S}_h(\bar{Z})](\cdot)|_{\Omega\times\{0\}}\|_{\mathbb{H}^\cdot(\Omega)}] \\ &\le Ch_{\mathscr{T}_\Omega}\mathbb{E}_Q\big[|\log(h_{\mathscr{T}_\Omega})|^s\big] + \big\|\mathbb{E}[[\mathscr{P}(\bar{\mathsf{z}})](\cdot)] - \mathbb{E}_Q[[\mathscr{P}(\bar{\mathsf{z}})](\cdot)]\big\|_{L^2(\Omega)},\end{aligned}$$

provided $\tau \sim \log(\#\mathscr{T}_\tau)$*.*

Proof. The proof is a consequence of Theorem 6.5, Remark 6.4(ii) and Theorem 6.2. ∎

7. Numerics.

We implement the optimal control problem in the iFEM library [23] within the MATLAB environment. The stiffness matrix is assembled exactly, and the forcing term is computed by a quadrature formula that is exact for polynomials of degree 4. The resulting state and adjoint systems are solved using backslash in MATLAB. In our numerical examples, we let $n = 2$, $\Omega = (0,1)^2$, $c \equiv 0$, and $A \equiv 1$. The eigenvalues and eigenfunctions of $\mathcal{L}$ are

$$\lambda_{k,l} = \pi^2(k^2 + l^2), \quad \varphi_{k,l}(x_1, x_2) = \sin(k\pi x_1)\sin(l\pi x_2), \quad k, l \in \mathbb{N}.$$

7.1. Example 1.

We consider

$$\mathcal{R} \equiv \mathbb{E}, \quad f(\cdot, \mathsf{u}) = \frac{1}{2}\|\mathsf{u} - \mathsf{u}_d(\cdot)\|^2_{L^2(\Omega)}, \quad g(\mathsf{z}) = \frac{\mu}{2}\|\mathsf{z}\|^2_{L^2(\Omega)},$$

where $\mu = 1$ and u_d is a given desired state exactly specified below. We discretize $\mathcal{R} \equiv \mathbb{E}$ using Gauss–Legendre quadrature of order 5. In order to be able to state an exact solution, we modify the state equation. More precisely, we consider

$$\mathcal{L}^s\mathsf{u} = \mathsf{f} + \bar{\mathsf{z}} \text{ in } \Omega, \quad \mathsf{u} = 0 \text{ on } \partial\Omega,$$

where we set $\mathsf{f} = \lambda^s_{2,2}\sin(2\pi x_1)\sin(2\pi x_2) - \bar{\mathsf{z}}$. Then $\mathsf{S}(\bar{\mathsf{z}}) = \sin(2\pi x_1)\sin(2\pi x_2)$. Letting $\mathsf{P}(\bar{\mathsf{z}}) = -\sin(2\pi x_1)\sin(2\pi x_2)$, we obtain that $\mathsf{u}_d = (1 + \lambda^s_{2,2})\sin(2\pi x_1)\sin(2\pi x_2)$. In view of the projection formula, we notice that $\bar{\mathsf{z}} = \min\{\mathsf{b}, \max\{\mathsf{a}, -\mathsf{P}(\bar{\mathsf{z}})\}\}$, where we set $\mathsf{a} = 0$ and $\mathsf{b} = 0.5$. Figure 1 shows the rate of convergence for the control as we refine the mesh in space. Clearly we obtain the theoretically expected rate of convergence.

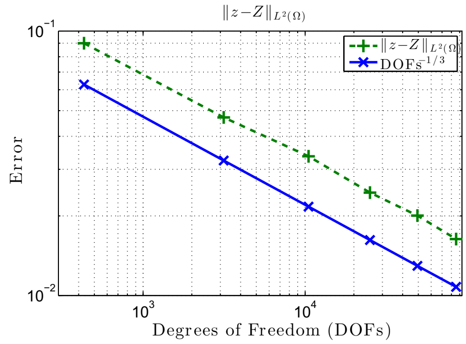

Fig. 1. *The panel illustrates the rate of convergence in space for the control* z *for a fixed Gauss quadrature rule of order* 5. *We recover the optimal* $|\mathrm{DoF}|^{-1/3}$ *rate of convergence.*

7.2. Example 2. In our second example, we compare the behavior of the cost functional when $\mathcal{R}(X) = \mathbb{E}[X]$ and $\mathcal{R}(X) = \beta^{-1}\log \mathbb{E}[\exp(\beta X)]$, respectively. We let

$$f(\cdot, \mathsf{u}) = \frac{1}{2}\|\mathsf{u} - \mathsf{u}_d\|^2_{L^2(\Omega)}, \quad g(\mathsf{z}) = \frac{\mu}{2}\|\mathsf{z}\|^2_{L^2(\Omega)}$$

with $\mu = 0.1$, $\mathsf{u}_d \equiv 1$, $\beta = 5$, $\mathsf{a} = 0$, $\mathsf{b} = 2$, and $\Xi = [0.05, 0.95]$. We use the Gauss–Legendre quadrature rule of order 5 and $\#\mathscr{T}_\tau = 3146$. Figure 2 shows the cumulative distribution function (CDF, left) and the probability density function (PDF, right) for $f(\cdot, [\mathsf{S}_h(\bar{Z})](\cdot))$ when sampled at 100,000 uniformly distributed random numbers. Moreover, Figure 3 shows the CDF and PDF for $f(\cdot, [\mathsf{S}_h(\bar{Z})](\cdot)) + g(\bar{Z})$. Figure 4 shows the difference between the controls for the two cases. The entropic risk measure is more conservative than $\mathbb{E}$ in the sense that

$$\beta^{-1}\log \mathbb{E}[\exp(\beta X)] > \mathbb{E}[X] \quad \forall \text{ nonconstant } X \in L^p(\Xi, \mathcal{B}, P).$$

As seen in Figure 2, the control computed using the entropic risk results in less variability of $f(\cdot, \mathsf{u})$; i.e., the support of the associated PDF is smaller than the support of the PDF for the expected value. Effectively, the entropic risk control reduces variability in the optimal objective function value. This is not without cost. Figure 2 demonstrates that the state-only objective function CDF corresponding to $\mathcal{R} = \mathbb{E}$ dominates the CDF corresponding to the entropic risk. However, when considering the total objective function, the entropic risk control appears to outperform the control for $\mathcal{R} = \mathbb{E}$ with respect to the approximately 45% largest scenarios (cf. Figure 3).

8. Conclusion. This paper has introduced a new class of optimal control problems for fractional diffusion equations where the fractional exponent s is taken as a random variable. Since the order of the fractional Sobolev space is itself now a random variable, the existing techniques to analyze and solve such optimal control problems are not directly applicable. We have introduced a risk-averse optimization framework for this class of optimal control problems, and we have shown existence of solutions as well as rigorously derived the first-order optimality conditions. We employ quadrature to approximate the random exponent and the finite element method to discretize in space. We have also derived the rate of convergence for the fully discrete optimal control problem to the continuous one in the risk-neural case.

We have considered optimal control problems governed by the most basic fractional PDE. As we mentioned in the introduction, one of our motivations is to extend

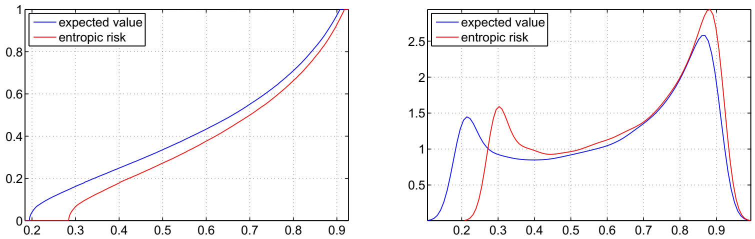

FIG. 2. *Left panel shows the cumulative distribution function (CDF) for* $\|[\mathsf{S}_h(\bar{Z})](\cdot)-\mathsf{u}_d\|^2_{L^2(\Omega)}$. *Right panel shows probability density function (PDF).*

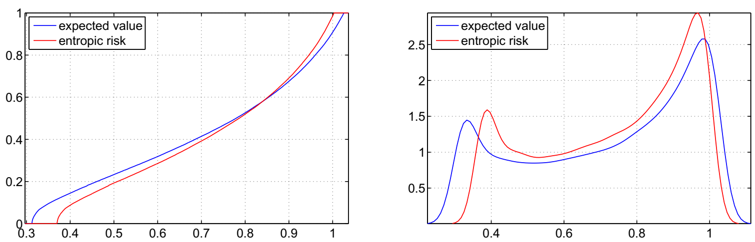

FIG. 3. *Left panel shows the cumulative distribution function (CDF) for* $\|[\mathsf{S}_h(\bar{Z})](\cdot)-\mathsf{u}_d\|^2_{L^2(\Omega)}+\mu\|\bar{Z}\|^2_{L^2(\Omega)}$. *Right panel shows probability density function (PDF).*

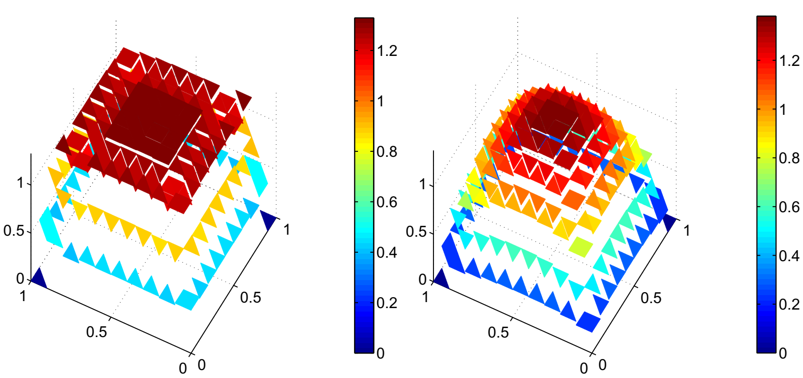

FIG. 4. *Left panel shows the control in case* $\mathcal{R}\equiv\mathbb{E}$ *and the right panel shows the control in case* $\mathcal{R}$ *is the entropic risk.*

our work to the fractional Helmholtz PDE [58]. In order to fully understand this problem, we must tailor estimates of the distribution of s to the specific application; see [58] for an initial discussion on this topic. In view of [7], for the fractional Helmholtz PDE and imaging science applications, it is also of interest to consider a spatial varying s, i.e., $s(x)$. In this case, the current problem becomes significantly more complicated, as the fractional exponent is a spatially varying random field.

REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Applied Mathematics 55, Superintendent of Documents, U.S. Govt. Printing Office, Washington, D.C., 1964.

[2] H. Antil and S. Bartels, *Spectral approximation of fractional PDEs in image processing and phase field modeling*, Comput. Methods Appl. Math., 17 (2017), pp. 661–678, https://doi.org/10.1515/cmam-2017-0039.

[3] H. Antil and E. Otárola, *A FEM for an optimal control problem of fractional powers of elliptic operators*, SIAM J. Control Optim., 53 (2015), pp. 3432–3456, https://doi.org/10.1137/140975061.

[4] H. Antil, E. Otárola, and A. J. Salgado, *Optimization with respect to order in a fractional diffusion model: Analysis, approximation and algorithmic aspects*, J. Sci. Comput., 77 (2018), pp. 204–224, https://doi.org/10.1007/s10915-018-0703-0.

[5] H. Antil, J. Pfefferer, and S. Rogovs, *Fractional operators with inhomogeneous boundary conditions: Analysis, control, and discretization*, Commun. Math. Sci., 16 (2018), pp. 1395–1426, https://doi.org/10.4310/CMS.2018.v16.n5.a11.

[6] H. Antil, J. Pfefferer, and M. Warma, *A note on semilinear fractional elliptic equation: Analysis and discretization*, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 2049–2067, https://doi.org/10.1051/m2an/2017023.

[7] H. Antil and C. N. Rautenberg, *Sobolev spaces with non-Muckenhoupt weights, fractional elliptic operators, and applications*, SIAM J. Math. Anal., 51 (2019), pp. 2479–2503, https://doi.org/10.1137/18M1224970.

[8] H. Antil and M. Warma, *Optimal control of the coefficient for fractional p-Laplace equation: Approximation and convergence*, RIMS Kôkyûroku, 2090 (2018), pp. 102–116.

[9] H. Antil and M. Warma, *Optimal control of the coefficient for the regional fractional p-Laplace equation: Approximation and convergence*, Math. Control Relat. Fields, 9 (2019), pp. 1–38.

[10] H. Antil and M. Warma, *Optimal control of fractional semilinear PDEs*, ESAIM Control Optim. Calc. Var., 26 (2020), 5, https://doi.org/10.1051/cocv/2019003.

[11] Ph. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, *Coherent measures of risk*, Math. Finance, 9 (1999), pp. 203–228, https://doi.org/10.1111/1467-9965.00068.

[12] H. Attouch, G. Buttazzo, and G. Michaille, *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*, 2nd ed., MOS-SIAM Ser. Optim. 17, SIAM, Philadelphia, 2014, https://doi.org/10.1137/1.9781611973488.

[13] E. Barkai, R. Metzler, and J. Klafter, *From continuous time random walks to the fractional Fokker-Planck equation*, Phys. Rev. E (3), 61 (2000), pp. 132–138, https://doi.org/10.1103/PhysRevE.61.132.

[14] A. Ben-Tal and M. Teboulle, *Expected utility, penalty functions, and duality in stochastic nonlinear programming*, Management Sci., 32 (1986), pp. 1445–1466, https://doi.org/10.1287/mnsc.32.11.1445.

[15] A. Ben-Tal and M. Teboulle, *An old-new concept of convex risk measures: The optimized certainty equivalent*, Math. Finance, 17 (2007), pp. 449–476, https://doi.org/10.1111/j.1467-9965.2007.00311.x.

[16] A. Bonito and J. E. Pasciak, *Numerical approximation of fractional powers of elliptic operators*, Math. Comp., 84 (2015), pp. 2083–2110, https://doi.org/10.1090/S0025-5718-2015-02937-8.

[17] A. Borzì and G. von Winckel, *Multigrid methods and sparse-grid collocation techniques for parabolic optimal control problems with random coefficients*, SIAM J. Sci. Comput., 31 (2009), pp. 2172–2192, https://doi.org/10.1137/070711311.

[18] A. Borzì and G. von Winckel, *A POD framework to determine robust controls in PDE optimization*, Comput. Vis. Sci., 14 (2011), pp. 91–103, https://doi.org/10.1007/s00791-011-0165-5.

[19] X. Cabré and J. Tan, *Positive solutions of nonlinear problems involving the square root of the Laplacian*, Adv. Math., 224 (2010), pp. 2052–2093, https://doi.org/10.1016/j.aim.2010.01.025.

[20] L. Caffarelli and L. Silvestre, *An extension problem related to the fractional Laplacian*, Comm. Partial Differential Equations, 32 (2007), pp. 1245–1260, https://doi.org/10.1080/03605300600987306.

[21] L. A. Caffarelli and P. R. Stinga, *Fractional elliptic equations, Caccioppoli estimates and regularity*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 33 (2016), pp. 767–807, https://doi.org/10.1016/j.anihpc.2015.01.004.

[22] A. Capella, J. Dávila, L. Dupaigne, and Y. Sire, *Regularity of radial extremal solutions for some non-local semilinear equations*, Comm. Partial Diferential Equations, 36 (2011), pp. 1353–1384, https://doi.org/10.1080/03605302.2011.562954.

[23] L. Chen, *iFEM: An Integrated Finite Element Methods Package in MATLAB*, Technical Report, University of California at Irvine, Irvine, CA, 2009.

[24] P. Chen and A. Quarteroni, *Weighted reduced basis method for stochastic optimal control problems with elliptic PDE constraint*, SIAM/ASA J. Uncertain. Quantif., 2 (2014), pp. 364–396, https://doi.org/10.1137/130940517.

[25] P. Chen, A. Quarteroni, and G. Rozza, *Stochastic optimal Robin boundary control problems of advection-dominated elliptic equations*, SIAM J. Numer. Anal., 51 (2013), pp. 2700–2722, https://doi.org/10.1137/120884158.

[26] P. Chen, A. Quarteroni, and G. Rozza, *Multilevel and weighted reduced basis method for stochastic optimal control problems constrained by Stokes equations*, Numer. Math., 133 (2016), pp. 67–102.

[27] J. B. Conway, *A Course in Functional Analysis*, Grad. Texts in Math. 96, Springer, New York, 2019.

[28] E. Di Nezza, G. Palatucci, and E. Valdinoci, *Hitchhiker's guide to the fractional Sobolev spaces*, Bull. Sci. Math., 136 (2012), pp. 521–573, https://doi.org/10.1016/j.bulsci.2011.12.004.

[29] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, Appl. Math. Sci. 159, Springer-Verlag, New York, 2004, https://doi.org/10.1007/978-1-4757-4355-5.

[30] H. Föllmer and A. Schied, *Convex measures of risk and trading constraints*, Finance Stoch., 6 (2002), pp. 429–447, https://doi.org/10.1007/s007800200072.

[31] C. Geldhauser and E. Valdinoci, *Optimizing the fractional power in a model with stochastic PDE constraints*, Adv. Nonlinear Stud., 18 (2018), pp. 649–669.

[32] V. Gol'dshtein and A. Ukhlov, *Weighted Sobolev spaces and embedding theorems*, Trans. Amer. Math. Soc., 361 (2009), pp. 3829–3850, https://doi.org/10.1090/S0002-9947-09-04615-7.

[33] E. Hille and R. S. Phillips, *Functional Analysis and Semi-groups*, revised ed., Amer. Math. Soc. Colloq. Publ. 31, American Mathematical Society, Providence, RI, 1957.

[34] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, Pure Appl. Math. 88, Academic Press, New York, London, 1980.

[35] D. P. Kouri, *A multilevel stochastic collocation algorithm for optimization of PDEs with uncertain coefficients*, SIAM/ASA J. Uncertain. Quantif., 2 (2014), pp. 55–81, https://doi.org/10.1137/130915960.

[36] D. P. Kouri, M. Heinkenschloss, D. Ridzal, and B. G. van Bloemen Waanders, *A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty*, SIAM J. Sci. Comput., 35 (2013), pp. A1847–A1879, https://doi.org/10.1137/120892362.

[37] D. P. Kouri, M. Heinkenschloss, D. Ridzal, and B. G. van Bloemen Waanders, *Inexact objective function evaluations in a trust-region algorithm for PDE-constrained optimization under uncertainty*, SIAM J. Sci. Comput., 36 (2014), pp. A3011–A3029, https://doi.org/10.1137/140955665.

[38] D. P. Kouri and A. Shapiro, *Optimization of PDEs with uncertain inputs*, in Frontiers in PDE-Constrained Optimization, H. Antil, D. P. Kouri, M. D. Lacasse, and D. Ridzal, eds., Springer, New York, 2018, pp. 41–81.

[39] D. P. Kouri and T. M. Surowiec, *Existence and optimality conditions for risk-adverse PDE-constrained optimization*, SIAM/ASA J. Uncertain. Quantif., 6 (2018), pp. 787–815, https://doi.org/10.1137/16M1086613.

[40] A. Kufner and B. Opic, *How to define reasonably weighted Sobolev spaces*, Comment. Math. Univ. Carolin., 25 (1984), pp. 537–554.

[41] P. D. Lax, *Functional Analysis*, John Wiley & Sons, New York, Chichester, Brisbane, Toronto, 2002.

[42] D. Meidner, J. Pfefferer, K. Schürholz, and B. Vexler, *hp-finite elements for fractional diffusion*, SIAM J. Numer. Anal., 56 (2018), pp. 2345–2374, https://doi.org/10.1137/17M1135517.

[43] R. Metzler and J. Klafter, *The random walk's guide to anomalous diffusion: A fractional dynamics approach*, Phys. Rep., 339 (2000), no. 1.

[44] K. S. Miller and S. G. Samko, *Completely monotonic functions*, Integral Transform. Spec. Funct., 12 (2001), pp. 389–402.

[45] S. A. Molchanov and E. Ostrovskii, *Symmetric stable processes as traces of degenerate diffusion processes*, Teor. Verojatnost. i Primenen., 14 (1969), pp. 127–130.

[46] A. S. Nekvinda, *Characterization of traces of the weighted Sobolev space $W^{1,p}(\Omega, d_M^\epsilon)$ on M*, Czechoslovak Math. J., 43(118) (1993), pp. 695–711.

[47] R. H. Nochetto, E. Otárola, and A. J. Salgado, *A PDE approach to fractional diffusion in general domains: A priori error analysis*, Found. Comput. Math., 15 (2015), pp. 733–791, https://doi.org/10.1007/s10208-014-9208-x.

[48] R. T. Rockafellar and S. Uryasev, *The fundamental risk quadrangle in risk management, optimization and statistical estimation*, Surv. Oper. Res. Manag. Sci., 18 (2013), pp. 33–53.

[49] R. T. Rockafellar, S. Uryasev, and M. Zabarankin, *Generalized deviations in risk analysis*, Finance Stoch., 10 (2006), pp. 51–74, https://doi.org/10.1007/s00780-005-0165-8.

[50] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, 2nd ed., MOS-SIAM Ser. Optim. 16, SIAM, Philadelphia, 2014, https://doi.org/10.1137/1.9781611973433.

[51] J. Sprekels and E. Valdinoci, *A new type of identification problems: Optimizing the fractional order in a nonlocal evolution equation*, SIAM J. Control Optim., 55 (2017), pp. 70–93, https://doi.org/10.1137/16M105575X.

[52] P. R. Stinga and J. L. Torrea, *Extension problem and Harnack's inequality for some fractional operators*, Comm. Partial Differential Equations, 35 (2010), pp. 2092–2122, https://doi.org/10.1080/03605301003735680.

[53] L. Tartar, *An Introduction to Sobolev Spaces and Interpolation Spaces*, Lect. Notes Unione Mat. Ital. 3, Springer, Berlin, 2007.

[54] F. Tröltzsch, *Optimal Control of Partial Differential Equations. Theory, Methods and Applications*, Grad. Stud. Math. 112, American Mathematical Society, Providence, RI, 2010, https://doi.org/10.1090/gsm/112.

[55] B. O. Turesson, *Nonlinear Potential Theory and Weighted Sobolev Spaces*, Springer, Berlin, 2000, https://doi.org/10.1007/BFb0103908.

[56] M. Warma, *A fractional Dirichlet-to-Neumann operator on bounded Lipschitz domains*, Commun. Pure Appl. Anal., 14 (2015), pp. 2043–2067, https://doi.org/10.3934/cpaa.2015.14.2043.

[57] M. Warma, *The fractional relative capacity and the fractional Laplacian with Neumann and Robin boundary conditions on open sets*, Potential Anal., 42 (2015), pp. 499–547, https://doi.org/10.1007/s11118-014-9443-4.

[58] C. J. Weiss, B. G. van Waanders, and H. Antil, *Fractional operators applied to geophysical electromagnetics*, Geophys. J. Int., 220 (2020), pp. 1242–1259.

[59] K. Yosida, *Functional Analysis*, Classics Math., Springer-Verlag, Berlin, 1995, https://doi.org/10.1007/978-3-642-61859-8.

[60] Z. Zou, D. P. Kouri, and W. Aquino, *A locally adapted reduced basis method for solving risk-averse PDE-constrained optimization problems*, in Proceedings of the AIAA Non-Deterministic Approaches Conference, 2018.