FULL LENGTH PAPER

Series B

Risk-neutral PDE-constrained generalized Nash equilibrium problems

Deborah B. Gahururu[1] · Michael Hintermüller[2,3] · Thomas M. Surowiec[1]

Abstract

A class of risk-neutral generalized Nash equilibrium problems is introduced in which the feasible strategy set of each player is subject to a common linear elliptic partial differential equation with random inputs. In addition, each player's actions are taken from a bounded, closed, and convex set on the individual strategies and a bound constraint on the common state variable. Existence of Nash equilibria and first-order optimality conditions are derived by exploiting higher integrability and regularity of the random field state variables and a specially tailored constraint qualification for GNEPs with the assumed structure. A relaxation scheme based on the Moreau-Yosida approximation of the bound constraint is proposed, which ultimately leads to numerical algorithms for the individual player problems as well as the GNEP as a whole. The relaxation scheme is related to probability constraints and the viability of the proposed numerical algorithms are demonstrated via several examples.

✉ Thomas M. Surowiec
  surowiec@mathematik.uni-marburg.de

  Deborah B. Gahururu
  gahururu@mathematik.uni-marburg.de

  Michael Hintermüller
  hint@math.hu-berlin.de

[1] Philipps-Universität Marburg, FB12 Mathematik und Informatik, Hans-Meerwein Straße 6, 35032 Marburg, Germany

[2] Weierstrass Institute, Mohrenstr. 39, 10117 Berlin, Germany

[3] Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

1 Introduction

Whether it be a consequence of noisy measurements, estimated parameter values, or model ambiguity, uncertainty is present in just about every mathematical model of real-world phenomena. Whenever the uncertainty is untreatable by deterministic quantities, it is best to assimilate it into our mathematical models via random variables, vectors, or elements. This allows us to find more robust solutions in the face of future uncertainty and guard against outlier events. Since many models in engineering and the natural sciences are defined by partial differential equations (PDEs), the inclusion of random inputs leads us to consider parametric or random PDEs as part of our optimization problems, cf. [9, 18, 49, 60, 62].

PDE-constrained optimization under uncertainty is a challenging area of mathematical optimization with many relevant applications in the engineering sciences. It is a growing field with many recent of contributions in theory and algorithms, see e.g. [3, 11–14, 20, 22, 39–41, 43–45, 57, 58, 63]. However, many practical problems require the simultaneous minimization of multiple objectives. By pitting these objectives against each other, i.e., treating the problem as a noncooperative game with each objective and separate control representing a player and its individual strategy, we naturally come to study PDE-constrained Nash equilibrium problems under uncertainty. In the deterministic setting, we mention here the pioneering works [16, 26, 51–54, 61]. It is important to note, however, that the models in these papers do not consider bound constraints, in particular there are no state constraints. This is an important distinction, as it makes the difference between modeling the game via a coupled PDEs (no bound constraints) versus a variational inequality (no state constraints) versus a quasivariational inequality (with state constraints).

As with their deterministic counterparts, it is often necessary to look for a control that forces the state (solution of the PDE) to satisfy certain bound constraints, e.g., below a maximum temperature threshold or above a physical obstacle. When uncertain inputs are involved, this problem of state constraints becomes much more challenging. This is due in part to a lack of smoothness with respect to the random parameters and missing compactness properties, which we would expect in a deterministic setting. Moreover, though an adjoint equation solely for the state equation can be introduced, an adjoint equation that combines the state equation as well as a multiplier for the state constraint cannot be derived without assuming additional properties. The inclusion of state constraints leads in fact to generalized Nash equilibrium problems in Banach spaces. For recent work in the deterministic setting, we refer the reader to [33, 34, 37, 38] and the references therein.

Summarizing the discussions above, we thus consider a class of risk-neutral PDE-constrained GNEPs under uncertainty subject to state constraints. In an abstract sense, this amounts to considering an N-player GNEP in which the ith player's problem takes the form

$$\min_{z_i \in Z_{\text{ad}}^i} \{\mathbb{E}_{\mathbb{P}}[J_i(z_i, z_{-i}, \cdot)] \,|\, S(z_i, z_{-i}, \omega) \in K \; \mathbb{P}\text{-a.s.}\}.$$

Here, $S(z)$ is the z-dependent random field solution of a linear elliptic PDE with uncertain inputs, Z_{ad}^i and K are closed convex sets and J_i is an appropriate convex disutility function for player i. We will make the appropriate data assumptions below. The term "risk-neutral" arises due to the fact that only the expected disutility is considered. Letting $\overline{z}$ be a Nash equilibrium for this problem, player i would expect $\overline{z}_i$ to be the best response to $\overline{z}_{-i}$ *on average*, i.e., if the game were played repeatedly. Since the literature is rather scarce on the treatment of state constraints in PDE-constrained optimization under uncertainty, see e.g., [19, 23] and the recent preprint [21], we pay special attention to the case where $N = 1$, as well. We comment further on the studies [19, 23] below, which make use of probability constraints. In contrast, the abstract results in [22] can be used for state constraints as considered in this paper. However, these results require a different kind of constraint qualification that may be difficult to verify in general.

The contributions of our paper are as follows:

1. We exploit existing results on elliptic regularity theory to prove higher integrability and regularity of the random field solutions $S(z)$.
2. Under appropriate constraint qualifications, we prove existence of solutions/equilibria and derive optimality conditions for the optimization problem and GNEP.
3. We extend the well-known Moreau-Yosida approach for state constraints to the stochastic case and rigorously prove that the approximations converge to the original GNEP.
4. The link between the Moreau-Yosida regularization technique and probability constraints is established using concentration inequalities.
5. We propose and demonstrate the viability of numerical algorithms for the optimization problem and GNEP.

The first contribution is crucial, as we need at least essential boundedness of the random field solutions in order to use techniques of convex optimization in Banach spaces to develop the optimality theory. In (2), we require a Slater-type condition for the optimization problem and the strict uniform feasible response (SUFR) condition introduced in [33] for the GNEP. The SUFR condition imposes a kind of hidden symmetry on the GNEP model. Although Moreau-Yosida regularization has been used successfully in deterministic settings, the stochastic setting poses additional pitfalls. Nevertheless, passing to the limit in the relaxation parameter is crucial for the justification of the numerical methods in the fully continuous setting. The link to probability constraints in (4) is interesting in its own right, since the approximating problems are much easier to solve than a similar problem with probability conditions. In addition,

we obtain a kind of probabilistic rate of convergence for the Moreau-Yosida relaxations, which is reflected in the properties of the out-of-sample controlled states in (5); even after solving with relatively small increasing batches and modest values of the relaxation parameter. The encouraging results in our numerical study (5) motivate a number of future research directions.

The rest of the paper is structured as follows. In Sect. 2, we pose a number of basic assumptions along with an analysis of the forward problem. In addition, the optimization problems and GNEP are introduced. Following this, we derive existence and optimality conditions in Sect. 3; using the underlying structure and basic constraint qualifications. Due to the low multiplier regularity in the optimality conditions and a lack of adjoint equation in the sense that the righthand side is the sum of the derivative of the objective with respect to the state and the Lagrange multiplier for the state constraint, we propose a Moreau-Yosida technique in Sect. 4. This allows us to formulate function-space-based numerical algorithms for both the optimization problems and the GNEP in Sect. 5. The potential of the algorithms is demonstrated via several numerical examples. In particular, we provide a brief, post-optimal analysis using the performance of the computed controls to derive a statistic on the violation of the state constraint.

2 Problem formulation

2.1 Notation, standing assumptions, and preliminary results

We start by defining the necessary function spaces. We assume that the physical domain $D \subset \mathbb{R}^d$ with $d = 1, 2$, or 3 is an open bounded set such that D is either a convex polyhedron or the boundary of D, denoted by ∂D, is of class $C^{1,1}$.

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ denotes a complete probability space, where Ω is the sample space of possible outcomes, $\mathcal{F}$ the Borel σ-alegra of Ω for a fixed topology on Ω and $\mathbb{P}$ is a probability measure.

Given a real-valued Banach space $(V, \|\cdot\|_V)$, Borel measure μ, and $p \in [1, \infty]$ we denote the usual Lebesgue-Bochner space $L_\mu^p(\Omega; V)$ of all strongly $\mathcal{F}$-measurable V-valued functions by

$$L_\mu^p(\Omega; V) = \{u : \Omega \to V : u \text{ strongly } \mathcal{F} - \text{measurable and } \|u\|_{L_\mu^p(\Omega;V)} < \infty\}$$

where

$$\|u\|_{L_\mu^p(\Omega;V)} = \begin{cases} \left(\int_\Omega \|u(\omega)\|_V^p \, d\mu(\omega)\right)^{1/p}, & p < \infty \\ \mu - \operatorname*{ess\,sup}_{\omega \in \Omega} \|u(\omega)\|_V, & p = \infty. \end{cases}$$

When $V = \mathbb{R}$, we set $L_\mu^p(\Omega; \mathbb{R}) = L_\mu^p(\Omega)$ the usual Lebesgue space with underlying measure μ. When the Lebesgue measure $\mu = \mathcal{L}$ is considered, we omit the subscript $\mathcal{L}$ and simply write $L^p(\Omega)$. We denote by $\mathcal{F}_{\mathcal{L}}$ the σ-algebra of Lebesgue measurable sets. We recall here that for $1 \leq p, q \leq \infty$ such that $1/p + 1/q = 1$, it is known that

the topological dual fulfills $L_\mu^p(\Omega; V)^* \simeq L_\mu^q(\Omega; V^*)$. If V is reflexive, then so is $L_\mu^p(\Omega; V)$ for $1 < p < \infty$. For further information see [28, Chapter III].

We show in the sections below that the multipliers for the stochastic state constraints are of very low regularity, i.e., bounded additive measures. We will need the space **ba**, which we recall here for ease of reference, cf. [27, 20.27 Definition] or [17].

Definition 1 Let $(\varXi, \mathcal{B}, \mu)$ be a σ-finite measure space. The space $\mathbf{ba}(\varXi, \mathcal{B}, \mu)$ denotes the set of all real-valued set-functions $\tau : \mathcal{B} \to \mathbb{R}$ such that

(i) $\sup\{|\tau(A)| : A \in \mathcal{B}\} < \infty$,
(ii) $\tau(A \cup B) = \tau(A) + \tau(B)$ for $A, B \in \mathcal{B}$ with $A \cap B = \emptyset$ and
(iii) $\tau(A) = 0$ if $A \in \mathcal{B}$ is μ-null, i.e. $\tau << \mu$.

The norm of $\tau \in \mathbf{ba}(\varXi, \mathcal{B}, \mu)$ is given by $|\tau|(\varXi)$, the total variation of τ on $\mathcal{B}$.

The key result for our analysis related to this space is the existence of an isometric isomorphism between $(L_\pi^\infty(\varXi))^*$ and $\mathbf{ba}(\varXi, \mathcal{B}, \pi)$, cf. [17, Thm. IV.8.16], where we use

$$\varXi = \Omega \times D, \quad \mathcal{B} = \mathcal{F} \otimes \mathcal{F}_{\mathcal{L}}, \quad \pi = \mathbb{P} \times \mathcal{L}.$$

Finally, we fix several notational conventions. For a (real) Banach space V we denote the expectation of a random element $X : \Omega \to V$ by

$$\mathbb{E}_\mathbb{P}[X] = \int_\Omega X(\omega)\, d\mathbb{P}(\omega) \in V.$$

For some nonempty subset $C \subset V$, $\mathcal{I}_C : V \to \mathbb{R} \cup \{\infty\}$ represents the standard indicator function, which satisfies $\mathcal{I}_C(x) = 0$ if $x \in C$ and $+\infty$ otherwise. For an arbitrary convex set K, we define the standard convex normal cone by

$$\mathcal{N}_K(x) = \begin{cases} \{x^* \in V^* \mid \langle x^*, y - x \rangle \leq 0, \quad \forall y \in K\}, & \text{if } x \in K, \\ \emptyset, & \text{otherwise.} \end{cases}$$

The (set-theoretic) characteristic function associated with some subset A is denoted by χ or χ_A, where $\chi_A(x) = 1$ if $x \in A$ and 0 otherwise. Strong convergence of a sequence is denoted by $\to$, weak-convergence by $\rightharpoonup$, and weak-*-convergence by $\overset{*}{\rightharpoonup}$. The *closed* ε-ball with center x in some normed space is denoted $\mathbb{B}_\varepsilon(x)$. The superscript $*$ is used to denote the adjoint operator or dual space. As usual $C \lesssim D$ means that C is bounded by D up to an independent constant. For two Banach spaces V and W, the set of all bounded linear operators from V to W will be denoted by $\mathcal{L}(V, W)$. We use the typical convention from game theory for a vector u with N components for emphasizing the ith component by writing $u = (u_i, u_{-i}) = (u_{-i}, u_i)$.

2.2 Risk-neutral PDE-constrained equilibrium problems

2.2.1 PDE-constrained equilibrium problems as strategic games

As mentioned above, our results apply to both PDE-constrained optimization problems under uncertainty as well as stochastic equilibrium problems with PDE-constraints. Whereas the solution concept for PDE-constrained optimization is obvious, there are several possibilities for equilibrium problems from the perspective of game theory. The notation in this brief section is chosen to reflect the references to the game theory literature.

We recall that a strategic game comprises a set of N players or agents, their sets of actions A^i, and a unique preference relation for each player over all possible profiles of actions $\boldsymbol{a} \in A := \times_{i=1}^{N} A^i$. In many cases, the preference relation can be described by the values of utility functions $u_i : A \to \mathbb{R}$ and the preferred solution concept for noncooperative behavior is often taken to be a Nash equilibrium; cf. [47]. The latter states that $\bar{\boldsymbol{a}} \in A$ is a (pure strategy) Nash equilibrium provided for all $i = 1, \ldots, N$ we have

$$u_i(\bar{\boldsymbol{a}}) \geq u_i(\bar{\boldsymbol{a}}_{-i}, a_i) \quad \forall a_i \in A^i, \tag{2.1}$$

see, e.g., [48] for more details. We will refer to games in which the solution concept is a Nash equilibrium as Nash Equilibrium Problems or NEPs.

We will take an analogous perspective for our PDE-constrained equilibrium problems. However, due to the presence of state constraints, the sets of actions are set-valued mappings $A^i(\boldsymbol{a}_{-i})$ that also depend on $\boldsymbol{a}_{-i}$ for each i. This leads to a natural extension, first introduced by Debreu [15], see also [6]: $\bar{\boldsymbol{a}} \in A$ is a (generalized) Nash equilibrium provided for all $i = 1, \ldots, N$ we have $\bar{a}_i \in A^i(\bar{\boldsymbol{a}}_{-i})$ and

$$u_i(\bar{\boldsymbol{a}}) \geq u_i(\bar{\boldsymbol{a}}_{-i}, a_i) \quad \forall a_i \in A^i(\bar{\boldsymbol{a}}_{-i}). \tag{2.2}$$

These games are significantly more difficult from both a theoretical as well as numerical perspective due to the embedded fixed point relation. We refer to games of this type as Generalized Nash Equilibrium Problems or GNEPs.

2.2.2 Linear elliptic random PDEs

Returning now to the context of PDE-constrained optimization, we introduce a class of linear elliptic random PDEs as our state system. Let

$$\mathcal{U} := L_{\mathbb{P}}^2(\Omega; H_0^1(D)).$$

Given $z \in L^2(D)$, we consider the following problem: Find $u \in \mathcal{U}$ such that

$$\mathbb{E}_{\mathbb{P}}\left[\int_D A(x, \cdot)\nabla u(x, \cdot) \cdot \nabla v(x, \cdot)\, \mathrm{d}x\right] = \mathbb{E}_{\mathbb{P}}\left[\int_D ((B(\cdot)z)(x) + f(x, \cdot))v(x, \cdot)\, \mathrm{d}x\right], \tag{2.3}$$

for all test functions $v \in \mathcal{U}$. Note that (2.3) can be equivalently written in a semi-weak form. Let u solve (2.3). Then using $v(x, \omega) = \chi_A(\omega)\varphi(x)$ such that $A \in \mathcal{F}$ and $\varphi \in H_0^1(D)$ (or $\varphi \in C_0^\infty(D)$) we have

$$
\int_A \int_D A(x, \omega)\nabla u(x, \omega) \cdot \nabla \varphi(x) \, dx d\mathbb{P}(\omega)
$$
$$
= \int_A \int_D ((B(\omega)z)(x) + f(x, \omega))\varphi(x) \, dx d\mathbb{P}(\omega), \tag{2.4}
$$

for every $\varphi \in H_0^1(D)$ and consequently

$$
\int_D A(x, \omega)\nabla u(x, \omega) \cdot \nabla \varphi(x) \, dx = \int_D ((B(\omega)z)(x) + f(x, \omega))\varphi(x) \, dx, \quad \mathbb{P}-a.s. \tag{2.5}
$$

for every $\varphi \in H_0^1(D)$. The reverse direction (from $\mathbb{P}$-pointwise weak solutions to a solution of (2.3)) can be easily adapted from the nonlinear setting in [42]. The key components of the argument are: Prove the existence of a solution for $\mathbb{P}$-a.e. ω, demonstrate measurability in ω using Fillipov's theorem for measurable selections, and obtain integrability using standard a priori estimates for elliptic PDEs. It is sometimes more convenient to work with one form versus the other as we will see below. For $z = 0$, we denote the solution of (2.3) by u_f and for $f \equiv 0$ we set $u = S(z)$. Hence, any solution u of (2.3) can be written

$$
u = S(z) + u_f. \tag{2.6}
$$

We will demonstrate below that $S(z)$ is a bounded linear operator in z between appropriate function spaces.

In order to ensure well-defined solutions and derive higher regularity results, we make the following additional assumptions on the problem data.

Assumption 1 In addition to the standing assumptions on D, ∂D, and $(\Omega, \mathcal{F}, \mathbb{P})$, the following sets of assumptions will be necessary below.

(i) (*Minimum Regularity*) The coefficient mapping $A : D \times \Omega \to \mathbb{R}$ is $(\mathcal{L} \times \mathbb{P})$-measurable and there exist constants $0 < \underline{A} < \overline{A}$ such that

$$
\underline{A} \leq A(x, \omega) \leq \overline{A} \quad (\mathcal{L} \times \mathbb{P})\text{-a.e. } (x, \omega) \in D \times \Omega
$$

The fixed bulk term f satisfies

$$
f \in L_{\mathbb{P}}^\infty(\Omega; L^2(D))
$$

(ii) (*Higher Regularity*) In addition to (i), $A \in L_{\mathbb{P}}^\infty(\Omega; C^{0,1}(\bar{D}))$.

(iii) (*Control Mapping*) The control mapping $B : \Omega \to \mathcal{L}(L^2(D)^N, L^2(D))$ is measurable and essentially bounded, i.e. $B \in L_{\mathbb{P}}^\infty(\Omega, \mathcal{L}(L^2(D)^N, L^2(D)))$. Moreover,

as a mapping from Ω to $\mathcal{L}(L^2(D), H^{-1}(D))$, B is completely continuous in the sense that for $\mathbb{P}$-a.e. $\omega \in \Omega$ we have

$$z_k \rightharpoonup z \text{ in } L^2(D)^N \implies B(\omega)z \to B(\omega)z \text{ in } H^{-1}(D).$$

Some remarks are in order. Assumption 1.(i) can be slightly weakened to allow for unbounded coefficients and still obtain the existence of solutions, cf. e.g., [24]. It is also possible to choose f and/or $B(\omega)z$ that is unbounded in ω. However, weakening these assumptions would mean that the solutions u to (2.3) are also not bounded. The latter property is essential for our treatment of state constraints. The Lipschitz continuity of $A(\omega, \cdot) : \bar{D} \to \mathbb{R}$ in Assumption 1.(ii) will be used to ensure boundedness of u in x. This along with the regularity assumption on the boundary ∂D can be slightly weakened to the extent that we can guarantee $u \in L^\infty_{\mathbb{P} \times \mathcal{L}}(\Omega \times D)$, e.g., we could relax Lipschitz to Hölder and work with $u(\cdot, \omega)$ in $W^{1,p}(D)$ with $p > d$. The properties in Assumption 1.(iii) are the weakest possible for our analysis. Using Assumption 1, we gather several essential properties of the mapping $z \mapsto u$ in the following result.

Proposition 1 *Let Assumption 1 hold. For any $z \in L^2(D)$, there exists a unique solution $u \in \mathcal{U}$ of (2.3). Moreover, $u \in L^\infty_{\mathbb{P}}(\Omega; H^2(D) \cap H^1_0(D))$ and the following a priori bound holds*

$$\|u\|_{L^\infty_{\mathbb{P}}(\Omega;H^2(D)\cap H^1_0(D))} \leq C \left(\|f\|_{L^\infty_{\mathbb{P}}(\Omega;L^2(D))} + \|B(\cdot)z\|_{L^\infty_{\mathbb{P}}(\Omega;L^2(D))} \right) \quad (2.7)$$

Here, C is independent of ω.

Proof Defining the bilinear form $b : \mathcal{U} \times \mathcal{U} \to \mathbb{R}$ by

$$b(u, v) := \mathbb{E}_{\mathbb{P}} \left[\int_D A(x, \cdot) \nabla u(x, \cdot) \cdot \nabla v(x, \cdot) \, \mathrm{d}x \right]$$

and z-dependent linear form $L(\cdot; z) : \mathcal{U} \to \mathbb{R}$ by

$$L(v; z) := \mathbb{E}_{\mathbb{P}} \left[\int_D ((B(\cdot)z)(x) + f(x, \cdot))v(x, \cdot) \, \mathrm{d}x \right],$$

we can view (2.3), as the variational problem: Find $u \in \mathcal{U}$ such that

$$b(u, v) = L(v; z) \quad \forall v \in \mathcal{U}.$$

It readily follows from Assumption 1 that b is a $\mathcal{U}$-coercive bilinear form. Then by the Lax-Milgram Lemma there exists a unique solution $u \in \mathcal{U}$. In light of the equivalence to (2.5), we immediately deduce from the standard a priori bound:

$$\|u(\cdot, \omega)\|_{H^1_0(D)} \leq C_1 \|B(\omega)z + f(\cdot, \omega)\|_{H^{-1}(D)} \quad (2.8)$$

that $u : \Omega \to H_0^1(D)$ is $\mathbb{P}$-essentially bounded. Due to the assumptions on A, C_1 does not depend on ω.

For the a priori bound (2.7), we need to consider two cases. We once again appeal to the equivalence between (2.3) and (2.5). If ∂D is of type $C^{1,1}$, then it follows from Assumption 1 along with Friedrichs' theorem, see e.g., [5, A12.2 Theorem], that for $\mathbb{P}$-a.e. $\omega \in \Omega$ we have

$$\|u(\cdot, \omega)\|_{H^2(D)} \le C(\omega) \left(\|u(\cdot, \omega)\|_{H_0^1(D)} + \|B(\omega)z\|_{L^2(D)} + \|f(\cdot, \omega)\|_{L^2(D)} \right). \tag{2.9}$$

Here, $C(\omega) = C\left(\partial D, d, \underline{A}, \|A(\cdot, \omega)\|_{C^{0,1}(\bar{D})} \right)$. The same estimate also holds when ∂D is nonsmooth, but D is a convex polyhedron, see Remark 1 below. The "constant" $C(\omega)$ is indeed a bounded and measurable function in ω. This follows from the fact that the term $\|A(\cdot, \omega)\|_{C^{0,1}(\bar{D})}$ is measurable, uniformly bounded away from zero, and $C(\omega)$ is a sum of rational functions of $\|A(\cdot, \omega)\|_{C^{0,1}(\bar{D})}$, where it appears in a numerator and a denominator. Continuing, for $\mathbb{P}$-a.e. $\omega \in \Omega$, we have

$$\begin{aligned} \|u(\omega, \cdot)\|_{H^2(D) \cap H_0^1(D)} &= \max \left\{ \|u(\omega, \cdot)\|_{H^2(D)}, \|u(\omega, \cdot)\|_{H_0^1(D)} \right\} \\ &\le \|u(\omega, \cdot)\|_{H^2(D)} + \|u(\omega, \cdot)\|_{H_0^1(D)} \\ &\le C(\omega) \Big(\|u(\omega, \cdot)\|_{H_0^1(D)} + \|B(\omega)z\|_{L^2(D)} \\ &\quad + \|f(\omega, \cdot)\|_{L^2(D)} \Big) \\ &\quad + C_1 \Big(\|B(\omega)z\|_{H^{-1}(D)} + \|f(\omega, \cdot)\|_{H^{-1}(D)} \Big). \end{aligned}$$

Furthermore, we obtain

$$\begin{aligned} \|u(\omega, \cdot)\|_{H^2(D) \cap H_0^1(D)} &\le C(\omega)C_1 \left(\|B(\omega)z\|_{H^{-1}(D)} + \|f(\omega, \cdot)\|_{H^{-1}(D)} \right) \\ &\quad + C(\omega) \left(\|B(\omega)z\|_{L^2(D)} + \|f(\omega, \cdot)\|_{L^2(D)} \right) \\ &\quad + C_1 \left(\|B(\omega)z\|_{H^{-1}(D)} + \|f(\omega, \cdot)\|_{H^{-1}(D)} \right). \end{aligned}$$

Finally, due to the Gelfand triple $H_0^1(D) \hookrightarrow L^2(D) \hookrightarrow H^{-1}(D)$, we have

$$\|u(\omega, \cdot)\|_{H^2(D) \cap H_0^1(D)} \le \widehat{C}(\omega) \left(\|B(\omega)z\|_{L^2(D)} + \|f(\omega, \cdot)\|_{L^2(D)} \right),$$

where

$$\widehat{C}(\omega) := 3 \max \{ C(\omega), C_1, C_{\mathrm{emb}} \}^3$$

and C_{emb} is the embedding constant for $L^2(D)$ into $H^{-1}(D)$. Passing to the $\mathbb{P}$-essential supremum yields

$$\operatorname*{ess\,sup}_{\omega \in \Omega} \|u(\omega, \cdot)\|_{H^2(D) \cap H^1_0(D)} \lesssim \operatorname*{ess\,sup}_{\omega \in \Omega} \|B(\omega)z\|_{L^2(D)}$$
$$+ \operatorname*{ess\,sup}_{\omega \in \Omega} \|f(\omega, \cdot)\|_{L^2(D)} < \infty.$$

Thus, $u \in L^\infty_{\mathbb{P}}(\Omega; H^1_0(D) \cap H^2(D))$ and (2.7) follows.   □

Remark 1 For details on the well-known regularity results for deterministic elliptic PDEs on nonsmooth domains, we refer to [25, Thm. 3.2.1.2] and especially to [25, Thm. 3.1.3.3, Lem. 3.1.3.2, Thm. 3.1.3.1] for the estimation bounds.

Proposition 1 justifies the decomposition in (2.6). In particular, we see that S is a bounded linear operator and $u_f \in L^\infty_{\mathbb{P}}(\Omega; H^1_0(D) \cap H^2(D))$. We deduce several additional properties in the following corollary.

Corollary 1 *Under the hypotheses of Proposition 1 we have:*

(i) *As a mapping from $L^2(D)$ to $L^q_{\mathbb{P}}(\Omega; H^1_0(D))$ with $q \in [1, \infty)$, S is completely continuous, bounded, and linear.*
(ii) *As a mapping from $L^2(D)$ to $L^\infty_{\mathbb{P}}(\Omega; H^1_0(D) \cap H^2(D)))$, S is bounded and linear.*

Proof Case (i) is a special case of [42, Prop 2.3]. In case (ii) linearity follows trivially from the definition of $S(z)$ whereas boundedness is a consequence of (2.7) and Assumption 1.(iii).   □

We end this section by introducing a convenient $\mathbb{P}$-pointwise notation that will aid in the derivation of optimality conditions below. We define

$$\mathbf{A} : \Omega \to \mathcal{L}(H^1_0(D) \cap H^2(D), L^2(D)) \text{ and } \mathbf{B} : \Omega \to \mathcal{L}(L^2(D))$$

to be the operators given by

$$\langle \mathbf{A}(\omega)u, v \rangle = \int_D A(x, \omega) \nabla u(x) \cdot \nabla v(x) \, \mathrm{d}x$$

for $u, v \in H^1_0(D) \cap H^2(D)$ and

$$\langle \mathbf{B}(\omega)z, v \rangle = \int_D ((B(\omega)z)(x)v(x) \, \mathrm{d}x,$$

respectively. Note that $A(\omega)$ is a linear isomorphism due to the regularity results above. Given A, B we can understand $S(z) + u_f$ $\mathbb{P}$-pointwise as

$$S(z)(\omega) + u_f(\omega) = A^{-1}(\omega)B(\omega)z + A^{-1}(\omega)f(\omega) \tag{2.10}$$

whenever we need to work with higher regularity.

2.2.3 A class of risk-neutral PDE-constrained optimization problems

In this section, we introduce a class of optimization problems that will serve as a template for the individual player problems in the PDE-constrained GNEP.

Assumption 2 We assume that

(i) (*Control Constraints*) $Z_{\mathrm{ad}} \subset L^2(D)$ is a nonempty, closed, bounded, and convex set.
(ii) (*Objective*) The cost parameter $\nu \geq 0$, $u_d \in L^2(D)$, $T \in \mathcal{L}(L^2(D))$, and $J : L^2(D) \times L^2(D) \to \mathbb{R}$ is defined by

$$J(u, z) := \frac{1}{2}\|Tu - u_d\|^2_{L^2(D)} + \frac{\nu}{2}\|z\|^2_{L^2(D)}. \tag{2.11}$$

(iii) (*State Constraint*) Given $\psi \in C(\overline{\Omega \times D})$ for which there exists $\varepsilon > 0$ such that

$$\psi|_{\partial D}(\omega) \leq -\varepsilon \quad \mathbb{P}\text{-a.s.,}$$

we define the state constraint by

$$S(z) + u_f \geq \psi \text{ for } (\mathcal{L} \times \mathbb{P})\text{-a.e. } (x, \omega) \in D \times \Omega. \tag{2.12}$$

(iv) (*Feasibility*) There exists $z \in Z_{\mathrm{ad}}$ such that (2.12) holds.

The boundedness in Assumption 2.(i) is only needed in the optimization setting if $\nu = 0$. However, it is unclear how to extend the existence proof for the GNEP, as the latter follows from an application of the Kakutani-Fan-Glicksberg theorem, which includes a compactness condition. It is not necessary for our analysis to restrict ourselves to the tracking-type objective in Assumption 2.(ii). We could proceed in a more general manner as suggested in [44] under appropriate convexity, continuity, and growth conditions. This would require further technical assumptions that we believe would detract from the main purpose of the text. The nonemptiness of the feasible set in our setting is assumed in Assumption 2.(iv). Provided Z_{ad} admits a $z > 0$ with sufficiently large $L^\infty(D)$-norm, then the existence of a feasible point can be guaranteed by the maximum principle in light of the regularity result in Proposition 1.

The inclusion of state constraints in PDE-constrained optimization in the form of (2.12) is new. An alternative way of interpreting (2.12) would be to consider either

$$\mathbb{P}(S(z)(x, \cdot) + u_f(x, \cdot) - \psi(x, \cdot) \geq 0) = 1 \quad \text{for } \mathcal{L}\text{-a.e. } x \in D$$

or

$$\mathbb{P}(S(z)(x, \cdot) + u_f(x, \cdot) - \psi(x, \cdot) \geq 0 \text{ for } \mathcal{L}\text{-a.e. } x \in D) = 1.$$

From the perspective of stochastic programming, this is rather restrictive and in general settings (beyond PDE-constrained optimization), may lead to empty feasible sets.

Typically one remedies this by selecting a minimum probability level $p \in (0, 1)$ and considering instead:

$$\mathbb{P}(S(z)(x, \cdot) + u_f(x, \cdot) - \psi(x, \cdot) \geq 0 \text{ for } \mathcal{L}\text{-a.e. } x \in D) \geq p. \qquad (2.13)$$

Several recent studies have considered this perspective, see [19, 23]. However, these approaches do not circumvent the fundamental difficulties encountered with state constraints in regards to multiplier regularity and mesh-independent numerical approaches. In addition, the functional

$$\phi(z) := \mathbb{P}(S(z)(x, \cdot) + u_f(x, \cdot) - \psi(x, \cdot) \geq 0 \text{ for } \mathcal{L}\text{-a.e. } x \in D)$$

is nontrivial to analyze and use in numerical algorithms. This usually requires $\mathbb{P}$ to admit a log-concave density and for $S(z)(x, \omega)$ to have a very specific structure with respect to ω. For more on probability constraints, we refer the reader to [50, 59] and the related references therein.

We may now formulate the optimization problem

$$\min_{z \in Z_{\text{ad}}} \left\{ \mathbb{E}_{\mathbb{P}} \left[J(S(z) + u_f, z) \right] \,\middle|\, S(z) + u_f \geq \psi \text{ for } (\mathcal{L} \times \mathbb{P})\text{-a.e. } (x, \omega) \in D \times \Omega \right\}. \qquad (2.14)$$

2.2.4 A class of risk-neutral PDE-constrained GNEPs

We now introduce a noncooperative game with N players by using the results of the previous section. The individual ith player is assumed to solve the following optimization problem

$$
\begin{aligned}
\min \quad & \mathbb{E}_{\mathbb{P}} \left[\frac{1}{2} \| T_i u - u_d^i \|_{L^2(D)}^2 \right] + \frac{v_i}{2} \| z_i \|_{L^2(D)}^2 \quad && \text{over } (z_i, u) \in Z_{\text{ad}}^i \times \mathcal{U} \\
\text{s.t.} \quad & A(\omega) u = B(\omega)(z_i, z_{-i}) + f(\omega) && \mathbb{P}\text{-a.s.} \\
& u \geq \psi && \text{for } (\mathcal{L} \times \mathbb{P})\text{-a.e. } (x, \omega) \in D \times \Omega.
\end{aligned}
$$

Here, the quantities Z_{ad}^i, T_i, v_i, and u_d^i are defined analogously to those in the standard optimization setting, where we again require Assumptions 1 and 2 for each $i = 1, \ldots, N$. In what follows, we denote the collective admissible set of controls by $\mathcal{Z}_{\text{ad}} = Z_{\text{ad}}^1 \times \cdots \times Z_{\text{ad}}^N$. The main difference for the individual player problems lies in the definition of the control mapping B. For the sake of reference, we make the following assumption.

Assumption 3 The operator B has the additive representation

$$B(\omega)(z_i, z_{-i}) = B_1(\omega) z_1 + \cdots + B_N(\omega) z_N, \quad \mathbb{P}\text{-a.s.}$$

where B_i satisfies Assumption 1 for $i = 1, \ldots, N$.

In light of the assumptions, we may also formulate the PDE-constrained GNEP in terms of the following reduced space problems.

$$\min_{z_i \in Z_{\text{ad}}^i} \left\{ \mathbb{E}_{\mathbb{P}} \left[J_i \left(S(z_i, z_{-i}) + u_f, (z_i, z_{-i}) \right) \right] \Big| S(z_i, z_{-i}) + u_f \geq \psi \text{ for } (\mathcal{L} \times \mathbb{P})\text{-a.e.} \right\}.$$

(2.15)

3 Existence and optimality conditions

We first prove existence of optimal solutions of (2.14) and provide optimality conditions. Then, by extending the arguments used in [34], we prove the existence of generalized Nash equilibria for (2.15). Optimality conditions for a certain type of equilibria are also derived. We will use the concept of variational equilibria, which is strongly related to the notion of normalized equilibrium due to Rosen [56]; although Rosen's concept of normalized equilibrium was formulated using Lagrange multipliers. This is a specific class of Nash equilibria that can in many cases be computed numerically.

3.1 Risk-neutral PDE-constrained optimization problems

For the risk-neutral PDE-constrained optimization problems the existence and optimality conditions are formulated as follows.

Theorem 4 *Let Assumptions 1 and 2 hold. Then (2.14) admits a solution $\bar{z}$. If $\nu > 0$, then $\bar{z}$ is unique. Moreover, if there exists a $z_0 \in Z_{\text{ad}}$ and a constant $\kappa > 0$ such that*

$$S(z_0) + u_f - \psi > \kappa$$

(3.1)

then there exists a measure $\bar{\mu} \in \mathbf{ba}(\Xi, \mathcal{B}, \pi)$ such that

(i) (**Nonpositivity**) $\bar{\mu}$ satisfies

$$\int_{\Xi} g(x, \omega) \, d\bar{\mu}(x, \omega) \leq 0, \quad \forall g \in L_{\pi}^{\infty}(\Xi)_{+}.$$

(ii) (**Complementarity**) $\bar{\mu}$ fulfills

$$\int_{\Xi} G(\bar{z})(x, \omega) \, d\bar{\mu}(x, \omega) = 0,$$

where

$$G(z) = \iota A^{-1} B z + \iota u_f - \psi$$

and

$$\iota : L_{\mathbb{P}}^{\infty}(\Omega; H_0^1(D) \cap H^2(D))) \to L_{\pi}^{\infty}(\Xi)$$

is the continuous embedding.

*(iii) (**Subgradient Condition**) The general inclusion holds*

$$0 \in \mathbb{E}_{\mathbb{P}}[\boldsymbol{B}^* \boldsymbol{A}^{-*} T^*(T S \bar{z} + T u_f - u_d)] + \nu \bar{z} + \mathcal{N}_{Z_{ad}}(\bar{z}) + \boldsymbol{B}^* \boldsymbol{A}^{-*} \iota^* \bar{\mu}.$$

Here, the latter term must be understood

$$\langle \boldsymbol{B}^* \boldsymbol{A}^{-*} \iota^* \bar{\mu}, \delta z \rangle = \int_{\Xi} (\boldsymbol{A}^{-1}(\omega) \boldsymbol{B}(\omega) \delta z)(x) \, \mathrm{d}\bar{\mu}(x, \omega)$$

for an arbitrary test function $\delta z \in L^2(D)$.

Conversely, if there exists a pair $(\bar{z}, \bar{\mu})$ such that (i)-(iii) hold, then $\bar{z}$ is an optimal solution of (2.14).

Remark 2 In this general setting, we cannot guarantee that $\bar{\mu}$ splits into a generalized product of measures that would allow us to write $\boldsymbol{B}^* \boldsymbol{A}^{-*} \iota^* \bar{\mu}$ using an expectation. We explain this in more detail following the proof. However, the subgradient condition in (iii) can be brought into a form slightly more familiar to PDE-constrained optimization by introducing an adjoint variable $\widehat{\lambda}$ that satisfies the pointwise adjoint equations

$$\int_D A(x, \omega) \nabla \widehat{\lambda}(x, \omega) \cdot \nabla \varphi(x) \, \mathrm{d}x = \int_D T^*(T S \bar{z} + T u_f - u_d) \varphi(x) \, \mathrm{d}x, \quad \mathbb{P}\text{-a.s.}$$

$$(3.2)$$

Using this term we can "unfold" the general subgradient condition into the inclusion

$$0 \in \mathbb{E}_{\mathbb{P}}[\boldsymbol{B}^* \widehat{\lambda}] + \nu \bar{z} + \mathcal{N}_{Z_{ad}}(\bar{z}) + \boldsymbol{B}^* \boldsymbol{A}^{-*} \iota^* \bar{\mu}$$

coupled to the adjoint Eq. (3.2), a state equation for $S\bar{z}$ and complementarity conditions for the state constraint. Similarly, in some settings we may choose sufficiently regular test functions φ and introduce an additional adjoint variable $\widetilde{\lambda}$ to simplify the term $\boldsymbol{B}^* \boldsymbol{A}^{-*} \iota^* \bar{\mu}$ using

$$\langle \widetilde{\lambda}, A\varphi \rangle = \int_{\Xi} \varphi \, \mathrm{d}\bar{\mu}(x, \omega).$$

Indeed, $\bar{\mu}$ defines a bounded linear functional on $L_{\pi}^{\infty}(\Xi)$. So if A can define a linear isomorphism between a subspace of $L_{\pi}^{\infty}(\Xi)$ and, e.g., $L_{\pi}^2(\Xi)$, then $\widetilde{\lambda}$ would be a (very weak) solution. These two observations could potentially be used in a numerical setting, especially if $\mathbb{P}$ is discrete.

Proof To prove existence, we need to argue that the feasible set is weakly sequentially closed and $F(z) := \mathbb{E}[J(S(z) + u_f, z)]$ is weakly sequentially lower semicontinuous

on $L^2(D)$. Since the assumptions on J imply F is convex and the latter component of J is deterministic and continuous, we concentrate on the properties of S and their relation to the first argument of J.

By Assumption 2, (2.14) admits a feasible point and consequently a minimizing sequence $\{z_k\} \subset Z_{\mathrm{ad}}$ such that (2.12) holds. Since Z_{ad} is bounded, closed, and convex, $\{z_k\}$ admits a weakly convergent subsequence $\{z_{k_l}\}$. For each l, we have

$$S(z_{k_l}) + u_f \geq \psi \text{ for } (\mathcal{L} \times \mathbb{P})\text{-a.e. } (x, \omega) \in D \times \Omega.$$

Since S is completely continuous as a mapping into $L^2_{\mathbb{P}}(\Omega; H^1_0(D))$, we have $S(z_{k_l}) \to S(\bar{z})$ strongly. Moreover, the Sobolev embedding theorem (see e.g. [2, 4.12 Theorem]) and the fact that $L^p(\Omega; X) \hookrightarrow L^q(\Omega; Y)$ if $X \hookrightarrow Y$ for $1 \leq q \leq p < \infty$ plus the equivalence of $L^1_{\mathbb{P}}(\Omega; L^1(D))$ and $L^1_\pi(\Xi)$ (see e.g. [36, Proposition 1.2.24]) imply that $S(z_{k_l}) \to S(\bar{z})$ in $L^1_\pi(\Xi)$. Therefore, there exists a subsequence $\{z_{k_{l_m}}\}$ such that $S(z_{k_{l_m}}) \to S(\bar{z})$ π-pointwise almost everywhere. It follows that

$$S(\bar{z}) + u_f \geq \psi \text{ for } (\mathcal{L} \times \mathbb{P})\text{-a.e. } (x, \omega) \in D \times \Omega.$$

Continuing, the integrand J induces a superposition operator that is continuous from the product space $L^2_{\mathbb{P}}(\Omega; H^1_0(D)) \times L^2(D)$ to $L^1_{\mathbb{P}}(\Omega)$, see e.g., [44, Ex. 3.2]. Then by combining the properties of S with this continuity result, we deduce the weak lower semicontinuity of F. It follows from the direct method that $\bar{z}$ is an optimal solution, which is of course unique if $\nu > 0$ as the objective would be strictly convex.

In order to derive first order optimality conditions for (2.14), we write

$$\min_{z \in Z_{\mathrm{ad}}} \left\{ \mathbb{E}_{\mathbb{P}} \left[J(A^{-1}Bz + u_f, z) \right] \mid G(z) \in K \right\}$$

and appeal to the general Lagrangian formalism in [10, Chap. 3]. Here, we set

$$G(z) = \iota A^{-1}Bz + \iota u_f - \psi \text{ and } K = L^\infty_\pi(\Xi)_+,$$

where $\iota : L^\infty_{\mathbb{P}}(\Omega; H^1_0(D) \cap H^2(D))) \to L^\infty_\pi(\Xi)$ is the continuous embedding and K is the convex cone of all positive essentially bounded $\mathcal{B}$-measurable functions. Note that we first use the continuous embedding of $L^\infty_{\mathbb{P}}(\Omega; H^1_0(D) \cap H^2(D))$ into $L^\infty_{\mathbb{P}}(\Omega; L^\infty(D))$ and then the continuous embedding of $L^\infty_{\mathbb{P}}(\Omega; L^\infty(D))$ into $L^\infty_\pi(\Xi)$ to define ι. The latter two spaces are not equivalent.

Since K has a nonempty interior and G is clearly convex with respect to the partial order induced by $(-K)$, (3.1) is equivalent to the constraint qualification $0 \in \mathrm{int} \{G(Z_{\mathrm{ad}}) - K\}$ (and therefore Robinson's CQ), cf. [10, Prop. 2.106]. It follows from [10, Thm. 3.6] that

$$0 \in \partial_z L(\bar{z}, \bar{\mu}) + \mathcal{N}_{Z_{\mathrm{ad}}}(\bar{z}) \text{ and } \bar{\mu} \in \mathcal{N}_K(G(\bar{z})),$$

where $L(z, \mu) = F(z) + \langle G(z), \mu \rangle$. Due to convexity, these are both necessary and sufficient for optimality. It remains to make the conditions more explicit.

Since K is a closed, convex cone, $\bar{\mu} \in \mathcal{N}_K(G(\bar{z}))$ yields assertions (i) and (ii). To obtain the form in (iii), we first note that

$$F'(\bar{z})(\delta z) = \mathbb{E}_{\mathbb{P}}[(TA^{-1}B\bar{z} + Tu_f - u_d, TA^{-1}B\delta z)] + \nu(\bar{z}, \delta z)$$

and

$$\langle G(\cdot), \bar{\mu}\rangle'(\bar{z}; \delta z) = \langle \iota A^{-1}B(\delta z), \bar{\mu}\rangle.$$

For the objective function F, we can exploit the equivalence with the pointwise adjoints and write

$$\mathbb{E}_{\mathbb{P}}[(B^*A^{-*}T^*(TA^{-1}B\bar{z} + Tu_f - u_d), \delta z)] + \nu(\bar{z}, \delta z).$$

Furthermore, the uniform integrability of the operators A, B, i.e. B^*, A^{-*} allows us to write via [28, Thm. 3.7.12]

$$F'(\bar{z})\delta z = (\mathbb{E}_{\mathbb{P}}[(B^*A^{-*}T^*(TA^{-1}B\bar{z} + Tu_f - u_d)] + \nu\bar{z}, \delta z). \tag{3.3}$$

This concludes the proof. $\qquad\square$

We caution the reader that the form of the duality pairing used for the μ-multiplier initially does not include the expectation with respect to $\mathbb{P}$. However, if μ is σ-finite and σ-additive, then by the Radon-Nikodym theorem, there exists a density ρ_μ such that

$$\mu(A \times B) = \int_{A \times B} \rho_\mu \mathrm{d}\pi = \int_{A \times B} \rho_\mu \mathrm{d}(\mathcal{L} \times \mathbb{P}) = \int_A \int_B \rho_\mu \, \mathrm{d}x \mathrm{d}\mathbb{P}.$$

In other words, we would have $\mathrm{d}\mu = \rho_\mu \mathrm{d}(\mathcal{L} \times \mathbb{P})$. Furthermore, The sign condition on μ carries over to ρ_μ, in which case $|\rho_\mu| = -\rho_\mu$. This would indicate that $\rho_\mu \in L^1_\pi(\Xi)$. We could then write

$$\int_\Xi (A^{-1}(\omega)B(\omega)\delta z)(x) \, \mathrm{d}\bar{\mu}(x, \omega) = \mathbb{E}_{\mathbb{P}}[(A^{-1}(\omega)B(\omega)\delta z, \rho_\mu)]$$

by Fubini's theorem.

Furthermore, note that [28, Theorem 3.8.1] together with [42, Lemma 2.1] ensures that B^* is a bounded linear operator from $\mathcal{U}$ into $L^p_{\mathbb{P}}(\Omega; L^2(D))$. This would then allow us to incorporate the density into the adjoint equation, which is formulated in a very weak sense. This is essential, as otherwise the dual pairing with ρ_μ and the test functions would not be defined.

3.2 Risk-neutral PDE-constrained GNEPs

In order to prove existence of at least one generalized Nash equilibrium and link the proof to a function-space-based numerical algorithm, we restrict ourselves to a

variational reformulation as mentioned earlier. The variational reformulation is based on the so-called Nikaido-Isoda function $\Psi : L^2(D)^N \times L^2(D)^N \to \mathbb{R}$. For our GNEP the Nikaido-Isoda function is given by

$$\Psi(z, v) = \sum_{i=1}^N E_\mathbb{P}\left[J_i(S(z_i, z_{-i}) + u_f, (z_i, z_{-i}))\right]$$
$$-E_\mathbb{P}\left[J_i(S(v_i, z_{-i}) + u_f, (v_i, z_{-i}))\right].$$

We then introduce the potentially set-valued function $\widehat{\mathcal{R}} : \mathcal{Z}_{ad} \rightrightarrows \mathcal{Z}_{ad}$ given by

$$\widehat{\mathcal{R}}(z) = \arg \max \left\{\Psi(z, v) \mid v \in \mathcal{Z}_{ad} \text{ such that } S(v_i, v_{-i}) + u_f \geq \psi\right\}.$$

This mapping acts as a collective best-response function to a strategy vector $z \in Z_{ad}$ for all players simultaneously. Next, we define variational equilibria by their characterization as fixed points of the best-response function $\widehat{\mathcal{R}}$. The nomenclature diverges somewhat from the literature, but it should be clear in context what is meant below.

Definition 2 A strategy vector $\bar{z}$ with $\bar{z} \in \mathcal{Z}_{ad}$ and $S(\bar{z}_i, \bar{z}_{-i}) + u_f \geq \psi$ for $(\mathcal{L} \times \mathbb{P})$-a.e. $(x, \omega) \in D \times \Omega$ is a variational equilibrium if and only if $\bar{z} \in \widehat{\mathcal{R}}(\bar{z})$.

Note that for jointly convex GNEPs, every variational equilibrium is also a Nash equilibrium [34, Theorem 3.2]. This characterization converts the proof of the existence of Nash Equilibria to a fixed point problem. The essential ingredient is the fixed point theorem of Kakutani-Fan-Glicksberg, see e.g. [4, Corollary 17.55].

Theorem 5 *Let Assumptions 1 and 3 hold. The set of variational equilibria of the jointly convex GNEP (2.15) is weakly compact and nonempty.*

Proof We proceed as in [34, Theorem 3.2], in order to apply the fixed point theorem of Kakutani-Fan-Glicksberg on $\widehat{\mathcal{R}}$. By adapting the proof to the current setting, it follows from Theorem 4 that $\widehat{\mathcal{R}}$ has nonempty and convex images.

To ensure compactness, we recast the problem in the space X_i, where X_i is $L^2(D)$ endowed with the weak topology. Note that X_i is a real locally convex topological space. The equivalence of weak and strong closure for convex sets in reflexive Banach spaces implies that Z_{ad}^i is closed in X_i. Moreover, the weak compactness of closed and bounded convex subsets in reflexive Banach spaces implies that each set Z_{ad}^i is convex and compact in X_i or equivalently sequentially compact (see [65, Satz VIII.6.1(Satz von Eberlein-Shmulyan)]). Consequently, if we take $\mathcal{Z}_{ad} = Z_{ad}^1 \times \cdots \times Z_{ad}^N$ and $X = X_1 \times \cdots \times X_N$, then $\mathcal{Z}_{ad} \subset X$, where $\mathcal{Z}_{ad}$ is also nonempty, convex and compact in X. Due to the latter property, the weak topology is metrizable on $\mathcal{Z}_{ad}$ (see [65, Lemma VIII.6.2]).

In order to see the closedness of the graph of $\widehat{\mathcal{R}}$, we introduce the set

$$\mathcal{X}_{ad} = \left\{v \in \mathcal{Z}_{ad} \text{ such that } S(v_i, v_{-i}) + u_f \geq \psi\right\}.$$

Now, we consider a closed subset $C \subset \mathcal{X}_{\mathrm{ad}}$ and a sequence $\{z^n\}_{n \in \mathbb{N}} \subset \widehat{\mathcal{R}}^{-1}(C)$ with $z^n \to \bar{z}$ in X (i.e. $z^n \rightharpoonup \bar{z}$ in $L^2(D)^N$). For every z^n we choose $v^n \in C \cap \hat{\mathcal{R}}(z^n)$. By a slight adaptation of the arguments in the proof of Theorem 4, we can show that $\mathcal{X}_{\mathrm{ad}}$ is sequentially compact. Hence, there exists a convergent subsequence $v^{n_k} \xrightarrow{X} \bar{v}$ with $\bar{v} \in C$.

For some arbitrary $w \in \mathcal{X}_{\mathrm{ad}}$ it holds that

$$\sum_{i=1}^{N} E_{\mathbb{P}} \left[J_i(S(v_i^{n_k}, z_{-i}^{n_k}) + u_f, (v_i^{n_k}, z_{-i}^{n_k})) \right]$$
$$\leq \sum_{i=1}^{N} E_{\mathbb{P}} \left[J_i(S(w_i, z_{-i}^{n_k}) + u_f, (w_i, z_{-i}^{n_k})) \right]$$

By adapting the proof of Theorem 4, we can argue that

$$\sum_{i=1}^{N} E_{\mathbb{P}} \left[J_i(S(\bar{v}_i, \bar{z}_{-i}) + u_f, (\bar{v}_i, \bar{z}_{-i})) \right]$$
$$\leq \liminf_{k \to \infty} \sum_{i=1}^{N} E_{\mathbb{P}} \left[J_i(S(v_i^{n_k}, z_{-i}^{n_k}) + u_f, (v_i^{n_k}, z_{-i}^{n_k})) \right]$$
$$\leq \liminf_{k \to \infty} \sum_{i=1}^{N} E_{\mathbb{P}} \left[J_i(S(w_i, z_{-i}^{n_k}) + u_f, (w_i, z_{-i}^{n_k})) \right].$$

This is a consequence of the properties of the expectation, the objectives J_i and the solution operator S. In particular, it is essential that S is completely continuous into $L_{\mathbb{P}}^1(\Omega; H_0^1(D))$. Using again the complete continuity, we have

$$\sum_{i=1}^{N} E_{\mathbb{P}} \left[J_i(S(\bar{v}_i, \bar{z}_{-i}) + u_f, (\bar{v}_i, \bar{z}_{-i})) \right]$$
$$\leq \liminf_{k \to \infty} \sum_{i=1}^{N} E_{\mathbb{P}} \left[J_i(S(w_i, z_{-i}^{n_k}) + u_f, (w_i, z_{-i}^{n_k})) \right]$$
$$\leq \sum_{i=1}^{N} E_{\mathbb{P}} \left[J_i(S(w_i, \bar{z}_{-i}) + u_f, (w_i, \bar{z}_{-i})) \right].$$

It follows that $\bar{v} \in \widehat{\mathcal{R}}(\bar{z})$, which proves the sequential closedness of the graph of $\widehat{\mathcal{R}}$ or equivalently the closedness in X ([65, Theorem B.1.2]). We now apply Kakutani-Fan-Glicksberg's fixed point theorem. The set of Nash equilibria of the GNEP is nonempty and compact in X and thus, weakly compact in $L^2(D)^N$. $\square$

The optimality conditions for a generalized Nash equilibria reads as follow. We adapt the same notation as Theorem 4.

Theorem 6 *Let Assumptions 1 and 3 hold. If there exists a $(z_i^0, z_{-i}^0) \in \mathcal{Z}_{ad}$ and a constant $\kappa > 0$ such that*

$$S(z_i^0, z_{-i}^0) + u_f - \psi > \kappa \tag{3.4}$$

then there exists a measure $\bar{\mu} \in \mathbf{ba}(\varXi, \mathcal{B}, \pi)$ such that

*(i) (**Nonpositivity**) $\bar{\mu}$ is an element of the polar cone of $L_\pi^\infty(\varXi)_+$.*
*(ii) (**Complementarity**) $\bar{\mu}$ fulfills*

$$\int_\varXi G(\bar{z}_i, \bar{z}_{-i})(x, \omega) \, d\bar{\mu}(x, \omega) = 0.$$

*(iii) (**Subgradient Conditions**) For $i = 1, \ldots, N$ the general inclusion holds*

$$0 \in \mathbb{E}_\mathbb{P}[\boldsymbol{B}_i^* \boldsymbol{A}^{-*} T_i^* (T_i S(\bar{z}_i, \bar{z}_{-i}) + T_i u_f - u_d^i)] + v_i \bar{z}_i + \mathcal{N}_{\mathcal{Z}_{ad}^i}(\bar{z}_i) + \boldsymbol{B}_i^*(\boldsymbol{A}^{-*}\iota^*\bar{\mu}).$$

Conversely, if there exists a pair $(\bar{z}, \bar{\mu})$ such that (i)-(iii) hold, then $\bar{z}$ is generalized Nash equilibrium of (2.15).

Proof Similar to the proof of Theorem 4, we work with the general Lagrangian formalism. We first note that $\bar{z} \in \widehat{\mathcal{R}}(\bar{z})$. This is equivalent to

$$\bar{z} \in \arg \max_{v \in \mathcal{Z}_{ad}} \left\{ \Psi(\bar{z}, v) \mid S(v_i, v_{-i}) + u_f \geq \psi \right\}$$

$$= \arg \min_{v \in \mathcal{Z}_{ad}} \left\{ \sum_{i=1}^N \mathbb{E}_\mathbb{P} \left[J_i(\boldsymbol{A}^{-1}\boldsymbol{B}(v_i, \bar{z}_{-i}) + u_f, (v_i, \bar{z}_{-i})) \right] \mid S(v_i, v_{-i}) + u_f \geq \psi \right\}$$

In order to derive first order optimality conditions for variational equilibria of (2.15), we recall some of the notation from the proof of Theorem 4 We again set

$$G(v_i, v_{-i}) = \iota \boldsymbol{A}^{-1}\boldsymbol{B}(v_i, v_{-i}) + \iota u_f - \psi \text{ and } K = L_\pi^\infty(\varXi)_+,$$

and define the continuous embedding $\iota : L_\mathbb{P}^\infty(\varOmega; H_0^1(D) \cap H^2(D))) \to L_\pi^\infty(\varXi)$. In the notation of [10], we set

$$F_{\bar{z}}(v_i, v_{-i}) := \sum_{i=1}^N \mathbb{E}_\mathbb{P} \left[J_i(\boldsymbol{A}^{-1}\boldsymbol{B}(v_i, \bar{z}_{-i}) + u_f, (v_i, \bar{z}_{-i})) \right],$$

which yields the parametric Lagrangian

$$L_{\bar{z}}(v_i, v_{-i}, \mu) = F_{\bar{z}}(v_i, v_{-i}) + \langle G(v_i, v_{-i}), \mu \rangle.$$

Since (3.4) is equivalent to the constraint qualification $0 \in \text{int} \{G(\mathcal{Z}_{ad}) - K\}$, it follows from [10, Thm. 3.6] that

$$0 \in \partial_z L_{\bar{z}}(\bar{z}_i, \bar{z}_{-i}, \bar{\mu}) + \mathcal{N}_{\mathcal{Z}_{ad}}(\bar{z}_i, \bar{z}_{-i}) \text{ and } \bar{\mu} \in \mathcal{N}_K(G(\bar{z}_i, \bar{z}_{-i})).$$

Assertions (i) and (ii) are implied by $\bar{\mu} \in \mathcal{N}_K(G(\bar{z}_i, \bar{z}_{-i}))$ since K is a closed, convex cone. To obtain the subgradient conditions in (iii), we first note that

$$\langle G(\cdot), \bar{\mu}\rangle'(\bar{z}; \delta z) = \langle \iota A^{-1} B(\delta z), \bar{\mu}\rangle.$$

For the objective function, it holds that

$$\partial F_{\bar{z}}(\bar{z}_i, \bar{z}_{-i}) = \partial \left(\sum_{i=1}^{N} \mathbb{E}_{\mathbb{P}} \left[J_i(A^{-1}B((\cdot)_i, \bar{z}_{-i}) + u_f, ((\cdot)_i, \bar{z}_{-i})) \right] \right) (\bar{z})$$

$$= \sum_{i=1}^{N} \partial \left(\mathbb{E}_{\mathbb{P}} \left[J_i(A^{-1}B((\cdot)_i, \bar{z}_{-i}) + u_f, ((\cdot)_i, \bar{z}_{-i})) \right] \right) (\bar{z})$$

$$= \prod_{i=1}^{N} \left(\partial_i \mathbb{E}_{\mathbb{P}} \left[J_i(A^{-1}B(\cdot, \bar{z}_{-i}) + u_f, (\cdot, \bar{z}_{-i})) \right] \right) (\bar{z}_i). \qquad (3.5)$$

In order to see that the sum of the subdifferentials equals the product in (3.5), we refer to the proof of [34, Theorem 3.7]. Analogously to (3.3), we can write

$$F_{\bar{z}}'(\bar{z}_i, \bar{z}_{-i})\delta z_i = (\mathbb{E}_{\mathbb{P}}[(B_i^* A^{-*} T_i^*(T_i A^{-1} B(\bar{z}_i, \bar{z}_{-i}) + T_i u_f - u_d^i)] + \nu_i \bar{z}_i, \delta z_i).$$

Moreover, [8, section 4.6] enables us to write the normal cones as

$$\mathcal{N}_{\mathcal{Z}_{\text{ad}}}(\bar{z}_i, \bar{z}_{-i}) = \mathcal{N}_{\prod_{i=1}^{N} Z_{\text{ad}}^i}(\bar{z}_i, \bar{z}_{-i}) = \prod_{i=1}^{N} \mathcal{N}_{Z_{\text{ad}}^i}(\bar{z}_i).$$

For $i = 1, \ldots, N$, we have the componentwise subgradient condition

$$0 \in \mathbb{E}_{\mathbb{P}}[B_i^* A^{-*} T_i^*(T_i S(\bar{z}_i, \bar{z}_{-i}) + T_i u_f - u_d^i)] + \nu_i \bar{z}_i + \mathcal{N}_{Z_{ad}^i}(\bar{z}_i) + B_i^*(A^{-*}\iota^* \bar{\mu}).$$

$$\square$$

4 A Moreau-Yosida regularization technique

The optimality conditions derived in Theorems 4 and 6 are not suitable for the development of algorithms. This is mainly due to the low regularity of the multiplier $\bar{\mu}$ for the state constraint. To remedy this issue, we propose a Moreau-Yosida (MY) regularization technique, similar to the studies [1, 30, 31, 34].

From the perspective of risk aversion, the MY-regularization can be seen as a measure of regret, e.g., as in [55], for the state constraint. We will also use several concentration inequalities below, which link MY-regularization to probability constraints. This further justifies the viability of the approach and provides a modeling solution for cases in which either the constraint qualification is hard to verify and/or it is not known if a feasible point for the original problem actually exists. To the best

of our knowledge, this is the first time that such concentration inequalities have been used in the context of MY-regularization for infinite dimensional optimization under uncertainty.

4.1 Approximation of the risk-neutral PDE-constrained GNEPs

More specifically, the γ-dependent regularized problem of player i in the risk-neutral PDE-constrained GNEP (2.15) reads as

$$\min_{z \in Z^i_{\text{ad}}} \left\{ \mathbb{E}_{\mathbb{P}} \left[J_i(S(z_i, z_{-i}) + u_f, (z_i, z_{-i})) + \frac{\gamma}{2} \|(\psi - (S(z_i, z_{-i}) + u_f))_+\|^2_{L^2(D)} \right] \right\},$$
(4.1)

where $\gamma > 0$. The usage of MY-regularization amounts to approximating the original GNEP by a more numerically tractable NEP. We will refer to this γ-dependent strategic game as NEP$_\gamma$.

4.2 Existence and optimality conditions

The existence of a Nash equilibrium for every $\gamma > 0$ follows by using almost identical arguments to those in Theorem 5. Moreover, the first-order conditions have a similar, but numerically more workable form. We state the following theorem for ease of reference.

Theorem 7 *Let Assumptions 1 and 3 hold. The set of variational equilibria of the jointly convex NEP$_\gamma$ (4.1) is weakly compact and nonempty. If $\bar{z}^\gamma \in \mathcal{Z}_{\text{ad}}$ is a Nash equilibrium, then we have the following necessary and sufficient optimality conditions: For each $i = 1, \ldots, N$*

$$0 \in \mathbb{E}_{\mathbb{P}}[\boldsymbol{B}^*_i \boldsymbol{A}^{-*} T^*_i (T_i S(\bar{z}^\gamma_i, \bar{z}^\gamma_{-i}) + T_i u_f - u^i_d)] + \nu_i \bar{z}^\gamma_i + \mathcal{N}_{Z^i_{ad}}(\bar{z}^\gamma_i)$$
$$+ \mathbb{E}_{\mathbb{P}}[\boldsymbol{B}^*_i (\boldsymbol{A}^{-*} \bar{\mu}^\gamma)],$$
(4.2)

where

$$\bar{\mu}^\gamma = -\gamma (\psi - (S(\bar{z}^\gamma_i, \bar{z}^\gamma_{-i}) + u_f))_+.$$

Corollary 7 allows us to introduce the adjoint variables $\bar{\lambda}^\gamma_i \in L^2_{\mathbb{P}}(\Omega; H^1_0(D))$ for $i = 1, \ldots, N$ and the associated adjoint equations:

$$\int_D A(x, \omega) \nabla \bar{\lambda}^\gamma_i (x, \omega) \cdot \nabla \varphi(x) \, dx$$
$$= \int_D (T^*_i (T_i \bar{u}^\gamma + T_i u_f - u^i_d) + \bar{\mu}^\gamma)) \varphi(x) \, dx, \quad \mathbb{P}\text{-a.s.}$$
(4.3)

$\varphi \in H_0^1(D)$ and $\bar{u} = S(\bar{z}_i^\gamma, \bar{z}_{-i}^\gamma)$. This simplifies (4.2) to

$$0 \in \mathbb{E}_{\mathbb{P}}[\boldsymbol{B}_i^* \bar{\lambda}_i^\gamma] + \nu_i \bar{z}_i^\gamma + \mathcal{N}_{Z_{ad}^i}(\bar{z}_i^\gamma).$$

4.3 Asymptotic considerations

We now investigate the behavior of NEP$_\gamma$ as $\gamma \to \infty$. This is important for both theoretical as well as numerical considerations. We closely follow the approach in [33]. In order to ensure consistency of the relaxed problems, we will require the fulfillment of a constraint qualification as introduced in [33].

Definition 3 We say that (2.15) satisfies the strict uniform feasible response constraint qualification (SUFR), if there exists an $\varepsilon > 0$ such that for all $i = 1, \ldots, N$ and $z_{-i} \in \mathcal{Z}_{ad}^{-i}$ there is a $v_i \in Z_{ad}^i$ that satisfies

$$S(v_i, z_{-i}) + u_f \geq \psi + \varepsilon \quad \mathbb{P}\text{-a.s., a.e. } D.$$

A few comments are in order. Traditional constraint qualifications such as the existence of a Slater point or in nonlinear programming the Mangasarian-Fromovitz constraint qualification (Robinson's CQ in infinite dimensions) were developed for optimization problems. They provide not only the existence of Lagrange multipliers, but also, they indicate a certain stability of the constraint set around the optimal solution. For example, the MFCQ gives us that the Lagrange multipliers associated with the point in question lie in a convex, compact polytope. In Theorem 4, it was enough to assume such a CQ without the need to adapt to the GNEP setting. However, for issues of approximation, we will see in the following that GNEPs require a much more robust CQ such as SUFR in order to exhibit the local stability needed to bound the dual variables; in this case the adjoint states and the constraint multipliers. From a game-theoretic perspective, we are requiring that each player has a feasible response to any strategy by its competitors such that the common state constraint is strictly uniformly fulfilled. Finally, as the current regularity assumptions on the random inputs only provide essential boundedness, we will need more regularity of the solutions.

Assumption 8 *(Higher Parametric Regularity)* The set Ω is a compact Polish space. The solution mapping $S(\cdot) + u_f$ is a continuous affine mapping from $L^2(D)^N$ into $C(\Omega; H_0^1(D) \cap H^2(D))$.

The need for Ω to be a compact Polish space will be evident in the following proof. Under weaker assumptions, we have already shown that the mapping $S(\cdot) + u_f$ is a continuous affine mapping from $L^2(D)^N$ into $L_{\mathbb{P}}^\infty(\Omega; H_0^1(D) \cap H^2(D))$. The continuity assumption is actually weaker than it appears and can be guaranteed under mild assumptions (continuity in ω) on $A(x, \omega)$, $B(\omega)$ and $f(x, \omega)$, cf. the results in [35, Section 6]. The main idea is to reformulate the random PDE as a parametric fixed point equation and apply classic results on parametric dependence of solutions to fixed point equations. We now state the main result of this section.

Theorem 9 *Suppose the GNEP (2.15) satisfies the Slater condition (3.1) and SUFR. If in addition Assumption 8 holds, then there exist sequences $\gamma_n \to \infty$ and*

- $\{z^{\gamma_n}\}_{n \in \mathbb{N}} \subset L^2(D)^N$,
- $\{u^{\gamma_n}\}_{n \in \mathbb{N}} \subset L^2_{\mathbb{P}}(\Omega; H^1_0(D) \cap H^2(D))$,
- $\{\lambda^{\gamma_n}\}_{n \in \mathbb{N}} \subset L^2_{\mathbb{P}}(\Omega; H^1_0(D) \cap H^2(D))^N$,
- $\{\eta^{\gamma_n}\}_{n \in \mathbb{N}} \subset L^2(D)^N$,
- $\{\mu^{\gamma_n}\}_{n \in \mathbb{N}} \subset L^2_\pi(\Xi)$,

such that for each $i = 1, \dots, N$, $(z_i^{\gamma_n}, u^{\gamma_n}, \lambda_i^{\gamma_n}, \eta_i^{\gamma_n}, \mu^{\gamma_n})$ satisfies (4.2) as stated in Corollary 7. This sequence admits a limit point

$$(z^*, u^*, \Lambda^*, \eta^*, \rho^*) \in L^2(D)^N \times L^2_{\mathbb{P}}(\Omega; H^1_0(D) \cap H^2(D)) \times L^2(D)^N \times L^2(D)^N \\ \times \mathcal{M}(\overline{\Xi}),$$

where, for all $i = 1, \dots, N$, we have

$$z^{\gamma_n} \to z^* \qquad \qquad in \ L^2(D)^N, \tag{4.4a}$$

$$u^{\gamma_n} \to u^* \qquad \qquad in \ C(\Omega; H^1_0(D) \cap H^2(D)), \tag{4.4b}$$

$$\mu^{\gamma_n} \overset{*}{\rightharpoonup} \rho^* \qquad \qquad in \ L^1_\pi(\Xi)^* \cong \mathcal{M}(\overline{\Xi}), \ i.e. \ \rho \in \mathcal{M}(\overline{\Xi}), \tag{4.4c}$$

$$\mathbb{E}_{\mathbb{P}}\left[B_i^* \lambda_i^{\gamma_n}\right] \rightharpoonup \Lambda_i^* \qquad in \ L^2(D), \tag{4.4d}$$

$$\eta_i^{\gamma_n} \rightharpoonup \eta_i^* \qquad \qquad in \ L^2(D). \tag{4.4e}$$

Moreover, the limit point satisfies

$$z^* \in \mathcal{Z}_{\mathrm{ad}} \tag{4.5a}$$

$$u^* = S(z_i^*, z_{-i}^*) + u_f \ and \ u^* \geq \psi \tag{4.5b}$$

$$\left(\Lambda_i^*, \varphi\right) = \left(\mathbb{E}_{\mathbb{P}}\left[B_i^* A^{-*} T_i^* (T_i(u^* + u_f) - u_d^i)\right], \varphi\right) \\ + \int_\Xi A^{-1}(\omega) B_i(\omega) \varphi \, d\rho^*(x, \omega) \tag{4.5c}$$

$$0 = \left(\Lambda_i^*, \varphi\right) + \nu_i \left(z_i^*, \varphi\right) + \left(\eta_i^*, \varphi\right) \ and \ \eta_i^* \in \mathcal{N}_{Z_{\mathrm{ad}}^i}(z_i^*) \tag{4.5d}$$

for an arbitrary test function $\varphi \in L^2(D)$. Finally, ρ^* satisfies

$$\langle \phi, \rho^* \rangle \leq 0, \quad \forall \phi \in C(\overline{\Xi}) : \phi \geq 0. \tag{4.6a}$$

$$\langle \psi - (u^* + u_f), \rho^* \rangle = 0 \tag{4.6b}$$

Note that (4.5c) and (4.5d) correspond to the subdifferential inclusion in Theorem 6. For readability, we split the proof over several partial results.

Lemma 1 *Under the assumptions of Theorem 9, there exists a sequence of MY parameters $\gamma_k \to \infty$ such that the associated sequence of Nash equilibria $\{z^{\gamma_k}\}_{k \to \infty}$ converges weakly to a feasible strategy of the GNEP, i.e. (4.5a) and (4.5b) hold.*

Proof Fix a sequence $\gamma_n \to \infty$ for $n \to \infty$. Since $\mathcal{Z}_{ad}$ is weakly compact in $L^2(D)^N$ and $z^{\gamma_n} \in \mathcal{Z}_{ad}$ for all γ^n, there exists a subsequence, denoted by $\gamma_k := \gamma_{n_k}$ and some element $z^* \in \mathcal{Z}_{ad}$ such that $z^{\gamma_k} \rightharpoonup z^*$ in $L^2(D)^N$. According to SUFR, there exists an $\varepsilon > 0$ and a sequence $\{v^{\gamma_k}\}_{k \to \infty} \subset \mathcal{Z}_{ad}$ such that $S(v_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f \geq \psi + \varepsilon$ $\mathit{\Xi}$-a.s. for all $i = 1, \ldots, N$. By definition of $\mathcal{Z}_{ad}$, $\{v^{\gamma_k}\}_{k \to \infty}$ is uniformly bounded in $L^2(D)^N$. Then for all γ_k, the non-negativity of the MY-term gives us the lower bound:

$$\mathbb{E}_{\mathbb{P}}\left[\frac{1}{2}\|T_i(S(z_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{v_i}{2}\|z_i^{\gamma_k}\|_{L^2(D)}^2$$
$$\leq \mathbb{E}_{\mathbb{P}}\left[\frac{1}{2}\|T_i(S(z_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{v_i}{2}\|z_i^{\gamma_k}\|_{L^2(D)}^2$$
$$+ \mathbb{E}_{\mathbb{P}}\left[\frac{\gamma_k}{2}\|(\psi - (S(z_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f)_+\|_{L^2(D)}^2\right].$$

Since $S(v_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f \geq \psi + \varepsilon$, it holds that

$$\mathbb{E}_{\mathbb{P}}\left[\frac{\gamma_k}{2}\|(\psi - (S(v_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f)_+\|_{L^2(D)}^2\right] = 0.$$

Furthermore, by definition of a Nash equilibrium we have the simple upper bound

$$\frac{1}{2}\mathbb{E}_{\mathbb{P}}\left[\|T_i(S(z_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{v_i}{2}\|z_i^{\gamma_k}\|_{L^2(D)}^2$$
$$+ \mathbb{E}_{\mathbb{P}}\left[\frac{\gamma_k}{2}\|(\psi - (S(z_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f)_+\|_{L^2(D)}^2\right]$$
$$\leq \mathbb{E}_{\mathbb{P}}\left[\|\frac{1}{2}T_i(S(v_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{v_i}{2}\|v_i^{\gamma_k}\|_{L^2(D)}^2$$

Using the fact that S is completely continuous into $L_\pi^2(\mathit{\Xi})$ and each individual feasible set is bounded, we deduce the existence of a constant M independent of i, γ_k such that

$$\mathbb{E}_{\mathbb{P}}\left[\|\frac{1}{2}T_i(S(v_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{v_i}{2}\|v_i^{\gamma_k}\|_{L^2(D)}^2 \leq M.$$

Combining these observations yields

$$\mathbb{E}_{\mathbb{P}}\left[\frac{1}{2}\|T_i(S(z_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{v_i}{2}\|z_i^{\gamma_k}\|_{L^2(D)}^2 \leq M.$$

Using the weak lower semicontinuity of the objective functions, it follows that the bound also holds for the limit

$$\mathbb{E}_{\mathbb{P}}\left[\frac{1}{2}\|T_i(S(z_i^*, z_{-i}^*) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{v_i}{2}\|z_i^*\|_{L^2(D)}^2$$
$$\leq \liminf_{k \to \infty}\left(\mathbb{E}_{\mathbb{P}}\left[\frac{1}{2}\|T_i(S(z_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{v_i}{2}\|z_i^{\gamma_k}\|_{L^2(D)}^2\right) \leq M.$$

As a result, $\mathbb{E}_{\mathbb{P}}\left[\frac{\gamma_k}{2}\|(\psi - (S(z_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f)_+\|_{L^2(D)}^2\right]$ is bounded. This can only hold if

$$\mathbb{E}_{\mathbb{P}}\left[\|(\psi - (S(z_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f)_+\|_{L^2(D)}^2\right] \to 0,$$

since $\gamma_k \to \infty$. Since $S(z_i^{\gamma_k}, z_{-i}^{\gamma_k})$ converges strongly to $S(z_i^*, z_{-i}^*)$ in $L_\pi^2(\Xi)$, we also have

$$\mathbb{E}_{\mathbb{P}}\left[\|(\psi - (S(z_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f)_+\|_{L^2(D)}^2\right] \to \mathbb{E}_{\mathbb{P}}\left[\|(\psi - (S(z_i^*, z_{-i}^*) + u_f)_+\|_{L^2(D)}^2\right].$$

We can conclude, that

$$\mathbb{E}_{\mathbb{P}}\left[\|(\psi - (S(z_i^*, z_{-i}^*) + u_f)_+\|_{L^2(D)}^2\right] = 0.$$

Thus, $z^* \in \mathcal{Z}_{\mathrm{ad}}$ such that $S(z_i^*, z_{-i}^*) + u_f \geq \psi$ π-a.e., i.e. z^* is a feasible strategy vector for the GNEP. $\qquad\square$

We note that for feasibility of z^*, it is not necessary for ε to be positive in the SUFR condition. In what follows, we discuss the convergence of the stationary points individually. We start by showing that z^* is also a generalized Nash equilibrium.

Lemma 2 *Suppose the assumptions of Theorem 9 hold. Let $\{\gamma_k\}$ be the sequence of MY parameters from the proof of Lemma 1. Then there exists a subsequence $\{\gamma_l\}$ with $\gamma_l := \gamma_{k_l} \to +\infty$ such that the weak limit point z^* is a generalized Nash equilibrium.*

Proof Define $X_i = \{v_i \in \mathcal{Z}_{\mathrm{ad}}^i : S(v_i, z_{-i}^*) + u_f \geq \psi$ π-a.s$\}$. Due to the SUFR condition, X_i is non-empty. Since for all γ_k the associated z^{γ_k} is a Nash equilibrium, it holds that

$$\frac{1}{2}\mathbb{E}_{\mathbb{P}}\left[\|T_i(S(z_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{\nu_i}{2}\|z_i^{\gamma_k}\|_{L^2(D)}^2$$
$$+ \mathbb{E}_{\mathbb{P}}\left[\frac{\gamma_k}{2}\|(\psi - (S(z_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f)_+\|_{L^2(D)}^2\right]$$
$$\leq \frac{\nu_i}{2}\|v_i\|_{L^2(D)}^2 + \frac{1}{2}\mathbb{E}_{\mathbb{P}}\left[\|T_i(S(v_i, z_{-i}^{\gamma_k}) + u_f) - u_d^i\|_{L^2(D)}^2\right]$$
$$+ \mathbb{E}_{\mathbb{P}}\left[\frac{\gamma_k}{2}\|(\psi - (S(v_i, z_{-i}^{\gamma_k}) + u_f)_+\|_{L^2(D)}^2\right]$$

for all $v_i \in X_i$. For any $v_i \in X_i$, we want to construct a strongly convergent sequence $\{v^{\gamma_k}\}_{k\to\infty}$ so such $v^{\gamma_k} \to v_i$ in $L^2(D)$ and $S(v_i^{\gamma_k}, z_{-i}^{\gamma_k}) + u_f \geq \psi$.

Due to the SUFR condition, there exists an $\varepsilon > 0$ and for all k, a $v_i^k \in \mathcal{Z}_{\mathrm{ad}}^i$ such that $S(v_i^k, z_{-i}^{\gamma_k}) + u_f \geq \psi + \varepsilon$, π-a.s. Clearly, $\{v_i^k\}_{k\to\infty}$ is uniformly bounded in $L^2(D)$. Since every admissible set of each player is convex, we have that

$$v_i^k(t) = tv_i^k + (1-t)v_i \tag{4.7}$$

lies in $\mathcal{Z}_{ad}^i$ for all $t \in (0, 1)$. Due to the linearity of the operator **A** and **B**, it holds that

$$
\begin{aligned}
S(v_i^k(t), z_{-i}^{\gamma_k}) + u_f &= S(tv_i^k + (1-t)v_i, z_{-i}^{\gamma_k}) + u_f \\
&= t(S(v_i^k, z_{-i}^{\gamma_k}) + u_f) + (1-t)(S(v_i, z_{-i}^{\gamma_k}) + u_f) \\
&\geq t(\psi + \varepsilon) + (1-t)(S(v_i, z_{-i}^{\gamma_k}) + u_f).
\end{aligned} \tag{4.8}
$$

We know, that for $\mathbb{P}$-a.e. $\omega \in \Omega$ the solution operator $S(v_i, \cdot)(\omega) + u_f(\omega)$ maps continuously from $L^2(D)^{N-1}$ into $H_0^1(D) \cap H^2(D)$. Due to the Sobolev and Rellich-Kondrachov theorem, the solution of the state equation can be continuously and compactly embedded into the space of continuous functions over $\bar{D}$ $\mathbb{P}$-a.s. Thus, $S(v_i, \cdot)(\omega) + u_f(\omega)$ maps from $L^2(D)^{N-1}$ into $C(\bar{D})$ for $\mathbb{P}$-a.e. $\omega \in \Omega$. Combining this with the regularity assumption on the solution of the state equation, we have $S(v_i, z_{-i}^{\gamma_k}) + u_f \to S(v_i, z_{-i}^*) + u_f$ in $C(\Omega; C(\bar{D}))$. Then by virtue of the nature of convergence in the $C(\Omega; C(\bar{D}))$-norm, we deduce the existence of a subsequence γ_{k_l}, denoted by γ_l, such that

$$
S(v_i, z_{-i}^{\gamma_l}) + u_f \geq \psi - 1/2^l
$$

on D for all k. Now, setting

$$
t_l = (1/2^l)/(\varepsilon + 1/2^l), \tag{4.9}
$$

then $t_l \to 0$ and $t_l \in (0, 1)$ for all l. Moreover, substituting (4.9) in (4.7) and due to (4.8), we have

$$
\begin{aligned}
S(v_i(t_l), z_{-i}^{\gamma_l}) + u_f &\geq t_l(\psi + \varepsilon) + (1-t_l)(S(v_i, z_{-i}^{\gamma_l}) + u_f) \\
&\geq t_l(\psi + \varepsilon) + (1-t_l)(\psi - 1/2^l) \\
&= \psi + t_l(\varepsilon + 1/2^l) - 1/2^l \\
&= \psi.
\end{aligned}
$$

Thus, $S(v_i(t_l), z_{-i}^{\gamma_l}) + u_f \geq \psi$ for all l. And finally, since

$$
\begin{aligned}
\|v_i(t_l) - v_i\|_{L^2(D)} &= \|t_l v_i^l + (1-t_l)v_i - v_i\|_{L^2(D)} \\
&= |t_l| \|v_i^l - v_i\|_{L^2(D)} \\
&\leq |t_l| \left(\|v_i^l\|_{L^2(D)} + \|v_i\|_{L^2(D)} \right)
\end{aligned}
$$

Passing to the limit as $l \to \infty$ yields $|t_l| \left(\|v_i^l\|_{L^2(D)} + \|v_i\|_{L^2(D)} \right) \to 0$ due to the boundedness of $\{v_i(t_l)\}_{l \to \infty}$ and that $\{t_l\}_{l \to \infty}$ is a null sequence. Thus, we have constructed a sequence $\{v_i(t_l)\}_{l \to \infty}$ such that $v_i(t_l) \to v_i$ in $L^2(D)$ and $S(v_i(t_l), z_{-i}^{\gamma_l}) + u_f \geq \psi$. Note that $v_i \in X_i$ was arbitrary.

Finally, by substitution, we have

$$\frac{1}{2}\mathbb{E}_{\mathbb{P}}\left[\|T_i(S(z_i^{\gamma_l}, z_{-i}^{\gamma_l}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{\nu_i}{2}\|z_i^{\gamma_l}\|_{L^2(D)}^2$$

$$+ \mathbb{E}_{\mathbb{P}}\left[\frac{\gamma_l}{2}\|(\psi - (S(z_i^{\gamma_l}, z_{-i}^{\gamma_l}) + u_f)_+\|_{L^2(D)}^2\right]$$

$$\leq \frac{1}{2}\mathbb{E}_{\mathbb{P}}\left[\|T_i(S(v_i(t_l), z_{-i}^{\gamma_l}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{\nu_i}{2}\|v_i(t_l)\|_{L^2(D)}^2$$

$$+ \mathbb{E}_{\mathbb{P}}\left[\frac{\gamma_l}{2}\|(\psi - (S(v_i(t_l), z_{-i}^{\gamma_l}) + u_f)_+\|_{L^2(D)}^2\right].$$

For all $i = 1, \ldots, N$, passing to the limit inferior yields the following inequality

$$\frac{1}{2}\mathbb{E}_{\mathbb{P}}\left[\|T_i(S(z_i^*, z_{-i}^*) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{\nu_i}{2}\|z_i^*\|_{L^2(D)}^2$$

$$\leq \frac{1}{2}\mathbb{E}_{\mathbb{P}}\left[\|T_i(S(v_i, z_{-i}^*) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{\nu_i}{2}\|v_i\|_{L^2(D)}^2$$

for all $v_i \in X_i$. Thus, (z_i^*, z_{-i}^*) is a generalized Nash equilibrium. $\qquad\square$

Here, we see that the uniformity in the SUFR condition is crucial to prove that z^* is in fact a Nash equilibrium. In the following result, we obtain a stronger form of convergence to z^*. This is necessary to derive the adjoint equation in the limit.

Lemma 3 *Under the assumptions of Theorem 9, (4.4a) holds.*

Proof First, we choose $z_i^* \in X_i$ in the construction of (4.7) with $t = t_l$ as in (4.9), then we have

$$v_i^*(t_l) = t_l v_i^k + (1 - t_l)z_i^*.$$

Recall that $v_i^*(t_l) \to z_i^*$ in $L^2(D)$ and $S(v_i^*(t_l), z_{-i}^{\gamma_l}) + u_f \geq \psi$ for all $l \in \mathbb{N}$. Then it holds that

$$\frac{1}{2}\mathbb{E}_{\mathbb{P}}\left[\|T_i(S(z_i^{\gamma_l}, z_{-i}^{\gamma_l}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{\nu_i}{2}\|z_i^{\gamma_l}\|_{L^2(D)}^2$$

$$\leq \frac{1}{2}\mathbb{E}_{\mathbb{P}}\left[\|T_i(S(z_i^{\gamma_l}, z_{-i}^{\gamma_l}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{\nu_i}{2}\|z_i^{\gamma_l}\|_{L^2(D)}^2$$

$$+ \mathbb{E}_{\mathbb{P}}\left[\frac{\gamma_l}{2}\|(\psi - (S(z_i^{\gamma_l}, z_{-i}^{\gamma_l}) + u_f)_+\|_{L^2(D)}^2\right]$$

$$\leq \frac{1}{2}\mathbb{E}_{\mathbb{P}}\left[\|T_i(S(v_i^*(t_l), z_{-i}^{\gamma_l}) + u_f) - u_d^i\|_{L^2(D)}^2\right] + \frac{\nu_i}{2}\|v_i^*(t_l)\|_{L^2(D)}^2.$$

Passing to the limit superior yields

$$\limsup_{l \to \infty} \frac{1}{2} \mathbb{E}_{\mathbb{P}} \left[\| T_i(S(z_i^{\gamma_l}, z_{-i}^{\gamma_l}) + u_f) - u_d^i \|_{L^2(D)}^2 \right] + \frac{\nu_i}{2} \| z_i^{\gamma_l} \|_{L^2(D)}^2$$

$$\leq \limsup_{l \to \infty} \frac{1}{2} \mathbb{E}_{\mathbb{P}} \left[\| T_i(S(v_i^*(t_l), z_{-i}^{\gamma_l}) + u_f) - u_d^i \|_{L^2(D)}^2 \right] + \frac{\nu_i}{2} \| v_i^*(t_l) \|_{L^2(D)}^2.$$

(4.10)

Due to the complete continuity of S, we have

$$S(z_i^{\gamma_l}, z_{-i}^{\gamma_l}) \to S(z_i^*, z_{-i}^*) \quad \text{and} \quad S(v_i^*(t_l), z_{-i}^{\gamma_l}) \to S(z_i^*, z_{-i}^*).$$

Then (4.10) reads as

$$\frac{1}{2} \mathbb{E}_{\mathbb{P}} \left[\| T_i(S(z_i^*, z_{-i}^*) + u_f) - u_d^i \|_{L^2(D)}^2 \right] + \frac{\nu_i}{2} \limsup_{l \to \infty} \| z_i^{\gamma_l} \|_{L^2(D)}^2$$

$$\leq \frac{1}{2} \mathbb{E}_{\mathbb{P}} \left[\| T_i(S(z_i^*, z_{-i}^*) + u_f) - u_d^i \|_{L^2(D)}^2 \right] + \frac{\nu_i}{2} \| z_i^* \|_{L^2(D)}^2. \qquad (4.11)$$

This implies that

$$\limsup_{l \to \infty} \| z_i^{\gamma_l} \|_{L^2(D)}^2 \leq \| z_i^* \|_{L^2(D)}^2.$$

Due to the weak convergence of $\{z_i^{\gamma_l}\}_{l \in \mathbb{N}}$, it holds that $\liminf_{l \to \infty} \| z_i^{\gamma_l} \|_{L^2(D)}^2 \geq \| z_i^* \|_{L^2(D)}^2$. This implies

$$\lim_{l \to \infty} \| z_i^{\gamma_l} \|_{L^2(D)}^2 = \| z_i^* \|_{L^2(D)}^2.$$

Together with the weak convergence, the assertion follows.   □

We proceed with the sequence of the state variables.

Lemma 4 *Under the assumptions of Theorem 9, (4.4b) and (4.5b) hold.*

Proof This directly follows from the assumption, that $S(\cdot, \cdot) + u_f : L^2(D)^N \to C(\Omega; H_0^1(D) \cap H^2(D))$ is continuous and the fact the sequences $\{z_i^{\gamma_l}\}_{l \in \mathbb{N}}$ converges strongly in $L^2(D)$ for all $i = 1, \dots, N$.   □

We note that the continuity in Ω is not really needed to prove a norm convergence result. Indeed since $\{z_i^{\gamma_l}\}_{l \in \mathbb{N}}$ is bounded, we still have that $\{u^{\gamma_l}\}_{l \in \mathbb{N}}$ is bounded in $L_{\mathbb{P}}^2(\Omega; H_0^1(D) \cap H^2(D))$. Then $u^{\gamma_{l_n}} \rightharpoonup u^*$ in $L_{\mathbb{P}}^2(\Omega; H_0^1(D) \cap H^2(D))$. By Corollary 1 we even know that $u^{\gamma_{l_n}} \to u^*$ in $L_{\mathbb{P}}^\infty(\Omega; H_0^1(D) \cap H^2(D))$ holds.

Next, we turn our attention to the sequence of the multipliers μ^γ for the state constraint. We will observe that the Slater condition is enough to obtain a bound on μ^γ. Recall that $\mu^\gamma = -\gamma(\psi - (S(\bar{z}_i^\gamma, \bar{z}_{-i}^\gamma) + u_f))_+$.

Lemma 5 *Suppose the assumptions of Theorem 9 hold. In particular, (3.1) is fulfilled. Then we have (4.4c).*

Proof We now prove the existence of a constant $c_0 > 0$ such that

$$|(\mu^\gamma, z)| \leq c_0 \tag{4.12}$$

for any $z \in \mathbb{B}_\varepsilon(0) \subset L_\pi^\infty(\Xi)$ and some fixed $\varepsilon > 0$. For the sake of readability, we set $\beta : L_\pi^2(\Xi) \to \mathbb{R}_+$ such that

$$u \mapsto \beta(u) := \mathbb{E}\left[\frac{1}{2}\|(\psi - (u + u_f))_+\|_{L^2(D)}^2\right].$$

Unless otherwise noted, $(\cdot, \cdot)$ denotes the inner product on $L_\pi^2(\Xi)$ throughout the proof.

One readily shows that β is convex and continuously differentiable and therefore, $\mu^\gamma = \gamma\beta'(u^\gamma)$. Since β is convex, differentiable, and nonnegative, we obtain for any

$$y \in \left\{w \in L_\pi^2(\Xi) \ : \ w \geq \psi - u_f \ \pi\text{-a.e.}\right\}$$

the equality

$$0 = \gamma\beta(y) \geq \gamma\beta(u^\gamma) + \left(\mu^\gamma, y - u^\gamma\right) \geq \left(\mu^\gamma, y - u^\gamma\right). \tag{4.13}$$

By the assumption (3.4) there exists $\varepsilon > 0$ and $z^0 \in \mathcal{Z}_{\text{ad}}$ such that for all $v \in \mathbb{B}_\varepsilon(0) \subset L_\pi^\infty(\Xi)$: we have

$$Sz^0 + u_f - \psi + v \geq 0.$$

Since $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space and the spatial domain D is bounded, the Lebesgue spaces are nested, and it holds that $v \in L_\pi^2(\Xi)$. Furthermore, $Sz^0 + u_f \in L_\pi^2(\Xi)$. Fixing an arbitrary $v \in \mathbb{B}_\varepsilon(0)$, we have

$$\left(\mu^\gamma, v\right) = \left(\mu^\gamma, Sz^0 + v - Sz^\gamma\right) + \left(\mu^\gamma, Sz^\gamma - Sz^0\right).$$

Due to (4.13), we have

$$\left(\mu^\gamma, v\right) \leq \left(\mu^\gamma, Sz^\gamma - Sz^0\right) = \left(\mu^\gamma, A^{-1}B(z^\gamma - z^0)\right).$$

The definition of the multiplier μ^γ and the operator B yield

$$(\mu^\gamma, A^{-1}B(z^\gamma - z^0))$$
$$= \sum_{i=1}^N (T_i^*\left(T_i u^\gamma + T_i u_f - u_d^i\right) + \mu^\gamma - T_i^*\left(T_i u^\gamma + T_i u_f - u_d^i\right), A^{-1}B_i(z_i^\gamma - z_i^0))$$
$$= \sum_{i=1}^N (A^*\lambda_i^\gamma - T_i^*T_i u^\gamma - T_i^*T_i u_f + T_i^*u_d^i, A^{-1}B_i(z_i^\gamma - z_i^0))$$

Substituting the adjoint equation and applying the adjoint operator $\boldsymbol{B}^* \boldsymbol{A}^{-*}$ yields

$$
(\mu^\gamma, v) \le \sum_{i=1}^{N} (A^* \lambda_i^\gamma - T_i^* T_i u^\gamma - T_i^* T_i u_f + T_i^* u_d^i, A^{-1} B_i (z_i^\gamma - z_i^0))
$$

$$
= \sum_{i=1}^{N} \mathbb{E}_{\mathbb{P}} \left[(\boldsymbol{B}_i^* \lambda_i^\gamma - \boldsymbol{B}_i^* A^{-*} (T_i^* T_i u^\gamma + T_i^* T_i u_f - T_i^* u_d^i), z_i^\gamma - z_i^0)_{L^2(D)} \right]
$$

$$
= \sum_{i=1}^{N} \mathbb{E}_{\mathbb{P}} \left[(\boldsymbol{B}_i^* \lambda_i^\gamma, z_i^\gamma - z_i^0)_{L^2(D)} - \left(\boldsymbol{B}_i^* A^{-*} (T_i^* T_i u^\gamma + T_i^* T_i u_f - T_i^* u_d^i), z_i^\gamma - z_i^0 \right)_{L^2(D)} \right]
$$

Applying [28, Thm. 3.7.12] yields

$$
(\mu^\gamma, z) \le \sum_{i=1}^{N} (\mathbb{E}_{\mathbb{P}} \left[\boldsymbol{B}_i^* \lambda_i^\gamma \right], z_i^\gamma - z_i^0)_{L^2(D)}
$$

$$
- \mathbb{E}_{\mathbb{P}} \left[\left(\boldsymbol{B}_i^* A^{-*} (T_i^* T_i u^\gamma + T_i^* T_i u_f - T_i^* u_d^i), z_i^\gamma - z_i^0 \right)_{L^2(D)} \right]
$$

Using $0 = v_i z_i^\gamma + \mathbb{E}_{\mathbb{P}} \left[\boldsymbol{B}_i^* \lambda_i^\gamma \right] + \eta_i^\gamma$ and the fact, that $z_i^0 \in Z_{\text{ad}}^i$ yields

$$
\sum_{i=1}^{N} \left(\mathbb{E}_{\mathbb{P}} \left[\boldsymbol{B}_i^* \lambda_i^\gamma \right], z_i^\gamma - z_i^0 \right)_{L^2(D)}
$$

$$
- \sum_{i=1}^{N} \mathbb{E}_{\mathbb{P}} \left[\left(\boldsymbol{B}_i^* A^{-*} (T_i^* T_i u^\gamma + T_i^* T_i u_f - T_i^* u_d^i), z_i^\gamma - z_i^0 \right)_{L^2(D)} \right]
$$

$$
= \sum_{i=1}^{N} \left(-v_i z_i^\gamma - \eta_i^\gamma, z_i^\gamma - z_i^0 \right)_{L^2(D)}
$$

$$
- \sum_{i=1}^{N} \mathbb{E}_{\mathbb{P}} \left[\left(\boldsymbol{B}_i^* A^{-*} (T_i^* T_i u^\gamma + T_i^* T_i u_f - T_i^* u_d^i), z_i^\gamma - z_i^0 \right)_{L^2(D)} \right]
$$

$$
\le \sum_{i=1}^{N} \left(-v_i z_i^\gamma, z_i^\gamma - z_i^0 \right)_{L^2(D)}
$$

$$
- \sum_{i=1}^{N} \mathbb{E}_{\mathbb{P}} \left[\left(\boldsymbol{B}_i^* A^{-*} (T_i^* T_i u^\gamma + T_i^* T_i u_f - T_i^* u_d^i), z_i^\gamma - z_i^0 \right)_{L^2(D)} \right]
$$

$$
\le c_0 < \infty.
$$

Here, the existence of c_0 is guaranteed, since the mappings

$$
u(z) \mapsto \mathbb{E}_{\mathbb{P}} \left[\frac{1}{2} \| T_i u(z) + T_i u_f - u_d^i \|_{L^2(D)}^2 \right] \quad \text{and} \quad z_i \mapsto \frac{v_i}{2} \| z_i \|_{L^2(D)}^2
$$

are continuously differentiable with uniformly bounded gradients on $\mathcal{Z}_{\text{ad}}$ for all $i = 1, \ldots, N$. This proves (4.12), since z was arbitrary. Using the fact that the L^1-norm

is positively homogeneous, subadditive and continuous, it follows from the Fenchel-Moreau theorem that the L^1-norm is equivalent to the bidual norm

$$
\begin{aligned}
\|\mu^\gamma\|_{L^1_\pi(\Xi)} &= \frac{1}{\varepsilon} \sup \left\{ \langle \mu^\gamma, z \rangle_{L^1_\pi(\Xi) \times L^\infty_\pi(\Xi)} \ : \ z \in \mathbb{B}_\varepsilon(0) \right\} \\
&= \frac{1}{\varepsilon} \sup \left\{ (\mu^\gamma, z)_{L^2_\pi(\Xi)} \ : \ z \in \mathbb{B}_\varepsilon(0) \right\} \\
&\leq \frac{1}{\varepsilon} c_0 < \infty.
\end{aligned}
$$

It follows that the sequence $\{\mu^\gamma\}_{\gamma \to \infty}$ is bounded in $L^1_\pi(\Xi)$. Therefore, by [17, Theorem IV.6.2] or [7, Corollary 2.4.3], we can extract a subsequence $\{\mu^{\gamma_l}\}_{l \in \mathbb{N}}$ which is weak* convergent to some regular countably additive Borel measure $\rho \in \mathcal{M}(\overline{\Xi})$. □

Next, we discuss the limit of the adjoint equation. We start by investigating the behavior of the expectation of the adjoint states. This leads to the derivation of a limiting adjoint state Λ^*.

Lemma 6 *Under the assumptions of Theorem 9, for all $i = 1, \ldots, N$, (4.4d) holds.*

Proof We start by constructing a specific test function. Let ϕ be the solution of the operator equation

$$
\boldsymbol{A}(\omega)\phi = \boldsymbol{B}_i(\omega)\varphi \quad \mathbb{P}\text{-a.e. } \omega \in \Omega
$$

for $\varphi \in L^2(D)$. Then by the assumptions, $\phi \in L^\infty_\mathbb{P}(\Omega; H^1_0(D) \cap H^2(D))$ and, by Assumption 8, $\phi \in C(\Omega; H^1_0(D) \cap H^2(D))$ holds. Using the adjoint state as a test function, we have

$$
\begin{aligned}
\left(\boldsymbol{A}^*(\omega)\lambda_i^{\gamma_l}(\omega), \phi(\omega) \right) &= \left(\lambda_i^{\gamma_l}(\omega), \boldsymbol{A}(\omega)\phi(\omega) \right) \\
&= \left(\lambda_i^{\gamma_l}(\omega), \boldsymbol{B}_i(\omega)\varphi \right) \qquad (4.14) \\
&= \left(\boldsymbol{B}_i(\omega)^* \lambda_i^{\gamma_l}(\omega), \varphi \right)
\end{aligned}
$$

Then due to the Cauchy-Schwarz inequality and Hölder-inequality, respectively, we obtain

$$
\begin{aligned}
\left(\boldsymbol{A}^*(\omega)\lambda_i^\gamma(\omega), \phi(\omega) \right) &= \left(T_i^* T_i(u^\gamma(\omega) + u_f(\omega)) - T_i^* u_d^i, \phi(\omega) \right) + (\mu^\gamma(\omega), \phi(\omega))_{L^2(D)} \\
&= \left(T_i^* T_i(u^\gamma(\omega) + u_f(\omega)) - T_i^* u_d^i, \phi(\omega) \right) \\
&\quad + \langle \mu^\gamma(\omega), \phi(\omega) \rangle_{L^1(D) \times L^\infty(D)} \\
&\leq \| T_i^* T_i(u^\gamma(\omega) + u_f(\omega)) - T_i^* u_d^i \|_{L^2(D)} \|\phi(\omega)\|_{L^2(D)} \\
&\quad + \|\mu^\gamma(\omega)\|_{L^1(D)} \|\phi(\omega)\|_{L^\infty(D)}.
\end{aligned}
$$

Due to the continuous embedding of $H^2(D) \cap H^1_0(D)$ into $L^2(D)$ and $L^\infty(D)$, respectively, we have

$$
\begin{aligned}
&\left(A^*(\omega)\lambda^\gamma_i(\omega), \phi(\omega)\right) \\
&\leq C_1 \|\phi(\omega)\|_{H^2(D) \cap H^1_0(D)} \left(\|T^*_i T_i(u^\gamma(\omega) + u_f(\omega)) - T^*_i u^i_d\|_{L^2(D)} + \|\mu^\gamma(\omega)\|_{L^1(D)}\right)
\end{aligned}
\tag{4.15}
$$

By the assumptions on the operators **A** and **B**, there exists $C_2 \in L^\infty_\mathbb{P}(\Omega)$ such that

$$
\|\phi(\omega)\|_{H^2(D) \cap H^1_0(D)} \leq C_2(\omega)\|\varphi\|_{L^2(D)}.
$$

Now, combining the latter with (4.15) and (4.14), we obtain

$$
\begin{aligned}
\left(\boldsymbol{B}_i(\omega)^* \lambda^\gamma_i(\omega), \varphi\right) &\leq C_1\|\phi(\omega)\|_{H^2(D) \cap H^1_0(D)} \left(\|T^*_i T_i(u^\gamma(\omega) + u_f(\omega))\right. \\
&\quad \left. - T^*_i u^i_d\|_{L^2(D)} + \|\mu^\gamma(\omega)\|_{L^1(D)}\right) \\
&\leq C_1 C_2(\omega)\|\varphi\|_{L^2(D)} \left(\|T^*_i T_i(u^\gamma(\omega) + u_f(\omega))\right. \\
&\quad \left. - T^*_i u^i_d\|_{L^2(D)} + \|\mu^\gamma(\omega)\|_{L^1(D)}\right) \\
&\leq C_1 C_2(\omega)\|\varphi\|_{L^2(D)} \left(C_3(\omega) + C_4(\omega)\right)
\end{aligned}
$$

for all $\varphi \in L^2(D)$. Here, $C_3 \in L^\infty_\mathbb{P}(\Omega)$ and $C_4 \in L^1_\mathbb{P}(\Omega)$. The existence of C_4 follows from the uniform bound on μ^γ in the $L^1_\pi(\Xi)$-norm. Taking the expectation and applying Fubini's theorem yield

$$
\left(\mathbb{E}_\mathbb{P}\left[\boldsymbol{B}^*_i \lambda^\gamma_i\right], \varphi\right) \leq \mathbb{E}_\mathbb{P}\left[C_1 C_2 C_3 + C_1 C_2 C_4\right]\|\varphi\|_{L^2(D)} < \infty
$$

for all $\varphi \in L^2(D)$. In other words, the sequence $\{\mathbb{E}_\mathbb{P}\left[\boldsymbol{B}^*_i \lambda^\gamma_i\right]\}_{\gamma \to \infty}$ is bounded in $L^2(D)$. Thus, there exists a weakly convergent subsequence $\{\mathbb{E}_\mathbb{P}\left[\boldsymbol{B}^*_i \lambda^{\gamma_l}_i\right]\}_{l \in \mathbb{N}}$ and a $\Lambda^*_i \in L^2(D)$ such that $\mathbb{E}_\mathbb{P}\left[\boldsymbol{B}^*_i \lambda^{\gamma_l}_i\right] \rightharpoonup \Lambda^*_i$ in $L^2(D)$. □

Remark 3 The adjoint state plays an important role in numerical methods. In particular, $\mathbb{P}$ is often replaced by an empirical measure $\mathbb{P}_N$, which is associated with an i.i.d. random sample of size N. Therefore the quantity

$$
\frac{1}{N} \sum_{n=1}^N B^*_i(\omega^n) \lambda^{\gamma_l}_i(\omega^n)
$$

is of practical interest. By the (Kolmogorov) strong law of large numbers, we have

$$
\frac{1}{N} \sum_{n=1}^N (B^*_i(\omega^n) \lambda^{\gamma_l}_i(\omega^n), \varphi)_{L^2(D)} \to (E_\mathbb{P}[B^*_i \lambda^{\gamma_l}_i], \varphi)_{L^2(D)}
$$

with probability 1 as $N \to +\infty$ for any $\varphi \in L^2(D)$. For readability, set

$$\overline{X}_{N,l} := \frac{1}{N} \sum_{n=1}^{N} B_i^*(\omega^n) \lambda_i^{\gamma_l}(\omega^n)$$

and recall that almost sure convergence implies convergence in probability. Then for fixed $l \in \mathbb{N}$ and any $\varepsilon > 0$, there exists $N_{l,\varepsilon} \in \mathbb{N}$ such that

$$\mathbb{P}\left(\left|(\overline{X}_{N,l}, \varphi)_{L^2(D)} - (\mathbb{E}_{\mathbb{P}}[B_i^*\lambda_i^{\gamma_l}], \varphi)_{L^2(D)}\right| > \varepsilon\right) < \frac{1}{2^l} \quad \forall N \geq N_{l,\varepsilon}.$$

On the other hand, the previous lemma gives us $\mathbb{E}_{\mathbb{P}}\left[\boldsymbol{B}_i^*\lambda_i^{\gamma_l}\right] \rightharpoonup \Lambda_i^*$ in $L^2(D)$ as $l \to +\infty$. It follows that for any $\varphi \in L^2(D)$ we have

$$\left|(\overline{X}_{N,l}, \varphi)_{L^2(D)} - (\Lambda^*, \varphi)_{L^2(D)}\right|$$
$$\leq \left|(\overline{X}_{N,l}, \varphi)_{L^2(D)} - (E_{\mathbb{P}}[B_i^*\lambda_i^{\gamma_l}], \varphi)_{L^2(D)}\right|$$
$$+ \left|(E_{\mathbb{P}}[B_i^*\lambda_i^{\gamma_l}], \varphi)_{L^2(D)} - (\Lambda^*, \varphi)_{L^2(D)}\right|.$$

This means that the set of all events for which

$$\left|(\overline{X}_{N,l}, \varphi)_{L^2(D)} - (\Lambda^*, \varphi)_{L^2(D)}\right| > 3\varepsilon/2$$

is contained in the set of all events for which

$$\left|(\overline{X}_{N,l}, \varphi)_{L^2(D)} - (E_{\mathbb{P}}[B_i^*\lambda_i^{\gamma_l}], \varphi)_{L^2(D)}\right| > 3\varepsilon/2 - \left|(E_{\mathbb{P}}[B_i^*\lambda_i^{\gamma_l}], \varphi)_{L^2(D)} - (\Lambda^*, \varphi)_{L^2(D)}\right|.$$

Therefore, fix $\varphi \in L^2(D)$ and $\varepsilon > 0$, and choose l such that

$$\left|(E_{\mathbb{P}}[B_i^*\lambda_i^{\gamma_l}], \varphi)_{L^2(D)} - (\Lambda^*, \varphi)_{L^2(D)}\right| < \varepsilon/2,$$

Then for all ε, there exists an l such that

$$\mathbb{P}\left(\left|(\overline{X}_{N_{l,\varepsilon},l}, \varphi)_{L^2(D)} - (\Lambda^*, \varphi)_{L^2(D)}\right| > 3\varepsilon/2\right) < 2^{-l}.$$

Thus, the diagonal sequence of sample averages of the adjoint variables weakly converges in probability to the limiting adjoint variable Λ^*. For a fully discrete scheme using a finite element discretization of the underlying deterministic state spaces, in which error estimates for the deterministic adjoint variables were available, we could derive a similar statement. This is part of the justification for the update heuristic in our algorithm and, in general, for any related numerical algorithm in which the sample sizes gradually increase with the MY-parameters.

Next, we turn our attention on the adjoint equation in the limit.

Lemma 7 *Under the assumptions of Theorem 9, (4.5c) holds.*

Proof As in the previous proof, we start by constructing a specific test function. In this case, let w be the solution of the operator equation

$$\boldsymbol{A}(\omega)w = \boldsymbol{B}_i(\omega)\varphi \quad \mathbb{P}\text{-a.e. } \omega \in \Omega$$

for all $\varphi \in L^2(D)$, then we know that $w \in C(\Omega; H_0^1(D) \cap H^2(D))$ holds. It follows that

$$\left(\boldsymbol{A}^*(\omega)\lambda_i^{\gamma_l}(\omega), w(\omega)\right) = \left(\boldsymbol{B}_i(\omega)^*\lambda_i^{\gamma_l}(\omega), \varphi\right)$$

Taking the expectation on both sides yields

$$\left(\mathbb{E}_{\mathbb{P}}\left[\boldsymbol{B}_i^*\lambda_i^{\gamma_l}\right], \varphi\right) = \mathbb{E}_{\mathbb{P}}\left[\left(T_i^*T_i(u^{\gamma_l} + u_f) - T_i^*u_d^i, w\right)\right] + \mathbb{E}_{\mathbb{P}}\left[(\mu^{\gamma_l}, w)\right]$$

We know that $\mu^{\gamma_l} \rightharpoonup^* \rho^*$ in $\mathcal{M}(\overline{\Xi})$. The right hand side reads as

$$\mathbb{E}_{\mathbb{P}}\left[\left(T_i^*T_i(u^{\gamma_l} + u_f) - T_i^*u_d^i, w\right)\right] + \langle\mu^{\gamma_l}, w\rangle_{\mathcal{M}(\overline{\Xi}),C(\overline{\Xi})}.$$

Passing to the limit $l \to \infty$ yields

$$
\begin{aligned}
(\Lambda_i^*, \varphi) &= \mathbb{E}_{\mathbb{P}}\left[\left(T_i^*T_i(u^* + u_f) - T_i^*u_d^i, w\right)\right] + \int_{\Xi} w(x, \omega)\,\mathrm{d}\rho^*(x, \omega) \\
&= \mathbb{E}_{\mathbb{P}}\left[\left(T_i^*T_i(u^* + u_f) - T_i^*u_d^i, \boldsymbol{A}^{-1}\boldsymbol{B}_i\varphi\right)\right] + \int_{\Xi} \boldsymbol{A}^{-1}(\omega)\boldsymbol{B}_i(\omega)\varphi\,\mathrm{d}\rho^*(x, \omega) \\
&= \mathbb{E}_{\mathbb{P}}\left[\left(\boldsymbol{B}_i^*\boldsymbol{A}^{-*}T_i^*(T_i(u^* + u_f) - u_d^i), \varphi\right)\right] + \int_{\Xi} \boldsymbol{A}^{-1}(\omega)\boldsymbol{B}_i(\omega)\varphi\,\mathrm{d}\rho^*(x, \omega)
\end{aligned}
$$

for all $\varphi \in L^2(D)$. $\qquad\square$

Next, we turn to the sequence $\{\eta^\gamma\}_{\gamma \to \infty} \subset \mathcal{N}_{Z_{\mathrm{ad}}^i}(z_i^\gamma)$.

Lemma 8 *Under the assumptions of Theorem 9, (4.4e) and (4.5d) hold.*

Proof Due (4.2), we can write

$$\eta_i^{\gamma_l} = -\mathbb{E}_{\mathbb{P}}\left[\boldsymbol{B}_i^*\lambda_i^{\gamma_l}\right] - \nu_i z_i^{\gamma_l}.$$

Then the boundedness of the sequence $\{\eta_i^{\gamma_l}\}_{l \in \mathbb{N}}$ in $L^2(D)$ directly follows from

$$
\begin{aligned}
\|\eta_i^{\gamma_l}\|_{L^2(D)} &= \| - \mathbb{E}_{\mathbb{P}}\left[\boldsymbol{B}_i^*\lambda_i^{\gamma_l}\right] - \nu_i z_i^{\gamma_l}\|_{L^2(D)} \\
&\leq \|\mathbb{E}_{\mathbb{P}}\left[\boldsymbol{B}_i^*\lambda_i^{\gamma_l}\right]\|_{L^2(D)} + \|\nu_i z_i^{\gamma_l}\|_{L^2(D)} < \infty.
\end{aligned}
$$

Thus, there exists a $\eta_i^* \in L^2(D)$ and a subsequence $\{\eta_i^{\gamma_{l_n}}\}_{n \in \mathbb{N}}$ such that the assertion holds. $\qquad\square$

Finally, we derive the complementarity system for the multiplier ρ^*.

Lemma 9 *Under the assumptions of Theorem 9, (4.6a) and (4.6b) hold.*

Proof As used several times above, there exists a subsequence of MY parameters $\gamma_l \to +\infty$ along which the multipliers $\{\mu^{\gamma_l}\}$ converge weak* in $\mathcal{M}(\overline{\Xi})$ to some $\rho^* \in \mathcal{M}(\overline{\Xi})$. For each fixed l we have π-a.s.:

$$\mu^{\gamma_l} = -\gamma_l(\psi - (S(\bar{z}_i^{\gamma_l}, \bar{z}_{-i}^{\gamma_l}) + u_f))_+ \leq 0.$$

Therefore, for any non-negative test function $\phi \in C(\overline{\Xi})$, we have

$$\langle \phi, \mu^{\gamma_l} \rangle = \int_\Omega \int_D \phi \mu^{\gamma_l} \mathrm{d}\pi \leq 0.$$

By definition, $\langle \phi, \mu^{\gamma_l} \rangle \to \langle \phi, \rho^* \rangle$ as $l \to +\infty$. Hence, ρ^* is a negatively signed measure. Moreover, setting

$$\phi_l := \psi - (S(\bar{z}_i^{\gamma_l}, \bar{z}_{-i}^{\gamma_l}) + u_f),$$

which is continuous and converges strongly in $C(\overline{\Xi})$ (by assumption) to

$$\phi^* := \psi - (S(\bar{z}_i^*, \bar{z}_{-i}^*) + u_f) \leq 0,$$

we have

$$\langle \phi^*, \rho^* \rangle \geq 0.$$

Furthermore, for each l, we have $\langle \phi_l, \mu^{\gamma_l} \rangle \leq 0$ and $\langle \phi_l, \mu^{\gamma_l} \rangle \to \langle \phi^*, \rho^* \rangle$. Whence we have the complementarity condition. □

This completes the derivation of Theorem 9.

4.4 Probability constraints and Moreau-Yosida regularization

In this final theoretical section, we wish to draw the link between Moreau-Yosida regularization and probability constraints. We do so only for the the risk-neutral PDE-constrained optimization problem (2.14), as the treatment of the GNEP would require further technical assumptions and somewhat obfuscate our main point. The main tools are basic concentration inequalities from probability theory. We recall again the γ-dependent optimization problem:

$$\min_{z \in Z_{\mathrm{ad}}} \left\{ \mathbb{E}_{\mathbb{P}} \left[J(S(z) + u_f, z) + \frac{\gamma}{2} \|(\psi - (S(z) + u_f))_+\|_{L^2(D)}^2 \right] \right\}, \quad (4.16)$$

where $\gamma > 0$. We note that yet another way of formulating the original state constraint is

$$\mathbb{P}\left(\|(\psi - (S(z) + u_f))_+\|_{L^2(D)}^2 \leq 0\right) = 1.$$

Ideally, we would use the $L^\infty(D)$-norm as opposed to the $L^2(D)$-norm, since the latter allows strong violation of the constraint on small subsets of positive measure for the weaker constraint

$$\mathbb{P}\left(\|(\psi - (S(z) + u_f))_+\|_{L^2(D)}^2 \leq \varepsilon\right) = 1,$$

for $\varepsilon > 0$, but arbitrarily small. However, in order to derive a result of the type in the following theorem with the $L^\infty(D)$-norm, we would need a careful analysis similar to [32]. This goes beyond the scope of the current paper.

Proposition 2 *Let z^γ be the unique minimizer of (4.16). Then for any $\varepsilon > 0$, we have*

$$\mathbb{P}\left(\|(\psi - (S(z^\gamma) + u_f))_+\|_{L^2(D)}^2 < \varepsilon\right) \geq 1 - \frac{2\alpha}{\gamma\varepsilon},$$

where $\alpha = \mathbb{E}_{\mathbb{P}}\left[J(S(z), z)\right]$ and z is the unique minimizer of (2.14).

Proof Using Markov's inequality, we have

$$\mathbb{P}\left(\|(\psi - (S(z^\gamma) + u_f))_+\|_{L^2(D)}^2 \geq \varepsilon\right) \leq \frac{\mathbb{E}_{\mathbb{P}}\left[\|(\psi - (S(z^\gamma) + u_f))_+\|_{L^2(D)}^2\right]}{\varepsilon}.$$

We use z^γ to obtain a simpler upper bound. By definition of z^γ, it holds that

$$\mathbb{E}_{\mathbb{P}}\left[J(S(z_\gamma) + u_f, z_\gamma)\right] + \frac{\gamma}{2}\mathbb{E}_{\mathbb{P}}\left[\|(\psi - (S(z_\gamma) + u_f))_+\|_{L^2(D)}^2\right]$$
$$\leq \mathbb{E}_{\mathbb{P}}\left[J(S(v) + u_f, v)\right] + \frac{\gamma}{2}\mathbb{E}_{\mathbb{P}}\left[\|(\psi - (S(v) + u_f))_+\|_{L^2(D)}^2\right].$$

for all $v \in Z_{\mathrm{ad}}$. In particular, we obtain the bound

$$\mathbb{E}_{\mathbb{P}}\left[J(S(z_\gamma) + u_f, z_\gamma)\right] + \frac{\gamma}{2}\mathbb{E}_{\mathbb{P}}\left[\|(\psi - (S(z_\gamma) + u_f))_+\|_{L^2(D)}^2\right] \leq \mathbb{E}_{\mathbb{P}}\left[J(S(v) + u_f, v)\right].$$

for all $v \in Z_{\mathrm{ad}}$ such that $S(v) + u_f \geq \psi$ for $(\mathcal{L} \times \mathbb{P})$-a.e. $(x, \omega) \in D \times \Omega$. Using the minimizer z of (2.14) leads to

$$\frac{\gamma}{2}\mathbb{E}_{\mathbb{P}}\left[\|(\psi - (S(z_\gamma) + u_f))_+\|_{L^2(D)}^2\right] \leq \mathbb{E}_{\mathbb{P}}\left[J(S(z) + u_f, z)\right] - \underbrace{\mathbb{E}_{\mathbb{P}}\left[J(S(z_\gamma) + u_f, z_\gamma)\right]}_{\geq 0}$$

$$\leq \mathbb{E}_{\mathbb{P}}\left[J(S(z) + u_f, z)\right] =: \alpha.$$

From this we obtain $\mathbb{E}_{\mathbb{P}}\left[\|(\psi - (S(z_\gamma) + u_f))_+\|^2_{L^2(D)}\right] \leq \frac{2\alpha}{\gamma}$. Then returning to Markov's inequality, we now have

$$\mathbb{P}\left(\|(\psi - (S(z^\gamma) + u_f))_+\|^2_{L^2(D)} \geq \varepsilon\right) \leq \frac{2\alpha}{\gamma\varepsilon}.$$

Finally, the complementary event is given by

$$\mathbb{P}\left(\|(\psi - (S(z^\gamma) + u_f))_+\|^2_{L^2(D)} < \varepsilon\right) \geq 1 - \frac{2\alpha}{\gamma\varepsilon}. \tag{4.17}$$

$\square$

Remark 4 Using the analysis from the previous sections, we know that there exists a sequence $\gamma_n \to +\infty$ such that the random variable

$$X_n := \|(\psi - (S(z^{\gamma_n}) + u_f))_+\|^2_{L^2(D)}$$

converges strongly in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ to

$$X^* := \|(\psi - (S(z^*) + u_f))_+\|^2_{L^2(D)}.$$

Since z^* is feasible, the state constraint holds and $X^* \equiv 0$. Therefore, there exists a subsequence $\gamma_k := \gamma_{n_k}$ along which $X_k := X_{n_k}$ converges almost surely to 0; and consequently in distribution as well. For each k, we can set $\varepsilon_k = 1/\sqrt{\gamma_k}$ and treat $Y_k := \varepsilon_k$ as a degenerate random variable, which clearly converges in distribution to 0. It follows from Slutsky's theorem that $X_k + Y_k$ converges in distribution to $\|(\psi - (S(z^*) + u_f))_+\|^2_{L^2(D)}$, i.e., 0 and since

$$\mathbb{P}\left(\|(\psi - (S(z^\gamma) + u_f))_+\|^2_{L^2(D)} < \varepsilon\right) \leq \mathbb{P}\left(\|(\psi - (S(z^\gamma) + u_f))_+\|^2_{L^2(D)} \leq \varepsilon\right),$$

the Portmanteau lemma yields

$$\mathbb{P}\left(\|(\psi - (S(z^*) + u_f))_+\|^2_{L^2(D)} \leq 0\right) \geq$$
$$\limsup_{k \to +\infty} \mathbb{P}\left(\|(\psi - (S(z^{\gamma_k}) + u_f))_+\|^2_{L^2(D)} - \varepsilon_k \leq 0\right) \geq \limsup_{k \to \infty} 1 - \frac{2\alpha}{\sqrt{\gamma_k}} = 1.$$

In this sense, Proposition 2 provides us with a probabilistic rate of convergence from Moreau-Yosida to feasibility for the original problem. We observe in the out-of-sample experiments in Sect. 5 almost exactly this behavior, i.e., for $\gamma_k = 1000$, the percent of out-of-sample states is between one and three percent.

5 Numerical experiments

In this final section, we provide a numerical study to indicate how stochastic PDE-constrained optimization problems subject to pointwise state constraints and PDE-constrained GNEPs under uncertainty might best be solved. To the best of our knowledge, this is the first attempt to solve such problems numerically. As a result, the focus will be on the numerical solution of the individual optimization problems. For the GNEP, a Krasnoselskii-Mann-type alternating method is employed in which the dueling agents use the solver from Sect. 5.1.

5.1 Solving the individual problems

The basic idea behind this algorithm derives from the success of semismooth Newton methods for solving deterministic PDE-constrained optimization problems subject to state constraints using Moreau-Yosida regularization and path-following for the parameter updates; see e.g., [30, 31]. Indeed, given $\gamma > 0$ and an iid sample of size M, we can replace the underlying probability distribution with the associated empirical probability measure $\mathbb{P}_M$ and consider

$$
\min_{z \in Z_{ad}} \left\{ \frac{1}{M} \sum_{m=1}^{M} \left[J(S(z)(\omega^m) + u_f(\omega^m), z) \right. \right.
$$
$$
\left. \left. + \frac{\gamma}{2} \|(\psi(\omega^m) - (S(z)(\omega^m) + u_f(\omega^m)))_+\|_{L^2(D)}^2 \right] \right\}. \tag{5.1}
$$

This is now a deterministic problem. In order to solve (5.1) with a semismooth Newton solver, we rewrite the first order optimality system as a single nonsmooth equation in z.

$$
z^{\gamma,M} = \text{Proj}_{Z_{ad}} \left[-\frac{1}{\nu M} \sum_{m=1}^{M} B^*(\omega^m) \lambda_m^\gamma \right], \tag{5.2}
$$

where for each $m = 1, \ldots, M$, $\lambda_m^\gamma \in H_0^1(D)$ solves

$$
\int_D A(x, \omega^m) \nabla \lambda_m^\gamma(x) \cdot \nabla \varphi(x) \, \mathrm{d}x = \int_D (T^*(Tu_m^\gamma + Tu_f(\omega^m) - u_d) + \mu_m^\gamma)) \varphi(x) \, \mathrm{d}x, \tag{5.3}
$$

for all $\varphi \in H_0^1(D)$, $u_m^\gamma = S(z^{\gamma,M})(\omega^m) \in H_0^1(D)$ solves

$$
\int_D A(x, \omega^m) \nabla u_m^\gamma(x) \cdot \nabla \varphi(x) \, \mathrm{d}x = \int_D (B(\omega^m) z^{\gamma,M})(x) \varphi(x) \, \mathrm{d}x, \tag{5.4}
$$

with the same test functions φ, and

$$\mu_m^\gamma = -\gamma(\psi(\omega^m) - (u_m^\gamma + u_f(\omega^m)))_+.$$

The fixed random terms $u_f(\omega^m)$ are defined analogously to u_m^γ. For readability, we denote the mapping $z \mapsto B^*\lambda^\gamma$ as $\Lambda(z)$ or $\Lambda(z, \omega)$ to indicate the dependence on ω. Moreover, we set

$$F_M^\gamma(z) := z - \text{Proj}_{Z_{ad}}\left[-\frac{1}{\nu M}\sum_{m=1}^M \Lambda(z, \omega^m)\right].$$

In the current setting, $F_M^\gamma : L^2(D) \to L^2(D)$ admits a Newton derivative $G_M^\gamma(z)$ of the form

$$G_M^\gamma(z)dz = \left[I + \frac{1}{\nu M}\sum_{m=1}^M \mathcal{G}\left[\Lambda(z, \omega^m)\right]\Lambda'(z, \omega^m)\right]dz,$$

where $\mathcal{G}$ is the Newton derivative of the projection operator. This allows us to apply a semismooth Newton method in $L^2(D)$ [29, 64], which is known to be locally superlinearly convergent for each M and $\gamma > 0$.

However, since γ must be taken to $+\infty$, such an algorithm would not be computationally efficient if M were chosen large for comparatively small γ. If M were to remain fixed, then we could use a strategy as in [1, 30, 31]. On the other hand, M should be ideally as large as possible or also treated as a parameter going to $+\infty$. To remedy this issue, we set a maximum allowable sample size $M_{\max} > 0$ and penalty parameter $\gamma_{\max} > 0$ and, starting with $M_0 \in \mathbb{N}$ and $\gamma_0 > 0$, we add samples to M_k every time γ_k passes a certain threshold. For our numerical experiments, we consider a heuristic, which is motivated by the previous section; in particular the convergence statements in the fully continuous setting along with Remarks 3 and 4. A full convergence analysis linking sampling, approximation and smoothing error goes beyond the scope of this paper. The full algorithm is given in Algorithm 1. A few comments are in order.

The operator $G_{M_k}^{\gamma_k}(z_l^k)$ is not explicitly given. Thus, it is necessary to use an iterative method to solve for the Newton steps dz_l^k, for which we use the tolerance $\text{tol}^{\text{newt}} \geq 0$. Since we are using a semismooth Newton iteration for pointwise bound constraints, the components of dz_l^k are fixed on the estimated active sets for each l and we only need to solve the linear systems on the potentially smaller inactive set. Here, it is important to note that each evaluation of $G_{M_k}^{\gamma_k}(z_l^k)dz_l^k$ requires the solution of the forward equation and two adjoint equations for every sample $m_k = 1, \ldots, M_k$. In our implementation, we employ a preconditioned conjugate gradient method. Therefore, the computational complexity of each Hessian-vector product involved must also be multiplied by M_k and take into account the cost of applying the preconditioner. Similarly, the evaluation of the residual $F_{M_k}^\gamma(z_l^k)$ requires a forward and adjoint solve for each sample. For our numerical examples, we use a direct solver for the linear elliptic PDEs.

Due to these facts, we suggest starting with a relatively small M_0 and increasing slowly with γ_k. Moreover, we suggest a relatively large $\texttt{tol}_0^{\text{res}} > 0$ and ρ^{res} close to 1. In step 13: of Algorithm 1, we simply set $\gamma_{k+1} = \phi(\gamma_k) = \gamma_k + 1$. More aggressive strategies may be possible, but empirical evidence suggests that this is not necessary and may even cause the Newton iteration to cycle. Finally, in step 15: of Algorithm 1, we link the increases of the sample sizes M_k to γ_k. For our implementation, we start with γ_0 and M_0 and increase M_k by 10 every time γ_k is divisible by 100. This is merely a heuristic and other strategies are possible.

Algorithm 1: SSN for Stochastic PDE-Constrained Optimization with State Constraints

1: **Input (Data):** $u_d \in L^2(D)$; $\nu > 0$; $a, b \in L^\infty(D)$ $a < b$; $\psi \in C(\overline{\varXi})$; $f \in L^\infty_{\mathbb{P}}(\Omega; L^2(D))$

2: **Input (Parameters):** $k := 0$, $\gamma_0 > 0$, $\gamma_{\max} \geq \gamma_0 > 0$, $M_0 \in \mathbb{N}$, $M_{\max} \geq M_0$, $\phi : \mathbb{R}_+ \to \mathbb{R}_+$, $\texttt{tol}_0^{\text{res}} > 0$, $\rho^{\text{res}} \in (0, 1)$, $\texttt{tol}^{\text{newt}} \geq 0$

3: **Input (Initial Values):** $z_0 \in L^2(D)$, $dz_0 \equiv 0 \in L^2(D)$

4: **while** $\gamma_k < \gamma_{\max}$ **do**

5:     Set $l := 0$

6:     Set $z_l^k := z_k$

7:     **while** $\|F_{M_k}^{\gamma_k}(z_l^k)\| > \texttt{tol}_k^{\text{res}}$ **do**

8:         Find $dz_l^k \in L^2(D)$ such that

$$\|G_{M_k}^{\gamma_k}(z_l^k)dz_l^k + F_{M_k}^{\gamma_k}(z_l^k)\|_{L^2(D)} \leq \texttt{tol}^{\text{newt}} \tag{5.5}$$

9:         Set $z_{l+1}^k := z_l^k + dz_l^k$

10:        Set $l := l + 1$

11:     **end while**

12:     Set $z_{k+1} = z_l^k$

13:     Set $\gamma_{k+1} = \phi(\gamma_k)$

14:     **if** "penalty-to-sample threshold' **then**

15:         Choose M_{k+1} such that $M_k \leq M_{k+1} \leq M_{\max}$

16:     **else**

17:        Set $M_{k+1} := M_k$

18:     **end if**

19:     Set $\texttt{tol}_{k+1}^{res} = \rho\texttt{tol}_k^{res}$

20:     Set $k := k + 1$

21: **end while**

5.2 Example: risk-neutral PDE-constrained optimization

In order to demonstrate the viability of the algorithm, we consider a model problem based on [43, Ex. 6.1, Ex. 6.2] and [35, Sec. 7.2]. Here, we set $\nu = 10^{-3}$, $D = (0, 1)$,

$\tilde{u}(x) = \sin(50.0 * x/\pi)$, and consider the optimal control problem

$$\underset{z \in L^2(D)}{\text{minimize}} \frac{1}{2} \mathbb{E}_{\mathbb{P}} \left[\|u - \tilde{u}_d\|_{L^2(D)}^2 \right] + \frac{\nu}{2} \|z\|_{L^2(D)}^2 \text{ over } z \in L^2(D) \qquad (5.6)$$

where $z \in Z_{\text{ad}} := \left\{ w \in L^2(D) \,|\, -0.75 \leq w(x) \leq 0.75 \text{ a.e. } x \in D \right\}$ and the solution of the random PDE $u = u(z) \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}; H^1(D))$ solves the weak form of

$$-\nu(\omega)\partial_{xx}u(\omega, x) = f(\omega, x) + z(x) \qquad (\omega, x) \in \Omega \times D, \qquad (5.7a)$$
$$u(\omega, 0) = d_0(\omega), \quad u(\omega, 1) = d_1(\omega) \qquad \omega \in \Omega. \qquad (5.7b)$$

In addition, we impose the state constraint

$$\mathbb{P}(\{\omega \in \Omega \mid u(\omega, x) \geq 0, \text{ for } \mathcal{L}\text{-a.e. } x \in D\}) = 1.$$

Furthermore, we suppose that

$$\nu(\omega) := \max(0.05, \xi_1(\omega)), \qquad f(\omega, x) := \frac{2\xi_2(\omega) - 1}{10}$$
$$d_0(\omega) := 0.75 + \frac{2\xi_3(\omega) - 1}{1000} \qquad d_1(\omega) := 0.5 + \frac{\xi_4(\omega)}{1000},$$

with random variables $\xi_i : \Omega \to \mathbb{R}$, $i = 1, 2, 3, 4$, such that the supports ξ_i, $i = 1, 2, 3, 4$, are $[0, 1]$. We assume here that each of these random variables is uniformly distributed. Following the usual change of variables, the forward problem (5.7) can be understood as

$$-\nu(\xi)\partial_{xx}u(\xi, x) = f(\xi, x) + z(x) \qquad (\xi, x) \in \Xi \times D, \qquad (5.8a)$$
$$u(\xi, 0) = d_0(\xi), \quad u(\xi, 1) = d_1(\xi) \qquad \xi \in \Xi. \qquad (5.8b)$$

with $\Xi = [0, 1]^4$, endowed with the associated uniform density. We define $\boldsymbol{\xi} := (\xi_1, \ldots, \xi_4) \in \Xi$. Since (5.8) is linear, we can use the superposition principle to lift the boundary conditions into the righthand side of (5.8). This allows us to transform the problem into the function space setting used throughout the paper.

Remark 5 (*Feasibility*) In such settings as considered here, the ability to guarantee the nonnegativity of the state with at least one feasible control relies on three factors: 1. whether the maximum principle can be applied (a.s.), 2. whether the random inputs f, d_0, and d_1 give rise to nonnegative solutions when solving the differential equations for each of these terms separately (a.s.), 3. the width of the bilateral bounds on z. If, for instance d_0, d_1 are nonnegative, but f is bounded and negative on some portions of D with positive probability, then with sufficiently wide bounds on the control z, the righthand side can be made positive almost surely. This is admittedly not ideal. It is, however, much weaker than standard assumptions in two-stage stochastic programming such as complete or relatively complete recourse, which would require such a property for all feasible z.

5.2.1 Discretization and implementation

The pointwise forward problem and the control space are discretized using piecewise finite elements on a uniform mesh with parameter $h = 1/(2^8 - 1)$. We use a standard Monte Carlo approximation for the random inputs $\xi_1, \ldots, \xi_4 \in [0, 1]$. We initialize the algorithm by choosing: $\gamma_0 = 1$, $\gamma_{\max} = 10^4$, $M_0 = 200$, $\mathtt{tol}_0^{\mathrm{res}} = 10^{-2}$, $\rho^{\mathrm{res}} = 0.9997$, $\mathtt{tol}^{\mathrm{newt}} = 10^{-8}$, $z_0 \equiv 0$. Once the penalty to sample threshold is reached, M_k is increased by 10 samples. As mentioned above, the discrete PDEs are solved via a direct solver and the Newton steps are calculated using a preconditioned conjugate gradients method (for the linear equation on the inactive set). As a preconditioner we use the localized mass matrix for the inactive set. In the current implementation, we use the ℓ^2-norm of the residual in the stopping criterion. Alternatively, one could use the proper discrete Riesz maps (i.e. the inverse mass matrix) to first obtain a representation of the discrete solution in the finite-dimensional subspace and then use the discrete L^2-norm. This would be especially important in a nested grid or AFEM approach in future numerical studies. For the nonsmooth operator in the adjoint equations, we utilize a mass-lumping approach to obtain the discrete operators. Otherwise, the differential operators and identity operators give rise to the usual stiffness and mass matrices subject to the random inputs. We have observed that a heuristic damping strategy in which we take $z_{k+1} = z_k + t_{\mathrm{damp}}d_k$ with $t_{\mathrm{damp}} \in (0, 1)$, where z_k is the current iteration and d_k is the full Newton step, adds significant robustness to the method. For our experiments, we set $t_{\mathrm{damp}} = 0.1$. Though this has a clear effect on the local rate of convergence, the wildly varying ξ-dependent states appear to make this necessary.

5.2.2 Performance of the algorithm

The performance of the algorithm can be seen in Fig. 1, where we plot the total number of PCG iterations per γ-update (k) and the total number of Newton iterations needed to reach $\gamma = 1000$. The number of PCG iterations remains relatively stable (between 200 and 400), whereas the number of Newton iterations per γ-update appears to be trending downward. As mentioned above, robust convergence of the inner Newton iteration was ensured using a heuristic damping step with factor 0.1. Nevertheless, the number of iterations trends downwards as γ increases. We also note that both the CG algorithm and inner loop used inexact solves throughout. Since we employed a relatively rough initial stopping tolerance and a small batch of samples (despite increasing by 10 everytime ($\gamma_k \mod 100) = 0$) the algorithm consistently produces a solution $\bar{z}$ that performs exactly as expected in light of the model (risk neutral objective) and theory (especially Remark 4). This is qualitatively illustrated in Fig. 2, where we observed that only 0.6% of the out-of-sample states violated the bound constraint. Due to the presence of the random viscosity term in the forward problem, the L^∞-norm of the sampled states can vary significantly. Finally, and perhaps due to the previous fact, we noticed that smaller batches sizes, e.g., on the order of 10, led to a failure of the Newton solver for γ near 1000.
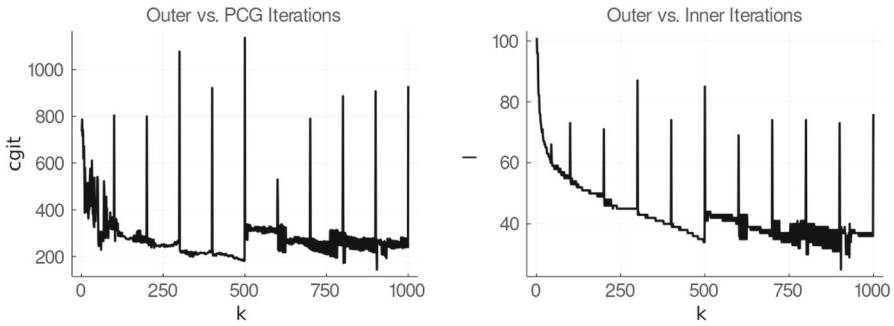
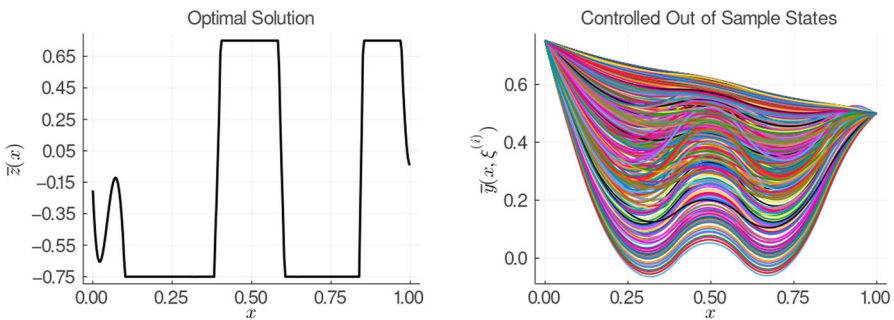**Fig. 1** (left) Total PCG iterations per outer iteration k. (right) Total Newton iterations per outer iteration k

Fig. 2 (left) Optimal solution $\bar{z}$ up to $\gamma = 1000$. (right) Controlled states using $\bar{z}$ for 2000 out of sample instances of ξ

5.3 Solving the risk-netural PDE-constrained GNEP

5.3.1 A general algorithm

As mentioned earlier, we employ a fixed point strategy to solve a two-player, risk-netural PDE-constrained GNEP. The fixed point iteration is derived from a standard Krasnoselskii-Mann iteration. We introduce the mappings $T^i(z_j)$ $i \neq j$, where

$$T^1(z_2) := \underset{z_1 \in Z_{\mathrm{ad}}^1}{\operatorname{argmin}} \left\{ \mathbb{E}_{\mathbb{P}}\left[J_1(S(z_1, z_2) + u_f, (z_1, z_2)) \right] | \mathbb{P}(S(z_1, z_2) + u_f \geq \psi) = 1 \right\}.$$

and $T^2(z_1)$ is defined analogously. The fixed point iteration is based on the following outer iteration:

1. Given $(z_1^{\mathrm{old}}, z_2^{\mathrm{old}}) \in L^2(D) \times L^2(D)$.
2. The first player determines $\widehat{z}_1 = T^1(z_2^{\mathrm{old}})$ and reveals this to the second player.
3. The second player then determines $\widehat{z}_2 = T^2(\widehat{z}_1)$ and reveals this to the first player.
4. The first player now determines $w_1 = T^1(\widehat{z}_2)$.
5. Choosing $\lambda \in (0, 1]$, the first player now updates their strategy by setting $z_1^{\mathrm{new}} := (1 - \lambda)z_1^{\mathrm{old}} + \lambda w_1$. The second player is assumed to choose $z_2^{\mathrm{new}} = T^2(z_1^{\mathrm{new}})$.

Obviously, (1)–(5) represents an ideal setting as the state constraint needs to be treated by a Moreau-Yosida approximation. In this context, we denote the γ-dependent mapping in steps (2)–(4) by T_γ^i. The full algorithm is depicted in Algorithm 2. Note that $\lambda = 1$ would correspond to a Gauss-Seidel iteration and $\lambda > 1$ to successive over-relation. As the evaluation of the T-mappings requires an iterative solver in practice, we suggest to initialize the nonlinear solvers by using z_1^{old} is used in (2), z_2^{old} in (3), $\widehat{z}_1$ in (4), and $\widehat{z}_2$ in (5).

Many of the inputs in Algorithm 2 are either self-explanatory or play the same role as in Algorithm 1. Here, we introduce $\texttt{tol}_0^{\text{km}} > 0$ and $\rho^{\text{km}} \in (0, 1]$, which allow us to successively reduce the tolerance used in the Krasnoselkskii-Mann iteration as γ_k (and consequently M_k) increase. We suppress the fact that certain fixed data and parameter values need to be passed to the T_γ-operators throughout the inner iterations.

It is again possible to adapt the tolerance used in the PCG solver for the Newton steps, but empirical evidence indicates that this value should be rather small (order at least 1e-6). Though the structure of Algorithm 2 is very similar to that of Algorithm 1, it is important to note that each evaluation of $T_{\gamma_k}^i$ is associated with a semismooth Newton solve for the current γ_k and sample of size M_k.

Remark 6 For each fixed γ_k and M_k, the algorithm is basically a Krasnoselskii-Mann iteration with inexact evaluations of the fixed point mapping. As such, convergence can be guaranteed if the latter can be shown to be nonexpansive. Such an analysis goes beyond the scope of the paper. Given the underlying individual problems are strongly convex, this property is most likely linked to the modulus of strong convexity of the individual cost functions.

Remark 7 Gauss-Seidel iterations (using $\lambda = 1$ above) have been considered in the context of GNEPs and NEPs in a number of texts, e.g., [33, 34], for more than two players. A similar adaptation using the Krasnoselskii-Mann-type approach for more than two players is therefore conceivable.

5.3.2 Examples

We wish to study the performance of the algorithm for two example GNEPs. The individual problems build on the model class used in Sect. 5.2. We start with an essentially symmetric game in which only the desired states $u_{d,i}$ differ and the controls z_1, z_2 are fully amenable in the sense that $B_1 = B_2$ is the identity on $L^2(D)$. As a second example we set

$$B_1 z_1 = \chi_{[0, \frac{1}{2})} z_1 \text{ and } B_2 z_2 = \chi_{[\frac{1}{2}, 1)} z_2.$$

In both examples, we set $u_{d,1} = \sin(50x/\pi)$ and $u_{d,2} = \cos(50x/\pi)$. In the case where B_1, B_2 are trivial, we take the bounds to be $a_1 = a_2 = -1$ and $b_1 = b_2 = 1$, whereas in the second example with restricted control action, we set $a_1 = a_2 = -3/4$ and $b_1 = b_2 = 3/4$. These choices were made to try to force larger active sets, which ensure that we are solving a truly nonsmooth equation.

Algorithm 2: A Fixed Point Iteration for a Stochastic PDE-Constrained GNEP

1: **Input (Data):** $(u_{d,1}, u_{d,2}) \in L^2(D)^2$, $(v_1, v_2) \in \mathbb{R}_{++}^2$, $(a_i, b_i) \in L^\infty(D)^2$ $a_i < b_i$ $i = 1, 2$,
  $\psi \in C(\overline{\Xi})$, $f \in L_{\mathbb{P}}^\infty(\Omega; L^2(D))$
2: **Input (Parameters):** $k := 0$, $\lambda \in (0, 1]$, $\gamma_0 > 0$, $\gamma_{\max} \geq \gamma_0 > 0$, $M_0 \in \mathbb{N}$, $M_{\max} \geq M_0$,
  $\phi : \mathbb{R}_+ \to \mathbb{R}_+$, $\texttt{tol}_0^{km} > 0$, $\rho^{km} \in (0, 1]$, $\texttt{tol}_0^{nwt} > 0$, $\rho^{nwt} \in (0, 1]$, $\texttt{tol}^{newt} \geq 0$
3: **Input (Initial Values):** $(z_{0,1}, z_{0,2}) \in L^2(D)^2$, $(dz_{0,1}, dz_{0,2}) \equiv 0 \in L^2(D)^2$
4: **(Initialize):** Set
$$\widehat{z}_{1,1} = T_{\gamma_k}^1(z_{0,2}), \quad \widehat{z}_{1,2} = T_{\gamma_k}^2(\widehat{z}_{1,1}), \quad w_{1,1} = T_{\gamma_k}^1(\widehat{z}_{1,2})$$

5: Set $z_{1,1} = (1 - \lambda)z_{0,1} + \lambda w_{1,1}$, $z_{1,2} = T_{\gamma_k}^2(z_{1,1})$, and $k := k + 1$
6: **while** $\gamma_k \leq \gamma_{\max}$ **do**
7:   Set $l := 0$
8:   Set $z_{l,1}^k := z_{k,1}$ and $z_{l,2}^k := z_{k,2}$
9:   Set
$$\widehat{z}_{l+1,1}^k = T_{\gamma_k}^1(z_{l,2}^k) \quad \widehat{z}_{l+1,2}^k = T_{\gamma_k}^2(\widehat{z}_{l+1,1}^k) \quad w_{l+1,1}^k = T_{\gamma_k}^1(\widehat{z}_{l+1,2}^k)$$
  Set $z_{l+1,1}^k = (1 - \lambda)z_{l,1}^k + \lambda w_{l+1,1}^k$, $z_{l+1,2}^k = T_{\gamma_k}^2(z_{l+1,1}^k)$, and $l := l + 1$
10:   **while** $\|z_{l,1}^k - z_{l-1,1}^k\|_{L^2(D)} > \texttt{tol}_k^{km}$ **do**
11:     Set
$$\widehat{z}_{l+1,1}^k = T_{\gamma_k}^1(z_{l,2}^k) \quad \widehat{z}_{l+1,2}^k = T_{\gamma_k}^2(\widehat{z}_{l+1,1}^k) \quad w_{l+1,1}^k = T_{\gamma_k}^1(\widehat{z}_{l+1,2}^k)$$
    Set $z_{l+1,1}^k = (1 - \lambda)z_{l,1}^k + \lambda w_{l+1,1}^k$, $z_{l+1,2}^k = T_{\gamma_k}^2(z_{l+1,1}^k)$, and $l := l + 1$
12:   **end while**
13:   Set $z_{k+1,1} = z_{l,1}^k$ and $z_{k+1,2} = z_{l,2}^k$
14:   Set $\gamma_{k+1} = \phi(\gamma_k)$
15:   **if** "penalty-to-sample threshold' **then**
16:     Choose $M_{k+1} \geq \min(M_k, M_{\max})$
17:   **else**
18:     Set $M_{k+1} := M_k$
19:   **end if**
20:   Set $\texttt{tol}_{k+1}^{km} = \rho^{km} \texttt{tol}_k^{km}$
21:   Set $k := k + 1$
22: **end while**

Given these assumptions, the template for the individual problems has the form

$$\underset{z \in L^2(D)}{\text{minimize}} \frac{1}{2} \mathbb{E}_{\mathbb{P}} \left[\|u - \widetilde{u}_{d,i}\|_{L^2(D)}^2 \right] + \frac{v_i}{2} \|z_i\|_{L^2(D)}^2 \text{ over } z_i \in L^2(D) \qquad (5.9)$$

where $z \in Z_{\text{ad}}^i := \left\{ w \in L^2(D) \,|\, a_i \leq w(x) \leq b_i \text{ a.e. } x \in D \right\}$ and the solution of the random PDE $u = u(z) \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}; H^1(D))$ solves the weak form of

$$-v(\omega)\partial_{xx}u(\omega, x) = f(\omega, x) + (B_1 z_1)(x) + (B_2 z_2)(x) \qquad (\omega, x) \in \Omega \times D, \tag{5.10a}$$

$$u(\omega, 0) = d_0(\omega), \quad u(\omega, 1) = d_1(\omega) \qquad \omega \in \Omega. \tag{5.10b}$$

As before, we impose the state constraint

$$\mathbb{P}(\{\omega \in \Omega \,|\, u(\omega, x) \geq 0, \text{ for } \mathcal{L}\text{-a.e. } x \in D\}) = 1,$$
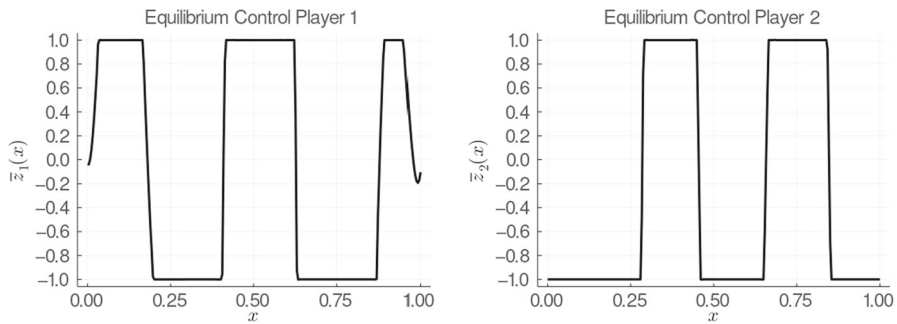
Fig. 3 Equilibrium controls $\bar{z}_1$ (left) and $\bar{z}_2$ (right) for $B_1 = B_2 = \mathrm{id}_{L^2(D)}$

where we set $D = (0, 1)$. The uncertainties are chosen as in Sect. 5.2.

5.3.3 Discretization and implementation

The discretization, sampling, γ-update strategy, and tolerance reduction for the Newton iterations are the same as in Sect. 5.2.1. We fixed $\lambda = 0.5$. Though further experiments demonstrated that successive over relaxation, i.e., $\lambda > 1$, does in fact work, the number of Krasnoselskii-Mann iterations remained roughly the same. The inner KM-iterations stopped once the discrete $L^2(D)$-norm of $z_1^{\mathrm{new}} - z_1^{\mathrm{old}}$ reached a tolerance of 1e-3.

5.3.4 Performance of the Algorithm

We have already investigated the performance of the sample-average based semismooth Newton solver in Sect. 5.2.2. As expected, the algorithm performs reliably in the GNEP setting, where it is called hundreds of times without failing to converge. The behavior of Algorithm 2 for the full control action is depicted in Figs. 3 and 4. For the second example, in which the controls are restricted to subsets of D, we point the reader to Figs. 5 and 6. In both cases, we observe non-trivial active sets for the equilibrium controls. In either case, the fixed point iteration requires a moderate number of iterations for the first few γ. This then rapidly tapers off as γ_k and M_k increased. The performance of the equilibrium controls is also demonstrated in Figs. 4 and 6. These plots correspond to an estimated violation (in the sense of the L^∞-norm) of the state constraints of 1.5% and 0.6% respectively. This is well within the usual tolerance of 95% often used for probability constraints.

6 Conclusions and outlook

In this paper, we proved existence of solutions/equilibria and derive optimality conditions for both stochastic PDE-constrained optimization and equilibrium problems subject to state constraints. For our analysis, higher regularity of the random states was proven using a priori estimates for deterministic elliptic PDE. This allowed us
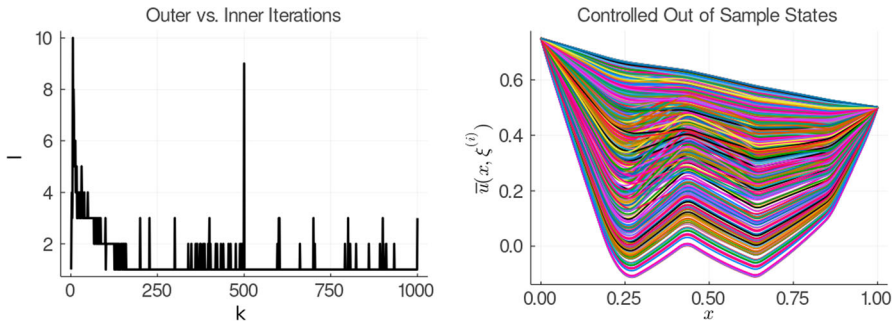
Fig. 4 (left) Number of Krasnoselskii-Mann iterations versus γ-updates k. (right) Controlled states using Nash equilibrium $(\bar{z}_1, \bar{z}_2)$ for 2000 out of sample instances of $\boldsymbol{\xi}$. $B_1 = B_2 = \mathrm{id}_{L^2(D)}$
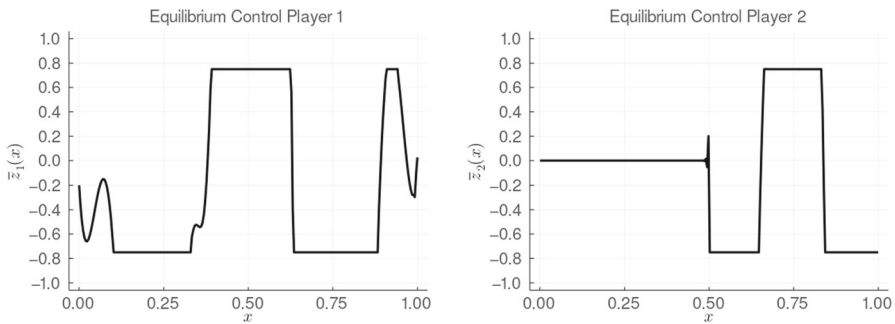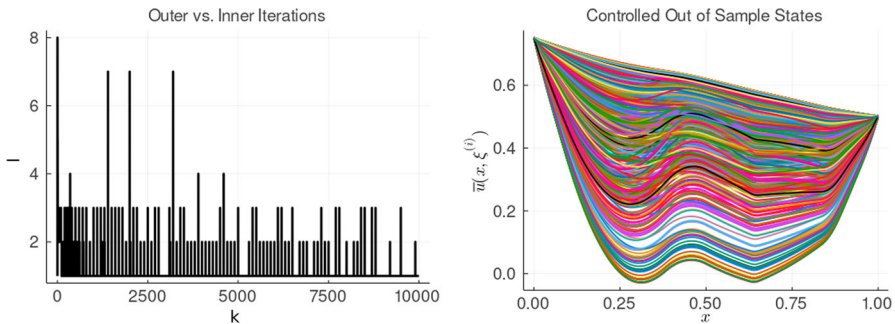

Fig. 5 Equilibrium controls $\bar{z}_1$ (left) and $\bar{z}_2$ (right) for $B_1 = \chi_{[0,0.5]}$, $B_2 = \chi_{[0.5,1]}$

Fig. 6 (left) Number of Krasnoselskii-Mann iterations versus γ-updates k. (right) Controlled states using Nash equilibrium $(\bar{z}_1, \bar{z}_2)$ for 2000 out of sample instances of $\boldsymbol{\xi}$ for $B_1 = \chi_{[0,0.5]}$, $B_2 = \chi_{[0.5,1]}$

to make use of the existing optimality theory for convex optimization problems. In the case of GNEPs, a GNEP-specific constraint qualification was crucial for the development of a relaxation approach on which both the theory and our numerical methods could be built. We saw that this condition is fundamentally different than the classical constraint qualifications from nonlinear programming such as the Mangasarian-Fromowitz CQ, which was originally introduced in [46]. Nevertheless,

the low regularity of the Lagrange multipliers still makes passing to the limit highly nontrivial.

After rigorously passing to the limit in the smoothing parameter, we provided further insight into the approximation technique using results on concentration inequalities and asymptotic statistics. Finally, we suggested two algorithms; the first for solving risk-neutral PDE-constrained optimization problems subject to state constraints and the second for the extension to GNEPs.

The algorithms performed well and the statistical properties of the solutions are comparable with what one would require of probability constraints; though our approach is much easier to treat theoretically and numerically. At least for a fixed sample, the optimization solver is known to converge locally superlinearly. A full convergence analysis linking sampling, adaptive finite elements, smoothing, and convergence of these algorithms (as least in a probabilistic sense) will be a future direction of research. The convergence of the GNEP solver is much more delicate and will require a fine analysis of the nonexpansivity of the underlying fixed point mapping. We postulate here that this is linked to the modulus of strong convexity of the underlying problems.

References

1. Adam, L., Hintermüller, M., Surowiec, T.M.: A semismooth Newton method with analytical path-following for the H^1-projection onto the Gibbs simplex. IMA J. Numer. Anal. **39**(3), 1276–1295 (2019). https://doi.org/10.1093/imanum/dry034
2. Adams, R.A., Fournier, J.J.F.: Sobolev spaces. In: Pure and Applied Mathematics Amsterdam, vol. 140, 2nd edn. Elsevier/Academic Press, Amsterdam (2003)
3. Alexanderian, A., Petra, N., Stadler, G., Ghattas, O.: Mean-variance risk-averse optimal control of systems governed by PDEs with random parameter fields using quadratic approximations. arXiv preprint arXiv:1602.07592 (2016)
4. Aliprantis, C.D., Border, K.C.: Infinite Dimensional Analysis, 3rd edn. Springer, Berlin (2006)
5. Alt, H.W.: Linear functional analysis. Universitext. Springer-Verlag, London Ltd, London: An application-oriented introduction. Translated from the German edition by Robert Nürnberg (2016). https://doi.org/10.1007/978-1-4471-7280-2
6. Arrow, K.J., Debreu, G.: Existence of an equilibrium for a competitive economy. Econometrica **22**, 265–290 (1954). https://doi.org/10.2307/1907353
7. Attouch, H., Buttazzo, G., Michaille, G.: Variational analysis in Sobolev and BV spaces. In: MPS/SIAM Series on Optimization, vol. 6. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA (2006). Applications to PDEs and optimization
8. Aubin, J.P.: Optima and Equilibria: An Introduction to Nonlinear Analysis. Graduate texts in mathematics, 2nd edn. Springer-Verlag, Berlin (1998). Translated from the French by Stephen Wilson

9. Bonfiglio, L., Perdikaris, P., Brizzolara, S., Karniadakis, G.: A multi-fidelity framework for investigating the performance of super-cavitating hydrofoils under uncertain flow conditions. In: 19th AIAA Non-Deterministic Approaches Conference, pp. 1–13. American Institute of Aeronautics and Astronautics (2017). https://doi.org/10.2514/6.2017-1328

10. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer, New York (2000)

11. Borzì, A., Schulz, V., Schillings, C., von Winckel, G.: On the treatment of distributed uncertainties in PDE constrained optimization. GAMM Mitt. **33**(2), 230–246 (2010)

12. Borzì, A., von Winckel, G.: A POD framework to determine robust controls in PDE optimization. Comput. Vis. Sci. **14**, 91–103 (2011)

13. Chen, P., Quarteroni, A.: Weighted reduced basis method for stochastic optimal control problems with elliptic PDE constraint. SIAM/ASA J. Uncertain. Quantif. **2**(1), 364–396 (2014)

14. Chen, P., Quarteroni, A., Rozza, G.: Stochastic optimal robin boundary control problems of advection-dominated elliptic equations. SIAM J. Numer. Anal. **51**(5), 2700–2722 (2013)

15. Debreu, G.: A social equilibrium existence theorem. Proc. Nat. Acad. Sci. USA **38**, 886–893 (1952). https://doi.org/10.1073/pnas.38.10.886

16. Désidéri, J.A.: Cooperation and competition in multidisciplinary optimization: application to the aero-structural aircraft wing shape optimization. Comput. Optim. Appl. **52**(1), 29–68 (2012). https://doi.org/10.1007/s10589-011-9395-1

17. Dunford, N., Schwartz, J.T.: Linear operators. Part I. Wiley Classics Library. John Wiley & Sons, Inc., New York (1988). General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication

18. Durlofsky, L.J., Chen, Y.: Uncertainty Quantification for Subsurface Flow Problems Using Coarse-Scale Models, chap 6, pp. 163–202. Springer, Berlin Heidelberg, Berlin, Heidelberg (2012)

19. Farshbaf-Shaker, M.H., Henrion, R., Hömberg, D.: Properties of chance constraints in infinite dimensions with an application to PDE constrained optimization. Set-Valued Var. Anal. **26**(4), 821–841 (2018). https://doi.org/10.1007/s11228-017-0452-5

20. Geiersbach, C., Pflug, G.C.: Projected stochastic gradients for convex constrained problems in Hilbert spaces. SIAM J. Optim. **29**(3), 2079–2099 (2019). https://doi.org/10.1137/18M1200208

21. Geiersbach, C., Wollner, W.: Optimality conditions for convex stochastic optimization problems in Banach spaces with almost sure state constraint (2020). WIAS Preprint No. 2755, https://doi.org/10.20347/WIAS.PREPRINT.2755

22. Geiersbach, C., Wollner, W.: A stochastic gradient method with mesh refinement for PDE-constrained optimization under uncertainty. SIAM J. Sci. Comput. **42**(5), A2750–A2772 (2020). https://doi.org/10.1137/19M1263297

23. Geletu, A., Hoffmann, A., Schmidt, P., Li, P.: Chance constrained optimization of elliptic PDE systems with a smoothing convex approximation. https://doi.org/10.1051/cocv/2019077. To appear in ESAIM: COCV, https://doi.org/10.1051/cocv/2019061

24. Graham, I.G., Kuo, F.Y., Nichols, J.A., Scheichl, R., Schwab, C., Sloan, I.H.: Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients. Numer. Math. **131**(2), 329–368 (2015). https://doi.org/10.1007/s00211-014-0689-y

25. Grisvard, P.: Elliptic problems in nonsmooth domains. Reprint of the 1985 hardback ed., vol. 69, reprint of the 1985 hardback ed. edn. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM) (2011)

26. Habbal, A., Petersson, J., Thellner, M.: Multidisciplinary topology optimization solved as a Nash game. Int. J. Numer. Methods Eng. **61**(7), 949–963 (2004). https://doi.org/10.1002/nme.1093

27. Hewitt, E., Stromberg, K.: Real and abstract analysis. Springer-Verlag, New York-Heidelberg (1975). A modern treatment of the theory of functions of a real variable, Third printing, Graduate Texts in Mathematics, No. 25

28. Hille, E., Phillips, R.S.: Functional analysis and semi-groups. American Mathematical Society, Providence, R. I. (1974). Third printing of the revised edition of 1957, American Mathematical Society Colloquium Publications, Vol. XXXI

29. Hintermüller, M., Ito, K., Kunisch, K.: The primal-dual active set strategy as a semismooth Newton method. SIAM J Optim. **13**(3), 865–888 (2002). https://doi.org/10.1137/S1052623401383558. (**2003**)

30. Hintermüller, M., Kunisch, K.: Feasible and noninterior path-following in constrained minimization with low multiplier regularity. SIAM J. Control Optim. **45**(4), 1198–1221 (2006). https://doi.org/10.1137/050637480

31. Hintermüller, M., Kunisch, K.: Path-following methods for a class of constrained minimization problems in function space. SIAM J. Optim. **17**(1), 159–187 (2006). https://doi.org/10.1137/040611598
32. Hintermüller, M., Schiela, A., Wollner, W.: The length of the primal-dual path in Moreau-Yosida-based path-following methods for state constrained optimal control. SIAM J. Optim. **24**(1), 108–126 (2014). https://doi.org/10.1137/120866762
33. Hintermüller, M., Surowiec, T.: A PDE-constrained generalized Nash equilibrium problem with pointwise control and state constraints. Pac. J. Optim. **9**(2), 251–273 (2013)
34. Hintermüller, M., Surowiec, T., Kämmler, A.: Generalized Nash equilibrium problems in Banach spaces:theory, Nikaido-Isoda-based path-following methods, and applications. SIAM J. Optim. **25**(3), 1826–1856 (2015). https://doi.org/10.1137/14096829X
35. Hoffhues, M., Römisch, W., Surowiec, T.M.: On quantitative stability in infinite-dimensional optimization under uncertainty (2020). Preprint SPP1962-131
36. Hytönen, T., van Neerven, J., Veraar, M., Weis, L.: Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory, *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*, vol. 63. Springer, Cham (2016)
37. Kanzow, C., Karl, V., Steck, D., Wachsmuth, D.: The multiplier-penalty method for generalized Nash equilibrium problems in Banach spaces. SIAM J. Optim. **29**(1), 767–793 (2019). https://doi.org/10.1137/17M114114X
38. Karl, V., Pörner, F.: On the uniqueness of non-reducible multi-player control problems. Optim. Methods Softw. (2019). https://doi.org/10.1080/10556788.2019.1694021
39. Kouri, D.P.: A multilevel stochastic collocation algorithm for optimization of PDEs with uncertain coefficients. SIAM/ASA J. Uncertain. Quantif. **2**(1), 55–81 (2014)
40. Kouri, D.P., Heinkenschloss, M., Ridzal, D., van Bloemen Waanders, B.G.: A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty. SIAM J. Sci. Comput. **35**(4), A1847–A1879 (2013)
41. Kouri, D.P., Heinkenschloss, M., Ridzal, D., van Bloemen Waanders, B.G.: Inexact objective function evaluations in a trust-region algorithm for PDE-constrained optimization under uncertainty. SIAM J. Sci. Comput. **36**(6), A3011–A3029 (2014)
42. Kouri, D.P., Surowiec, T.M.: Risk-averse optimal control of semilinear elliptic PDEs. https://doi.org/10.1051/cocv/2019061. To appear in ESAIM: COCV
43. Kouri, D.P., Surowiec, T.M.: Risk-averse PDE-constrained optimization using the conditional value-at-risk. SIAM J. Optim. **26**(1), 365–396 (2016)
44. Kouri, D.P., Surowiec, T.M.: Existence and optimality conditions for risk-averse pde-constrained optimization. SIAM/ASA J. Uncertain. Quantif. **6**(2), 787–815 (2018). https://doi.org/10.1137/16M1086613
45. Kouri, D.P., Surowiec, T.M.: Epi-regularization of risk measures. Math. Oper. Res. **45**(2), 774–795 (2020). https://doi.org/10.1287/moor.2019.1013
46. Mangasarian, O., Fromovitz, S.: The fritz john necessary optimality conditions in the presence of equality and inequality constraints. J. Math. Anal. Appl. **17**(1), 37–47 (1967). https://doi.org/10.1016/0022-247X(67)90163-1
47. Nash, J.F.: Noncooperative games. Ann. Math. **2**(54), 286–295 (1951). https://doi.org/10.2307/1969529
48. Osborne, M.J., Rubinstein, A.: A Course in Game Theory. MIT Press, Cambridge, MA (1994)
49. Petra, N., Zhu, H., Stadler, G., Hughes, T.J.R., Ghattas, O.: An inexact Gauss-Newton method for inversion of basal sliding and rheology parameters in a nonlinear Stokes ice sheet model. J. Glaciol. **58**(211) (2012)
50. Prékopa, A.: Stochastic programming. In: Mathematics and its Applications, vol. 324. Kluwer Academic Publishers Group, Dordrecht (1995)
51. Ramos, Á.M.: Numerical methods for Nash equilibria in multiobjective control of partial differential equations. In: Analysis and Optimization of Differential Systems (Constanta, 2002), pp. 333–344. Kluwer Acad. Publ, Boston, MA (2003)
52. Ramos, A.M., Glowinski, R., Periaux, J.: Nash equilibria for the multiobjective control of linear partial differential equations. J. Optim. Theory Appl. **112**(3), 457–498 (2002). https://doi.org/10.1023/A:1017981514093

53. Ramos, Á.M., Glowinski, R., Periaux, J.: Pointwise control of the Burgers equation and related Nash equilibrium problems: computational approach. J. Optim. Theory Appl. **112**(3), 499–516 (2002). https://doi.org/10.1023/A:1017907930931
54. Ramos, Á.M., Roubíček, T.: Nash equilibria in noncooperative predator-prey games. Appl. Math. Optim. **56**(2), 211–241 (2007). https://doi.org/10.1007/s00245-007-0894-5
55. Rockafellar, R.T., Uryasev, S.: The fundamental risk quadrangle in risk management, optimization and statistical estimation. Surv. Oper. Res. Manag. Sci. **18**(1–2), 33–53 (2013). https://doi.org/10.1016/j.sorms.2013.03.001
56. Rosen, J.B.: Existence and uniqueness of equilibrium points for concave n-person games. Econometrica **33**, 520–534 (1965). https://doi.org/10.2307/1911749
57. Rosseel, E., Wells, G.N.: Optimal control with stochastic PDE constraints and uncertain controls. Comput. Methods Appl. Mech. Eng. **213**, 152–167 (2012)
58. Schulz, V., Schillings, C.: On the nature and treatment of uncertainties in aerodynamic design. AIAA J. **47**(3), 646–654 (2009)
59. Shapiro, A., Dentcheva, D., Ruszczyński, A.: Lectures on stochastic programming, *MOS-SIAM Series on Optimization*, vol. 9, 2nd edn. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA (2014). Modeling and theory
60. Stadler, G., Gurnis, M., Burstedde, C., Wilcox, L.C., Alisic, L., Ghattas, O.: The dynamics of plate tectonics and mantle flow: from local to global scales. Science **329**(5995), 1033–1038 (2010)
61. Tang, Z., Désidéri, J.A., Périaux, J.: Multicriterion aerodynamic shape design optimization and inverse problems using control theory and Nash games. J. Optim. Theory Appl. **135**(3), 599–622 (2007). https://doi.org/10.1007/s10957-007-9255-4
62. Tartakovsky, D.M., Guadagnini, A., Riva, M.: Stochastic averaging of nonlinear flows in heterogeneous porous media. J. Fluid Mech. **492**, 47–62 (2003)
63. Tiesler, H., Kirby, R.M., Xiu, D., Preusser, T.: Stochastic collocation for optimal control problems with stochastic PDE constraints. SIAM J. Control Optim. **50**(5), 2659–2682 (2012)
64. Ulbrich, M.: Semismooth Newton methods for operator equations in function spaces. SIAM J. Optim. **13**(3), 805–842 (2002). https://doi.org/10.1137/S1052623400371569. (**(2003)**)
65. Werner, D.: Funktionalanalysis, extended Springer-Verlag, Berlin (2000)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.