Numerical Solution of an Optimal Control Problem with Probabilistic and Almost Sure State Constraints

Caroline Geiersbach[1] · René Henrion[1] · Pedro Pérez-Aros[2]

Abstract

We consider the optimal control of a PDE with random source term subject to probabilistic or almost sure state constraints. In the main theoretical result, we provide an exact formula for the Clarke subdifferential of the probability function without a restrictive assumption made in an earlier paper. The focus of the paper is on numerical solution algorithms. As for probabilistic constraints, we apply the method of spherical radial decomposition. Almost sure constraints are dealt with a Moreau–Yosida smoothing of the constraint function accompanied by Monte Carlo sampling of the given distribution or its support or even just the boundary of its support. Moreover, one can understand the almost sure constraint as a probabilistic constraint with safety level one which offers yet another perspective. Finally, robust optimization can be applied efficiently when the support is sufficiently simple. A comparative study of these five different methodologies is carried out and illustrated.

Communicated by Wim van Ackooij.

Extended author information available on the last page of the article

Published online: 27 December 2024

🌿 Springer

1 Introduction

Many physics-based systems that can be described mathematically as optimization problems contain inputs or parameters that are unknown. Ignoring an available model for the uncertain values, for example a probability distribution, can result in severely sub-optimal solutions. In the context of PDE-constrained optimization under uncertainty, the framework of stochastic optimization has proved useful due to its rich theory and wealth of numerical methods [15]. The simplest ansatz is to optimize with respect to a desired *average* outcome [15, 20, 26, 42] of the underlying random PDE, but robust [3, 34] and risk-averse formulations have also been proposed for engineering applications [16, 46] with theoretical investigations in [5, 23, 36, 37, 39].

In certain applications, additional constraints on the solution to the random PDE are also desirable, leading to *state constraints*. Systems involving an almost sure or robust model for state constraints have recently been investigated in [20, 22, 23, 26]. Probabilistic constraints offer the possibility to deal with uncertain restrictions in a robust way that is interpretable with respect to probability. They have been introduced by Charnes et al. [10]. Fundamental algorithmic and theoretical contributions are due to Prékopa, see [45]. More recent presentations are provided in [48] or [52]. In the last years, one may observe a growing interest in probabilistic constraints as part of PDE-constrained optimization, see e.g., [9, 18, 19, 27, 29, 38, 44, 47, 49, 55]. These works include proposals for numerical approaches, structural investigations, and applications to problems of energy management. For instance, [18] considers existence of solutions and their stability with respect to perturbations of the random distribution in a PDE constrained control problem subject to. In [49], structural properties and convergence approaches for chance-constrained optimization of boundary-value elliptic PDEs have been analyzed. A problem with boundary Neumann control of the 1D wave equation subject to a chance constraint with respect to terminal energy was in the focus of [19]. Numerical approaches based on kernel density estimators or alternative smoothing methods were discussed in [9, 27, 47]. Among practical applications, the optimal control of gas networks with transient flow played a prominent role, see [29, 43, 47].

A central challenge in optimization under uncertainty is in the numerical solution. Classical approaches involve discretization of the stochastic space [7, 21], stochastic collocation [12, 13, 35, 58], or using a sample average approximation [2, 30, 31, 42, 57]. Stochastic approximation, which dynamically samples over the course of optimization, has also gained attention [24–26, 38, 40, 41]. Other innovations include the use of surrogate functions constructed using Taylor approximations of the objective and constraint function [1, 11]. Very few approaches exist that can handle state constraints. A Monte Carlo approximation was used in combination with a Moreau–Yosida penalty in [20]. In [38], almost sure state constraints were relaxed to an expectation constraint and a stochastic approximation approach was proposed. In [6], random fields are approximated by the tensor-train decomposition and state constraints are handled using a Moreau–Yosida-type penalty with a softplus approximation for the positive part function. All of these approaches rely on some form of regularization, and are not intended for solving a problem with probabililistic state constraints. In this paper, we tackle this case by working with derived formulas for the Clarke subdifferential of the corresponding probability function. While obtaining such a formula comes at the price

of a more restrictive setting for our model, we are able to handle randomness directly and without the use of regularization. For the almost sure case, we are forced to resort to the same kind of regularization as is used in the previously-mentioned papers. We note that the numerical approaches investigated here may also be applied to problems with probabilistic state constraints and time-dependent PDEs as in [55].

The current work is a follow up paper to [22], where optimality conditions for a risk-averse PDE-constrained optimization problem with uncertain state constraint was considered. Risk aversion was modeled by probabilistic and almost sure constraints. In the current paper, we pick up the same optimization problem (subject to Poisson's equation with a distributed control and an additional random source term on the right-hand side), but focus on its numerical solution. The paper is organized as follows: after presenting the model in Sect. 2, we analyze the problem in Sect. 3. The main result is an improvement of an exact formula for the Clarke subdifferential of the probability function provided in [22] in that it omits a restrictive assumption on boundedness of the set of feasible realizations of the random vector. Section 4 is then devoted to the numerical solution of the control problem under probabilistic (uniform) state constraints in 1D and 2D. Our approach for dealing with probabilistic constraints is based on the well-studied spherical radial decomposition. In Sect. 5, we pass to almost sure constraints. For their numerical treatment, we follow a different methodology than for probabilistic constraints, namely sampling of the distribution (or its support) and applying a Moreau–Yosida approximation. Four different approaches are compared with a reference solution obtained by robust optimization.

2 The Model

In this paper, we discuss numerical approaches for solving the following risk-averse PDE-constrained optimization problem under uncertainty:

$$\min_{u \in L^2(D)} \quad F(u) \tag{1a}$$

$$\text{s.t.} \quad -\Delta y(x, \omega) = u(x) + f(x, \xi(\omega)), \quad x \in D \quad \mathbb{P}\text{-a.s.,} \tag{1b}$$

$$y(x, \omega) = 0, \quad x \in \partial D \quad \mathbb{P}\text{-a.s.,} \tag{1c}$$

$$\mathbb{P}(y(x, \omega) \leq \alpha \quad \forall x \in D) \geq p. \tag{1d}$$

Here, $D \subseteq \mathbb{R}^d$ ($d = 1, 2, 3$) is an open and bounded set. Moreover, $F \colon L^2(D) \to \mathbb{R}$ is a convex, Fréchet differentiable cost function, $\xi \sim \mathcal{N}(0, \Sigma)$ is a centered m-dimensional Gaussian random vector defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $p \in (0, 1]$ is some given probability level, and $\alpha > 0$ is some upper threshold for the random state $y(\cdot, \omega)$. The function $f \colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}$ is a random source term. Inequality (1d) is a probabilistic constraint expressing the condition that the random state y stays below α uniformly on D with probability at least p. Throughout this paper, we will make the following assumptions on the PDE (1b)–(1c). In the following, the notation meas($\cdot$) refers to the d-dimensional Lebesgue measure.

Assumption 2.1 The open and bounded set $D \subseteq \mathbb{R}^d$ $(d = 1, 2, 3)$ is of class S,[1] meaning that there exist constants $\gamma \in (0, 1)$ and $r_0 > 0$ such that $\operatorname{meas}(B_r(x) \backslash D) \geq \gamma \operatorname{meas}(B_r(x))$ for all $x \in \partial D$ and for all $r < r_0$. Additionally, the function $f : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}$ is defined by

$$f(x, z) := f_0(x) + \sum_{i=1}^{m} z_i \phi_i(x) \tag{2}$$

for some given $f_0, \phi_i \in L^2(D)$.

Note that $\mathbb{E}f(\cdot, \xi) = f_0$ on account of ξ being centered. We emphasize that the basic structure of the optimization problem introduced above is intentionally kept simple in order to focus the attention on the aspect of probabilistic state constraints. The possible dependence of the objective on the state y (which would include the tracking-type function frequently used in modeling) is not explicitly stated in (1). Provided the corresponding objective in the reduced formulation is convex and Fréchet differentiable, however, the results in this work apply. We note that the structure imposed on the term in (2) is chosen in such a way so that assumption (GA) in the next section is satisfied, which is needed for obtaining the Clarke subdifferential of the probability function. This structure, however, may come from a truncated Karhunen–Loève expansion, a model frequently used in simulation. We also omit randomness in the diffusion term in (1b) so that (GA) is satisfied. We stress, however, that these structural assumptions are not needed for the almost sure and robust cases. Moreover, it is no problem to pass to more general random distribution settings for ξ, such as: alternative multivariate distributions (Gaussian-like, elliptically symmetric). We will occasionally touch on some of these aspects in subsequent sections.

3 Analytical Properties of the Problem

3.1 General Statements on Optimization Problems with Probabilistic Constraints

We start by embedding problem (1) into a more general framework, which is given by

$$\min_{u \in U} \ F(u) \ \text{s.t.} \ \varphi(u) \geq p \quad (p \in (0, 1]). \tag{3}$$

Here, U is a reflexive and separable Banach space, $F : U \to \mathbb{R}$ is some convex, Fréchet differentiable cost function, and $\varphi : U \to \mathbb{R}$ denotes a probability function defined by $\varphi(u) := \mathbb{P}(\omega \mid g(u, \xi(\omega)) \leq 0)$. In this last expression, ξ is an m-dimensional Gaussian random vector defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and having a centered Gaussian distribution $\mathcal{N}(0, \Sigma)$ with covariance matrix Σ, and $g : U \times \mathbb{R}^m \to \mathbb{R}$ is

---

[1] For the domain, we use the terminology from [33] and note that class S covers many cases, including Lipschitz or convex domains.

some constraint function. We make the following general assumption on g:

$$g \text{ is locally Lipschitzian and } g(u, \cdot) \text{ is convex for all } u \in U. \qquad \text{(GA)}$$

We shall say that g satisfies the *condition of moderate growth* at $\bar{u} \in U$, if

$$\exists l > 0 \ \forall d \in U : g^\circ(\cdot, z)(u; d) \leq l \, \|z\|^{-m} \exp\left(\frac{\|z\|^2}{2\|\Sigma^{1/2}\|^2}\right) \|d\| \qquad (4)$$

$$\forall u \in \mathbb{B}_{1/l}(\bar{u}), \forall z : \|z\| \geq l$$

Here, $\|\cdot\|$ denotes the ℓ_2-norm of vectors and matrices, and $g^\circ(\cdot, z)(u; d)$ refers to the Clarke directional derivative of the locally Lipschitzian (by (GA)) partial function $g(\cdot, z)$ at the argument u in direction d. Moreover, $\Sigma^{1/2}$ denotes a root of Σ. As a consequence of the *spherical-radial decomposition* of Gaussian random vectors (e.g., [17, 28]), the (total) probability function φ can be represented as a spherical integral with respect to the uniform distribution μ_ζ on $\mathbb{S}^{m-1}$

$$\varphi(u) = \int_{\mathbb{S}^{m-1}} e(u, v) \, \mathrm{d}\mu_\zeta(v) \quad (u \in U) \qquad (5)$$

over a one-dimensional radial probability function $e \colon U \times \mathbb{S}^{m-1} \to \mathbb{R}$ defined by

$$e(u, v) := \mu_\chi(\{r \geq 0 \mid g(u, r\Sigma^{1/2}v) \leq 0\}), \qquad (6)$$

where μ_χ is the one-dimensional Chi-distribution with m degrees of freedom. We refer to [32, 50–52, 54, 56] for more details and generalizations of such decomposition to other classes of distributions. The following result on subdifferentiation under the integral sign holds true:

Theorem 3.1 *([32], Theorem 5, Corollary 2 and Proposition 6) Under the basic assumptions (GA), let $\bar{u} \in U$ be given such that $g(\bar{u}, 0) < 0$ and (4) is satisfied. Then, φ and the $e(\cdot, v)$, $(v \in \mathbb{S}^{m-1})$, are locally Lipschitzian around $\bar{u}$ and, with ∂^C denoting the Clarke subdifferential, it holds that*

$$\partial^C \varphi(\bar{u}) \subseteq \int_{\mathbb{S}^{m-1}} \partial^C_u e(\bar{u}, v) \, d\mu_\zeta(v). \qquad (7)$$

Here, the integral on the right-hand side is multivalued and understood in the sense of Aumann. $\partial^C \varphi(\bar{u})$ reduces to a singleton and equality holds in (7), if additionally, the condition

$$\mu_\zeta(\{v \in \mathbb{S}^{m-1} \mid \#\partial_u^C e(\bar{u}, v) \geq 2\}) = 0 \tag{8}$$

is satisfied (here, #A means the cardinality of A). As a consequence, φ is strictly differentiable at $\bar{u}$ in the Hadamard sense [14, p. 30].

We note that differentiability of φ, and thus condition (8), may be violated in general (see [22, Example 2.15]). Therefore, the question arises if there are weaker conditions than those enforcing differentiability as in (8), which still guarantee equality in (7). It was shown in [22, Theorem 2.3] that this holds true under the additional assumptions that g is jointly (!) convex in both variables and that the set

$$\{z \in \mathbb{R}^m \mid g(\bar{u}, z) \leq 0\} \tag{9}$$

is bounded. The condition (9) turns out to be quite restrictive (for instances where it is violated, see [22, Examples 2.15 and 2.16]) and it is difficult to check, in general. Now, we are going to state a result, which allows us to avoid the aforementioned conditions at the price of a growth condition which, however, turns out to be automatically satisfied in our control problem (1). We shift the rather lengthy proof to the appendix. In what follows, for a function $h\colon U \times \mathbb{R}^m \times K \to \mathbb{R}$, we denote by $D_u h(u, z, w)$ and $D_z h(u, z, w)$ the partial derivatives of h with respect to u and z, respectively.

Theorem 3.2 *Let some $\bar{u} \in U$ be given. For $(u, z) \in U \times \mathbb{R}^m$ assume that $g(u, z) = \sup_{w \in K} h(u, z, w)$, where $K \subseteq \mathbb{R}^l$ is compact and $h\colon U \times \mathbb{R}^m \times K \to \mathbb{R}$ satisfies the conditions*

1. *h is continuous, Fréchet differentiable in its first two arguments, and convex in its second argument;*
2. *the mapping $(u, z, w) \mapsto (D_u h(u, z, w), D_z h(u, z, w))$ is continuous;*
3. *$h(\bar{u}, 0, w) < 0 \quad \forall w \in K$;*
4. *$\exists c > 0 : \|D_u h(u, z, w)\| \leq c \exp(\|z\|) \ \forall u : \|u - \bar{u}\| \leq c^{-1} \ \forall z : \|z\| \geq c \ \forall w \in K$.*

Then, $-\varphi$ is regular at $\bar{u}$ in the sense of Clarke and equality holds in (7).

3.2 (Sub-)Differential of the Probability Function in the Concrete Optimal Control Problem

We now want to calculate the exact subdifferential of the probability function associated with problem (1). A corresponding result has been previously obtained in [22, Theorem 2.13]. However, the restrictive condition requiring the set (9) to be bounded was imposed there. The results of the previous section allow us to get rid of this assumption, which we will now show. To this end, denote by $S\colon L^2(D) \times \mathbb{R}^m \to \mathcal{C}(\bar{D})$ the parameterized control-to-state operator assigning to each $u \in L^2(D)$ and each $z \in \mathbb{R}^m$ the solution y of the PDE (1b)-(1c) with right-hand side $u + f(\cdot, z)$. Due to Assumption 2.1 and [22, Lemma 1.2], the operator S is well-defined and bounded. Then the control problem (1) may be recast in the general form of (3) upon defining the function

$g(u, z) := \sup_{x \in D}[S(u, z)](x) - \alpha$. Since S maps into $\mathcal{C}(\bar{D})$, we may write

$$g(u, z) := \max_{x \in \bar{D}}[S(u, z)](x) - \alpha. \tag{10}$$

Indeed, with $y = S(u, z)$ being the state associated with control u and paramater z, in the PDE (1b)–(1c), the inequality $g(u, z) - \alpha \leq 0$ is equivalent with $y(x) \leq \alpha$ for all $x \in D$. Now, identifying the parameter z with the realization $\xi(\omega)$ of the random vector ξ, the definition of φ yields that the constraint $\varphi(u) \geq p$ in (3) is equivalent with (1d).

Due to linearity, one has the decomposition

$$S(u, z) = \bar{y}_{(u)} + \sum_{i=1}^{m} z_i y^{(i)} \quad (u \in L^2(D), z \in \mathbb{R}^m), \tag{11}$$

where $\bar{y}_{(u)}$ is the solution of the PDE

$$-\Delta y(x) = u(x) + f_0(x) \quad x \in D; \quad y(x) = 0, \quad x \in \partial D \tag{12}$$

and the $y^{(i)}$ are the solutions of the PDEs

$$-\Delta y(x) = \phi_i(x) \quad x \in D; \quad y(x) = 0, \quad x \in \partial D \quad (i = 1, \ldots, m). \tag{13}$$

We shall refer to $\bar{y}_{(u)}$ as the mean state associated with the control u (because it relates to the mean value f_0 of the random source term) and to the $y^{(i)}$ as the basic states (because they relate to the basic functions ϕ_i in the random source term with no control in action). Moreover, to each $x \in D$ we assign a dual element $\mathbf{u}_x \in (L^2(D))^*$ by $\mathbf{u}_x(h) := y^h(x)$ for all $h \in L^2(D)$, where y^h is the solution of the control-only PDE

$$-\Delta y(x) = h(x) \quad x \in D; \quad y(x) = 0, \quad x \in \partial D. \tag{14}$$

Theorem 3.3 *In the control problem (1), fix some point of interest $\bar{u} \in L^2(D)$. Assume that the mean state $\bar{y}_{(\bar{u})}$ associated with $\bar{u}$ satisfies the condition*

$$\bar{y}_{(\bar{u})}(x) < \alpha \quad \forall x \in \bar{D}. \tag{15}$$

Then, the probability function φ associated with our control problem is locally Lipschitzian at $\bar{u}$ and has the exact Clarke subdifferential

$$\partial^C \varphi(\bar{u}) = - \int_{\{v \in \mathbb{S}^{m-1} : \rho(v) < \infty\}} \mathrm{clco}\left\{ \frac{f_\chi(\rho(v))}{\varkappa(v, x)} \cdot \mathbf{u}_x \mid x \in M^*(v) \right\} d\mu_\zeta(v),$$

where "clco" means the closed convex hull and for each $v \in \mathbb{S}^{m-1}$ and $x \in \bar{D}$, f_χ refers to the density of the one-dimensional Chi-distribution μ_χ with m degrees of freedom and

$$\varkappa(v, x) := \sum_{i=1}^{m} (\Sigma^{1/2} v)_i \, y^{(i)}(x),$$

$$\rho(v) := \max\{r \geq 0 \mid \bar{y}_{(\bar{u})}(x) + r\varkappa(v, x) \leq \alpha \ \forall x \in \bar{D}\},$$

$$M^*(v) := \{x \in \bar{D} \mid \varkappa(v, x) > 0, \ \bar{y}_{(\bar{u})}(x) + \rho(v)\varkappa(v, x) = \alpha\}.$$

Proof It has been shown in [22, Lemma 2.11, Lemma 2.12] that the integral

$$\int_{\mathbb{S}^{m-1}} \partial_u^C e(\bar{u}, v) \, \mathrm{d}\mu_\zeta(v)$$

coincides with the right-hand side of the identity in the statement of this theorem. Therefore, it will be sufficient to apply Theorem 3.2 to our control problem (1) in order to use (7) as an equality. Clearly, this would yield the asserted exact formula. Hence, we only have to check the assumptions of Theorem 3.2. Given the definition (10), we set $K := \bar{D}, w := x, h(u, z, x) := [S(u, z)](x) - \alpha$ and observe that we are in the setting of Theorem 3.2. We check the assumptions of that theorem: as for assumption 1., let us first verify the continuity of h by considering a sequence $(u_n, z_n, x_n) \to (u, z, x)$ in the set $L^2(D) \times \mathbb{R}^m \times \bar{D}$. As shown as part of the proof in [22, Lemma 2.7], there exists a constant $C > 0$ such that for all $(u_1, z_1), (u_2, z_2) \in L^2(D) \times \mathbb{R}^m$,

$$\|[S(u_1, z_1)] - [S(u_2, z_2)]\|_{\mathcal{C}(\bar{D})} \leq C(\|u_1 - u_2\|_{L^2(D)} + \|z_1 - z_2\|). \qquad (16)$$

Accordingly,

$$|h(u_n, z_n, x_n) - h(u, z, x)| \leq$$
$$|[S(u_n, z_n)](x_n) - [S(u, z)](x_n)| + |[S(u, z)](x_n) - [S(u, z)](x)| \leq$$
$$C(\|u_n - u\|_{L^2(D)} + \|z_n - z\|) + |[S(u, z)](x_n) - [S(u, z)](x)|.$$

Here, both terms on the right converge to zero (the second one thanks to $S(u, z) \in \mathcal{C}(\bar{D})$). This shows the continuity of h. As observed in [22, p.3], the operator S admits for some $y_0 \in \mathcal{C}(\bar{D})$ a decomposition

$$S(u, z) = P(u, z) + y_0 \quad \forall (u, z) \in L^2(D) \times \mathbb{R}^m,$$

where $P : L^2(D) \times \mathbb{R}^m \to \mathcal{C}(\bar{D})$ is some continuous linear operator. Consequently, for fixed $x \in \bar{D}$, the function $(u, z) \mapsto h(u, z, x)$ is the sum of a constant $y_0(x) - \alpha$ and a continuous linear function $(u, z) \mapsto [P(u, z)](x)$. This implies that h is Fréchet differentiable and convex in its first two arguments. Moreover, the partial Fréchet derivative $D_{(u,z)}h$ at (u, z, x) equals the continuous linear function $[P(\cdot, \cdot)](x)$. Altogether, we have shown the validity of assumptions 1. and 2. of Theorem 3.2. Assumption 3. follows immediately from (15). In order to check assumption 4., observe that by (16), the function $u \mapsto h(u, z, x)$ is locally Lipschitzian with some common modulus C at

all $u \in L^2(D)$ and for all $(z, x) \in \mathbb{R}^m \times \bar{D}$. Consequently,

$$\|D_u h(u, z, x)\| \leq C \leq c e^{\|z\|} \; \forall (u, z, x) \in L^2(D) \times \mathbb{R}^m \times \bar{D} : \|z\| \geq c := \max\{1, \log C\}.$$

This shows the validity of assumption 4. and finishes the proof. $\qquad \square$

Remark 3.1 One may easily derive from [50, Prop. 3.11] that condition (15) is satisfied whenever the probability $\varphi(\bar{u})$ is larger than or equal to 0.5. Since typically the probability level p is chosen close to one, this condition will be automatically fulfilled for most iterations of the control. Hence, the exact formula for the Clarke subdifferential in Theorem 3.2 will basically hold true unconditionally.

Remark 3.2 If the sets $M^*(v)$ introduced in Theorem 3.3 satisfy the condition

$$\#M^*(v) = 1 \quad \mu_\zeta - \text{a.e.} \quad v \in \mathbb{S}^{m-1}, \tag{17}$$

then the integral in the formula for the Clarke subdifferential, and hence the Clarke subdifferential itself, reduce to singletons. Therefore, the probability function is strictly differentiable in the Hadamard sense (see [14, Prop. 2.2.4]) and its derivative is given by

$$D\varphi(\bar{u}) = - \int\limits_{\{v \in \mathbb{S}^{m-1} : \rho(v) < \infty\}} \frac{f_\chi(\rho(v))}{\varkappa(v, x^*(v))} \cdot \mathbf{u}_{x^*(v)} \, d\mu_\zeta(v), \tag{18}$$

where, for μ_ζ-almost every $v \in \mathbb{S}^{m-1}$, $x^*(v)$ is defined as the unique element in the set $M^*(v)$.

Simple examples show that in the setting of Theorem 3.3 the probability function may fail to be differentiable, hence condition (17) is violated in general. Evidently, (17) is difficult to verify for concrete data. In [53, Lemma 4.3], some easy to understand constraint qualification (so-called "rank 2-CQ") has been shown to imply (17) for finite random inequality systems. A generalization to our setting with infinite systems (uniform state constraints) does not seem to be straightforward, so that the verification of differentiability for the probability function remains an open problem. On the other hand, in the numerical solution of the given control problem, one may be forced to replace the uniform state constraint (1d) by evaluating it on a finite subset $\tilde{D} \subseteq D$ of the domain. Then, the aforementioned rank 2-CQ reduces to the following verifiable condition at some fixed control $\bar{u} \in U$:

$$\text{rank} \{Y(x^a), Y(x^b)\} = 2 \quad \forall x^a, x^b \in \tilde{D} :$$
$$x^a \neq x^b, \quad \bar{y}_{(\bar{u})}(x^a) + \langle z, Y(x^a) \rangle = \bar{y}_{(\bar{u})}(x^b) + \langle z, Y(x^b) \rangle = \alpha$$
$$\forall z \in \mathbb{R}^m : \bar{y}_{(\bar{u})}(x) + \langle z, Y(x) \rangle \leq \alpha \; \forall x \in \tilde{D}.$$

Here, for $x \in \tilde{D}$, the vector $Y(x) := (y^{(1)}(x), \ldots, y^{(m)}(x))$ collects in its components all values of the basic functions $y^{(i)}$ at x defined in (13). This condition can be evidently strengthened to the much easier condition

$$\text{rank}\,\{Y(x^a), Y(x^b)\} = 2 \quad \forall x^a, x^b \in \tilde{D} : x^a \neq x^b$$

3.3 Convexity Properties Under Gaussian and Truncated Gaussian Distributions

In this section, we show that problem (1) is convex under Gaussian and truncated (to convex sets) Gaussian distributions. Given the assumed convexity of the objective F, it is sufficient to verify convexity of the constraint. For pure Gaussian distributions, this has been done (without explicit statement) in [22]. We refer to the abstract form (3) of problem (1), where φ is defined as the Gaussian probability function associated with the random inequality $g(u, \xi(\omega)) \leq 0$. In the context of our problem (1), this constraint function takes the form (10). Similarly, we may consider problem (1) under a truncated Gaussian distribution and represent it in the abstract form (3) with φ being replaced by the truncated Gaussian probability function $\tilde{\varphi}$ associated with the same constraint function g from (10). More precisely, if $\tilde{\xi} \sim \mathcal{TN}(\mu, \Sigma, \Xi)$ designates a random vector having a Gaussian distribution $\mathcal{N}(\mu, \Sigma)$ truncated to a closed set $\Xi := \{z \in \mathbb{R}^m \mid s(z) \leq 0\}$ ($s : \mathbb{R}^m \to \mathbb{R}$ continuous), then

$$\tilde{\varphi}(u) = \mathbb{P}(\omega \mid g(u, \tilde{\xi}(\omega)) \leq 0) = \frac{\mathbb{P}(\omega \mid g(u, \xi(\omega)) \leq 0, \; s(\xi(\omega)) \leq 0)}{\mathbb{P}(\omega \mid s(\xi(\omega)) \leq 0)}.$$

The last equation above expresses the fact that the truncated distribution is nothing but the original distribution conditioned to a subset. Consequently, the probabilistic constraint $\tilde{\varphi}(u) \geq p$ in (3) (with φ replaced by $\tilde{\varphi}$) can be equivalently written as $\mathbb{P}(\omega \mid \tilde{g}(u, \xi(\omega)) \leq 0) \geq \tilde{p}$, where

$$\tilde{g}(u, z) := \max\{g(u, z), s(z)\}, \quad \tilde{p} := p \cdot \mathbb{P}(\omega \mid s(\xi(\omega)) \leq 0). \tag{19}$$

Proposition 3.1 *The optimization problem (1) is convex if the random vector ξ follows a Gaussian distribution or a Gaussian distribution truncated to a set $\Xi := \{z \in \mathbb{R}^m \mid s(z) \leq 0\}$ where $s : \mathbb{R}^m \to \mathbb{R}$ is convex, hence continuous.*

Proof The convexity of the constraint $\varphi(u) \geq p$ in case of a Gaussian distribution follows in the abstract setting from the convexity of the constraint function $g(u, z)$ (jointly in both variables), see [22, Lemma 2.4]. That this property holds true for the concrete function g defined in (10), has been shown in [22, Lemma 2.7]. As for the case of a truncated distribution, we have seen above that we may represent the corresponding chance constraint as in the Gaussian case but with the modified constraint function $\tilde{g}$ and the modified probabilty level $\tilde{p}$ defined in (19). Evidently, $\tilde{g}$ is jointly convex in (u, z) since g is so and s is convex. Consequently, we can apply again the previous argument to prove the convexity of the feasible set also under truncated Gaussian distribution. $\qquad\square$

4 Numerical Results for the Control Problem with Probabilistic Constraints

4.1 Implementation of the Spherical-Radial Decomposition

In order to deal with the probabilistic constraint $\varphi(u) \geq p$ in the general problem (3) numerically, one has to appropriately approximate the probability function φ and, assuming differentiability, its derivative $D\varphi$. Since φ is defined as the spherical integral (5), it can be approximated by a finite sum

$$\varphi(u) \approx \hat{\varphi}(u) := \frac{1}{K} \sum_{k=1}^{K} e(u, v^{(k)}) \quad \forall u \in U,$$

where $\{v^{(k)}\}_{k=1}^{K}$ is a sample of size K of the uniform distribution on the sphere $\mathbb{S}^{m-1}$. A simple way to create such a sample consists in normalizing to unit length a quasi-Monte Carlo (QMC) sample of the given Gaussian distribution. It was shown in [22, eq. (19)] that, for any fixed $\bar{u}$ with $g(\bar{u}, 0) < 0$, there exists a neighborhood $\mathcal{N}$ of $\bar{u}$ such that

$$e(\bar{u}, v) = \begin{cases} F_\chi(\rho^*(v)) \text{ if } \rho^*(v) < \infty \\ 1 \quad\quad\quad\text{ if } \rho^*(v) = \infty \end{cases} \quad \forall u \in \mathcal{N} \; \forall v \in \mathbb{S}^{m-1}, \tag{20}$$

where F_χ is the cumulative Chi-distribution function introduced before and

$$\rho^*(v) := \sup\{r \geq 0 \mid g(\bar{u}, r\Sigma^{1/2}v) \leq 0\} \quad (v \in \mathbb{S}^{m-1}).$$

It follows from [22, eq. (29)] that ρ^* coincides with the function ρ defined in the statement of Theorem 3.3. Consequently, the probability function can be approximated at some $\bar{u}$ with $g(\bar{u}, 0) < 0$ by

$$\begin{aligned} \varphi(\bar{u}) \approx \hat{\varphi}(\bar{u}) = {} & \frac{1}{K} \#\{k \in \{1, \ldots, K\} \mid \rho(v^{(k)}) = \infty\} \\ & + \frac{1}{K} \sum_{\{k \in \{1, \ldots, K\}: \rho(v^{(k)}) < \infty\}} F_\chi(\rho(v^{(k)})). \end{aligned}$$

Should φ be differentiable, the spherical integral in (18) can be approximated by the sample-based quantity

$$D\varphi(\bar{u}) \approx D\hat{\varphi}(\bar{u}) := - \sum_{\{k \in \{1, \ldots, K\}: \rho(v^{(k)}) < \infty\}} \frac{f_\chi(\rho(v^{(k)}))}{\varkappa(v^{(k)}, x^*(v^{(k)}))} \cdot \mathbf{u}_{x^*(v^{(k)})},$$

whose ingredients are defined in the statement of Theorem 3.3 and below (18), respectively.

It is essential to observe that values and derivatives (in a specified direction $h \in L^2(D)$) can be simultaneously updated by sharing the same sample $v^{(k)}$. In this way, the potentially time-consuming computation of $\rho(v^{(k)})$ may be carried out only once, when it comes to updating the probabilities themselves, whereas it does not have to be redone during the update of the gradient. Altogether, this leads us to the algorithmic scheme in Algorithm 1.

Algorithm 1 Algorithmic approximation of probabilities and their gradients via spherical-radial decomposition

---

Input: $D \subset \mathbb{R}^d$ (domain), f_0 (expected source term), ϕ_i $(i = 1, \ldots m)$ (basic source functions), Σ (covariance matrix of Gaussian distribution for random coefficients), $\bar{u}$ (control of interest satisfying (15)), h (direction of interest to apply the gradient to).

1. Initialization
    (a) Fix sample size K and create a sample $\{v^{(k)}\}_{k=1}^K$ uniformly distributed on the sphere $\mathbb{S}^{m-1}$.
    (b) $prob := 0, deriv := 0, k := 1$.
    (c) Determine functions $y^{(i)}$ $(i = 1, \ldots m)$ as the solutions of (13).
    (d) Determine function $\bar{y}_{(\bar{u})}$ as the solution to (12) (with $u := \bar{u}$).
    (e) Determine function y^h as the solution to (14).

2. Iteration on k
    (a) Define a function $\varkappa := \sum\limits_{i=1}^{m} (\Sigma^{1/2} v^{(k)})_i \, y^{(i)}$
    (b) Calculate $\rho := \max\{r \geq 0 \mid \bar{y}_{(\bar{u})}(x) + r \varkappa(x) \leq \alpha \ \forall x \in \bar{D}\}$ (e.g., by evaluation on a grid over $\bar{D}$).
    (c) Determine a point $x^* \in \bar{D}$ satisfying $\bar{y}_{(\bar{u})}(x^*) + \rho \varkappa(x^*) = \alpha$
        (if x^* is not unique then the probability function fails to be differentiable and the output will yield a subgradient rather than a gradient).
    (d) If $\rho < \infty$ then $prob := prob + F_\chi(\rho)$ else $prob := prob + 1$.
    (e) If $\rho < \infty$ then $deriv := deriv + \frac{f_\chi(\rho)}{\varkappa(x^*)} y^h(x^*)$.

3. Termination: If $k < K$ then $k := k + 1$. Goto Step 2.
4. Output: $prob := K^{-1} prob, deriv := K^{-1} deriv$

Output: Approximation of probability $\varphi(\bar{u}) \approx prob$ and of its directional derivative $D\varphi(\bar{u})(h) \approx deriv$.

4.2 Results in Dimension One

We consider problem (1) with the following data:

$$d = 1, D = (0, 1), m = 6, \alpha = 0.2, F(u) = \|u\|^2_{L^2(D)},$$

$$p = 0.9, \Sigma_{i,j} = 9 \cdot 0.6^{|i-j|} \ (i, j = 1, \ldots m),$$

$$f_0(x) = 5x^2, \phi_{2i-1}(x) = \sin(ix) \ (i = 1, 2, 3), \phi_{2i-2}(x) = \cos(x/i) \ (i = 2, 3, 4).$$
$$\tag{21}$$

We use a finite difference discretization over a subdivision of the domain into 120 intervals. The values and—assuming differentiability—derivatives of the probability function were obtained as described in Algorithm 1 on the basis of a QMC sample on the unit sphere of size 512. The optimization problem (1) (with the data described
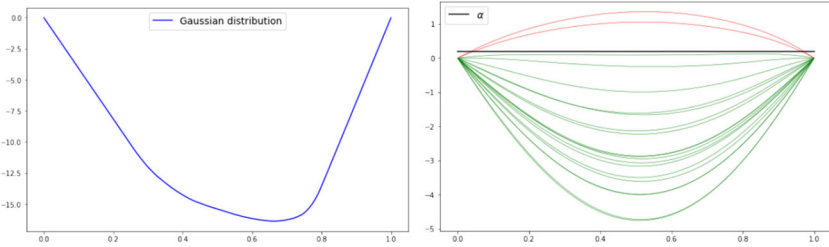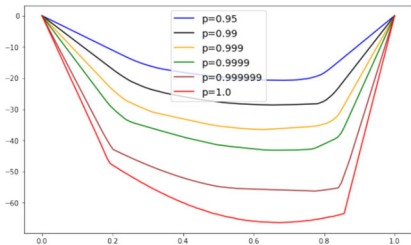
Fig. 1 Numerical solution of (1). Optimal control for probability level $p = 0.9$ (left) and associated states for twenty scenarios of the random source term (right)

Fig. 2 Solution of (1) for Gaussian distribution truncated to an ellipsoid and for increasing probability levels

in (21)) was numerically solved with the "SLSQP" method from the Python standard minimzation routine *scipy.optimize.minimize*. Figure 1 (left) shows the optimal control under the indicated Gaussian distribution.

Figure 1 (middle) plots twenty states associated with the optimal control and with twenty scenarios for the random source term $f(x, \xi(\omega))$ in (1b). It can be seen that only two of them exceed the desired threshold α occasionally, whereas the remaining ones stay below it uniformly over the domain, which corresponds well to the imposed probability level of $p = 0.9$.

In order to include the almost sure case ($p = 1$) for later comparison, the problem was solved again with the Gaussian distribution truncated to an ellipsoid $\mathcal{E}$ defined by $\mathcal{E} := \{z \in \mathbb{R}^m \mid z^T \Sigma^{-1} z \leq 36\}$ (the choice of the Mahalanobis norm in this definition facilitates computations). In this way, the support of the distribution becomes compact so that the problem with an almost-sure constraint has a chance to have a feasible solution. As can be seen from Fig. 2, the optimal solutions of the probabilistically constrained problems seem to converge to that with almost sure constraints for $p \to 1$ (for a rigorous statement, see Sect. 5.1).

Note that passing to truncated Gaussian distributions already points to the possibility of using alternative distributions. We do not present details here but note that the same methodology could be applied to multivariate lognormal or Student or Gaussian mixture distributions.

4.3 Results in Dimension Two

In this section, we revisit problem (1) but with some increased complexity when compared to the previous section: the domain will be in two dimensions, the dimension
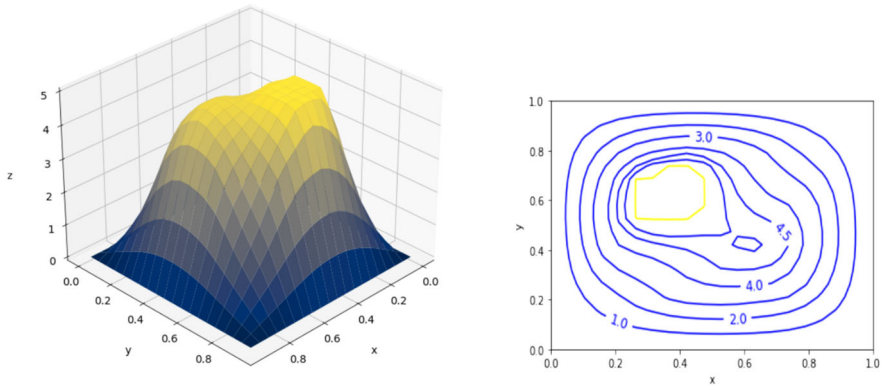
Fig. 3 Optimal solution of the control problem in 2D (left). For graphical reasons the negative control is plotted. Contourplot (right)

of the random parameter will increase from 6 to 30, and additional bound constraints will be imposed on the control. More precisely, we consider the following data:

$$d = 2, D = (0, 1)^2, m = 30, \alpha = 0.2, F(u) = \|u\|_{L^2(D)}^2,$$
$$p = 0.9, \Sigma_{i,j} = 9 \cdot 0.6^{|i-j|} \ (i, j = 1, \ldots m),$$
$$f_0(x_1) = 5x_1 x_2, \phi_i(x_1, x_2) = \sin(i x_1) \cos(i x_2) \ (i = 1, \ldots, m).$$

In addition, we impose the following bounds on the control:

$$-5 \le u(x_1, x_2) \le 0 \quad \text{a.e. } (x_1, x_2) \in D.$$

The PDE was solved using finite differences on a 20x20 grid of the domain and the random parameter was approximated by using $2^{13} = 8192$ QMC samples on the sphere. Figure 3 shows the resulting optimal control (linearly interpolated). It can be seen in both plots that the control bound becomes active in a certain small region of the domain (yellow).

5 Almost Sure Constraints

In this section, we want to consider the extreme case of risk-averse decision making, namely the optimal control with constraints that should be satisfied almost surely. In this case, our PDE-constrained optimization problem becomes

$$\min_{u \in L^2(D)} \quad F(u) \tag{22a}$$

$$\text{s.t.} \quad -\Delta y(x, \omega) = u(x) + f(x, \xi(\omega)), \qquad x \in D \quad \mathbb{P}\text{-a.s.,} \tag{22b}$$

$$\qquad y(x, \omega) = 0, \qquad x \in \partial D \quad \mathbb{P}\text{-a.s.,} \tag{22c}$$

$$\qquad y(x, \omega) \le \alpha, \qquad x \in D \quad \mathbb{P}\text{-a.s.} \tag{22d}$$

We shall assume in this section that the support $\Xi \subseteq \mathbb{R}^m$ of the random vector ξ is compact, because otherwise there may not exist a feasible solution that satisfies (22d). This assumption will meet the setting illustrated in Fig. 1 (right), where the distribution was assumed to be Gaussian, truncated to an ellipsoid.

5.1 Almost Sure Constraints as Limit of Probabilistic Constraints

Evidently, by equivalence of (22d) with (1d) when choosing $p = 1$, this optimization is equivalent with the previously analyzed probabilistically constrained problem (1) when choosing the maximum possible probability level. We emphasize that, in spite of this equivalence, necessary and sufficient optimality conditions can be obtained for the formulation (22) ([22, Theorem 3.10]) while they cannot for the formulation (1) ([22, Remark 2.6 and Example 3.1]). On the other hand, we were able numerically to find a candidate solution to (1) for $p = 1$ and observed a convergence of solutions for $p \to 1$ towards this candidate (see Fig. 2). Two questions arise: is the candidate we found the true solution of the almost sure problem (22) and can we prove the mentioned convergence? The first question will be considered in the next section, whereas the second question will be answered after the following technical preparation.

Lemma 5.1 *Consider problem (1) but with the random vector ξ having a Gaussian distribution truncated to a compact convex set Ξ as in Proposition 3.1. Let $u_{p_k} \in L^2(D)$ be optimal solutions of (1) with $p_k \in (0, 1)$ and $p_k \to 1$. In addition, suppose that $u_{p_k} \rightharpoonup \bar{u} \in L^2(D)$. Then, $\bar{u}$ is an optimal solution of problem (1) with $p = 1$ and $F(u_{p_k}) \to F(\bar{u})$.*

Proof As shown in Sect. 3.3, the constraint of (1) can be rewritten as $\tilde{\varphi}(u) \geq \tilde{p}$, where $\tilde{\varphi}(u) = \mathbb{P}(\omega \mid \tilde{g}(u, \xi(\omega)) \leq 0)$ is a probability function related to an untruncated Gaussian random vector and $\tilde{g}, \tilde{p}$ are defined in (19). By [22, eq. (5)], the functions $(u, z) \mapsto [S(u, z)](x)$ are affine linear (hence weakly lower semicontinuous) for each $x \in \bar{D}$. Consequently, the function g in (10) is weakly lower semicontinuous as a sum of a constant and a maximum of such functions. On the other hand, the function $(u, z) \mapsto s(z)$ with s from Proposition 3.1 is convex and continuous, hence weakly lower semicontinuous. It follows once more that $\tilde{g}$ as the maximum of two weakly lower semicontinuous functions shares this property. Therefore, using [18, Lemma 2], we get that the function φ defined in (3) is weakly sequentially upper semicontinuous, whence $\varphi(\bar{u}) \geq \limsup_{k\to\infty} \varphi(u_{p_k}) = 1$. Now, let $u' \in L^2(D)$ be any control function satisfying $\varphi(u') = 1$. Since in problem (3) with probability level p_k, the control u' is feasible while the control u_{p_k} is optimal, it follows that $F(u_{p_k}) \leq F(u')$. In particular, it shows that $F(u_{p_k}) \leq F(\bar{u})$ and consequently $\limsup_{k\to\infty} F(u_{p_k}) \leq F(\bar{u})$. Therefore, utilizing the weak lower semicontinuity of the objective function, we can conclude that $F(\bar{u}) \leq \liminf_{k\to\infty} F(u_{p_k}) \leq F(u')$, thereby demonstrating the optimality of $\bar{u}$ and $\lim_{k\to\infty} F(u_{p_k}) = F(\bar{u})$. $\qquad\square$

Proposition 5.1 *In addition to the assumptions of Lemma 5.1, assume that the objective F in (1) is strictly convex. Then, any bounded sequence $(u_{p_k})_{k\in\mathbb{N}}$ of minimizers of (1) with $p_k \in (0, 1)$ weakly converges to the unique solution of (1) with $p = 1$. If the objective happens to be even strongly convex (as in the numerical examples), then any*

bounded sequence $(u_{p_k})_{k \in \mathbb{N}}$ of minimizers of problem (1) with $p_k \in (0, 1)$ strongly converges to the unique solution of problem (1) with $p = 1$.

Proof As proven in Proposition 3.1, problem (1) has a convex feasible set for all $p \in (0, 1]$. Therefore, if F is strictly convex, the problem has a unique minimizer. Hence, if u_{p_k} is bounded, then as per Lemma 5.1, the only possible accumulation point is the unique minimizer of (1) for $p = 1$.

Strong convexity of F (with modulus μ) gives

$$F(u_{p_k}) - F(\bar{u}) \geq \langle \nabla F(\bar{u}), u_{p_k} - \bar{u} \rangle + \frac{\mu}{2} \|u_{p_k} - \bar{u}\|^2. \tag{23}$$

From Lemma 5.1, we have $\lim_{k \to \infty} F(u_{p_k}) = F(\bar{u})$. Also $\langle \nabla F(\bar{u}), u_{p_k} - \bar{u} \rangle \to 0$ as $k \to \infty$ by weak convergence of u_{p_k}. Strong convergence of the same sequence follows by (23). □

Note that the last Proposition explains why in Fig. 2 we observe strong and not just weak convergence.

5.2 Moreau–Yosida Approximation for Dealing with Almost Sure Constraints

A difficulty in the numerical solution to problem (22) is in the enforcement of the state constraints. One possibility to handle this computationally is to penalize this constraint in the objective. Note that $(H_0^1(D), L^2(D), H^{-1}(D))$ is a Gelfand triple. Moreover, the cone $\mathcal{K} = \{y \in L_{\mathbb{P}}^\infty(\Omega, H_0^1(D)) \mid y(x, \omega) \leq 0 \text{ a.e.-}\mathbb{P}\text{ -a.s.}\}$ is compatible with the cone $\mathcal{K}_H = \{y \in L_{\mathbb{P}}^2(\Omega, L^2(D)) \mid y(x, \omega) \leq 0 \text{ a.e.-}\mathbb{P}\text{ -a.s.}\}$ in the sense that $\mathcal{K}_H \cap L_{\mathbb{P}}^\infty(\Omega, H_0^1(D)) = \mathcal{K}$. This justifies the penalization on the weaker (Hilbert) space $H = L_{\mathbb{P}}^2(\Omega, L^2(D))$ where the projection $\pi_{\mathcal{K}_H}$ onto the cone $\mathcal{K}_H$ is possibly computationally cheaper. The Moreau–Yosida regularization (or envelope) for the indicator function on the cone $\mathcal{K}_H$ has the formula

$$\hat{\beta}^\gamma(k) = \gamma \|k - \pi_{\mathcal{K}_H}(k)\|_H^2;$$

see, e.g., [23, Section 4.1]. Let $\boldsymbol{\alpha} \in L_{\mathbb{P}}^\infty(\Omega, H^1(D))$ be the function that is equal to α a.e. in D and $\mathbb{P}$-a.s. Penalizing the constraint (22d) amounts to adding $\hat{\beta}^\gamma(y - \boldsymbol{\alpha})$ to the objective, leading to the modified problem

$$\min_{u \in L^2(D)} \left\{ f^\gamma(u) := F(u) + \gamma \mathbb{E}[\| \max(0, y - \boldsymbol{\alpha})\|_{L^2(D)}^2] \right\} \tag{24a}$$

$$\text{s.t.} \quad -\Delta y(x, \omega) = u(x) + f(x, \xi(\omega)), \qquad x \in D \quad \mathbb{P}\text{ -a.s.,} \tag{24b}$$

$$y(x, \omega) = 0, \qquad x \in \partial D \quad \mathbb{P}\text{ -a.s.} \tag{24c}$$

The results from [23] focus on the consistency of the optimality conditions for (24) to the optimality conditions for (22), which require an interior point condition

(and therefore higher regularity of the solution to the PDE). Here, we will focus on the consistency of the primal problem and will work with the weakest regularity available. Let $A: H_0^1(D) \to H^{-1}(D)$ represent the Laplacian, which is a linear isomorphism thanks to Assumption 2.1 and the Lax–Milgram lemma. We denote with $\mathcal{A}^{-1}: L_\mathbb{P}^\infty(\Omega, L^2(D)) \to L_\mathbb{P}^\infty(\Omega, H_0^1(D))$ the superposition defined by $[\mathcal{A}^{-1}y](\omega) = A^{-1}y(\cdot, \omega)$. Let $\mathcal{B}: L^2(D) \to L_\mathbb{P}^\infty(\Omega, L^2(D))$ be the canonical embedding. Let $\tilde{f}$ be defined by $\tilde{f}(x, \omega) := f(x, \xi(\omega))$. It will be convenient to define the (affine linear) control-to-state operator $\tilde{\mathcal{S}}: L^2(D) \to L_\mathbb{P}^\infty(\Omega, H_0^1(D))$ by $\tilde{\mathcal{S}}(u) = \mathcal{A}^{-1}(\mathcal{B}u + \tilde{f})$, which is bounded by [22, Lemma 10]. Due to the continuous canonical embedding $\iota: L_\mathbb{P}^\infty(\Omega, H_0^1(D)) \to L_\mathbb{P}^2(\Omega, L^2(D))$, the operator $\mathcal{S} := \iota \circ \tilde{\mathcal{S}}$ is continuous from $L^2(D)$ to $L_\mathbb{P}^2(\Omega, L^2(D))$.

We demonstrate how solutions to (24) converge to (22) in the limit as $\gamma \to \infty$. First we need the following.

Lemma 5.2 *For any $\gamma \geq 0$, the reduced functional $\beta^\gamma: L^2(D) \to \mathbb{R}$ defined by*

$$\beta^\gamma(u) = \gamma \mathbb{E}[\|\max(0, \mathcal{S}(u) - \boldsymbol{\alpha})\|_{L^2(D)}^2]$$

is convex and (weakly) lower semicontinuous.

Proof Since $\mathcal{S}$ is continuous and the map $y \mapsto \mathbb{E}[\|\max(0, y - \boldsymbol{\alpha})\|_{L^2(D)}^2]$ is evidently continuous on $L_\mathbb{P}^2(\Omega, L^2(D))$, β^γ is continuous on $L^2(D)$. Moreover, since $\mathcal{S}$ is (affine) linear, β^γ is convex. Weak lower semincontinuity follows from the continuity and convexity of β^γ. □

We now have the following result:

Lemma 5.3 *If F is (in addition to being convex and Fréchet differentiable) coercive, then there exists a solution u^γ to (24) for all $\gamma > 0$. Additionally, if there exists a feasible point for problem (22), then given a sequence (γ_n) with $\gamma_n \to \infty$, weak limit points of u^{γ_n} solve (22).*

Proof Let $F_{\text{ad}} = \{u \in L^2(D) \mid [\mathcal{S}(u)](x, \omega) \leq \alpha \text{ a.e.-} \mathbb{P}\text{-a.s.}\}$ denote the feasible set for problem (22). It is straightforward to show that this set is convex; it is closed thanks to the continuity of $\mathcal{S}$. Thus, F_{ad} is weakly closed (see [8, Theorem 2.23]). Coercivity of F (in combination with the other assumptions) implies that f^γ attains its minimum over F_{ad}. In particular, there exists a minimizing sequence $(u_n)_{n \in \mathbb{N}} \subset F_{\text{ad}}$ such that $\lim_{n \to \infty} f^\gamma(u_n) = \inf_{u \in F_{\text{ad}}} f^\gamma(u)$. This sequence is also bounded due to the coercivity of F, so we can extract a subsequence $(u_{n_k})_{k \in \mathbb{N}}$ such that $u_{n_k} \rightharpoonup u^\gamma$. Now, we have (since F is weakly lower semicontinuous, as a convex and continuous function)

$$\liminf_{k \to \infty} F(u_{n_k}) + \beta^\gamma(u_{n_k}) \geq \liminf_{k \to \infty} F(u_{n_k}) + \liminf_{k \to \infty} \beta^\gamma(u_{n_k}) \geq F(u^\gamma) + \beta^\gamma(u^\gamma),$$

meaning u^γ solves problem (24). In other words, a solution exists for any $\gamma > 0$.

Now, let $(u^{\gamma_n})_{n\in\mathbb{N}}$ be a sequence such that $\gamma_n \to \infty$ as $n \to \infty$. Let $\bar{u} \in F_{\text{ad}} \neq \emptyset$ be a minimizer to (22). Then, by optimality of u^{γ_n} and feasibility of $\bar{u}$, we have

$$F(u^{\gamma_n}) \leq F(u^{\gamma_n}) + \beta^{\gamma_n}(u^{\gamma_n}) \leq F(\bar{u}) + \beta^{\gamma_n}(\bar{u}) = F(\bar{u}). \tag{25}$$

This means u^{γ_n} is a minimizing sequence for F, which is bounded thanks to the coercivity of F. Therefore, there exists a subsequence $(u^{\gamma_{n_k}})_{k\in\mathbb{N}}$ and point $\hat{u}$ such that $u^{\gamma_{n_k}} \rightharpoonup \hat{u}$. Weak lower semicontinuity of F implies that

$$F(\hat{u}) \leq \liminf_{k\to\infty} F(u^{\gamma_{n_k}}) \leq F(\bar{u}). \tag{26}$$

It only remains to show that $\hat{u}$ is feasible for (22). Since F is bounded over F_{ad}, (25) implies that there exists a constant $c > 0$ such that

$$\beta^{\gamma_{n_k}}(u^{\gamma_{n_k}}) \leq c \quad \Leftrightarrow \quad \mathbb{E}[\|\max(0, \mathcal{S}(u^{\gamma_{n_k}}) - \boldsymbol{\alpha})\|_{L^2(D)}^2] \leq \frac{c}{\gamma_{n_k}}.$$

Since $\gamma_{n_k} \to \infty$, we obtain $\mathbb{E}[\|\max(0, \mathcal{S}(u^{\gamma_{n_k}}) - \boldsymbol{\alpha})\|_{L^2(D)}^2] \to 0$, implying that $[\mathcal{S}(\hat{u})](x, \omega) \leq \alpha$ a.e. $\mathbb{P}$-a.s. By definition of $\mathcal{S}$, the function $\mathcal{S}(u)$ satisfies (22b)–(22c). □

Corollary 5.1 *Under the same conditions as Lemma 5.3, suppose that F is strongly convex. Then any bounded sequence $(u^{\gamma_n})_{n\in\mathbb{N}}$ of minimizers of problem (24) with $\gamma_n \to \infty$ converges to the unique solution of problem (22).*

Proof Since there can only be one solution to (22), (26) together with the feasibility of $\hat{u}$ and $\bar{u}$ imply that $\lim_{n\to\infty} F(u^{\gamma_n}) = F(\bar{u})$. Strong convergence follows with the same reasoning used in Proposition 5.1. □

Remark 5.1 While the above proof is based on [23, Proposition 3.8], we used the reduced formulation here, which greatly simplified certain arguments; we did not need to rely on arguments using the weak* topology, for one. Moreover, we did not require an interior point condition to establish consistency (which was assumed in [23] but not used in this proof). We also note that similar arguments are made in [20] for a related problem in the context of generalized Nash equilibrium problems. The single-player case resembles our problem but different assumptions are made on the underlying PDE and objective and therefore do not cover the theory presented here. Moreover, the argument in [20] relies on a constraint qualification that we do not need here.

5.3 Numerical Results by "Sampling the Distribution," "Sampling the Support," or "Sampling the Boundary" of the Support

In the following numerical tests, we use gradient descent to solve a sample average approximation of the subproblem (24). Given a fixed number N of samples $z_1, \ldots, z_N$ of the random vector ξ, the corresponding SAA problem is

$$\min_{u \in L^2(D)} \left\{ \hat{f}^\gamma(u) := F(u) + \frac{\gamma}{N} \sum_{i=1}^{N} [\|\max(0, y_i - \boldsymbol{\alpha})\|_{L^2(D)}^2] \right\} \tag{27a}$$

$$\text{s.t.} \quad -\Delta y_i(x) = u(x) + f(x, z_i), \quad x \in D \quad \forall i = 1, \ldots, N, \tag{27b}$$

$$y_i(x) = 0, \quad x \in \partial D \quad \forall i = 1, \ldots, N. \tag{27c}$$

The gradient of $\hat{f}^\gamma$ can be computed using standard techniques, resulting in

$$\nabla \hat{f}^\gamma(u) = \nabla F(u) - \frac{1}{N} \sum_{i=1}^{N} p_i, \tag{28}$$

where p_i is the solution to the adjoint equation

$$-\Delta p_i(x) = -2\gamma \max(0, y_i(x) - \alpha), \quad x \in D \quad \text{and} \quad p_i(x) = 0 \quad x \in \partial D, \tag{29}$$

and y_i is the solution to (27b)–(27c) for a fixed sample z_i. For numerical tests, we use the same setup as described at the beginning of Sect. 4.2. However, rather than considering the purely Gaussian distribution $\mathcal{N}(0, \Sigma)$ defined there, we truncate it to the ellipsoid $\mathcal{E}$ defined below the original data in connection with Fig. 1 (right). The sampling of this truncated Gaussian distribution is carried out by accepting/rejecting samples of the underlying original Gaussian distribution according to whether or not these samples belong to the ellipsoid $\mathcal{E}$. The PDEs (27b)–(27c) and (29) were solved using FEniCS [4] with a finite element discretization over 29 intervals. We use a path-following approach, where in an outer iteration, the penalty was increased and in inner iterations, gradient descent on a subproblem was performed until a given tolerance. A starting value for the control was set to $u \equiv -1$. In each outer iteration $k \in \{0, \ldots, 8\}$, $\gamma = 10^k$, an increasing batch of 3^k m-dimensional vectors was used. The step-size for each outer iteration was chosen to be $t_\ell = 4/\ell$, which was informed by the strong convexity of the problem; see [24] for a discussion of this choice in the context of PDE-constrained optimization. Each inner iteration was terminated when $\|\nabla \hat{f}^\gamma(u_j)\|_{L^2(D)} < 10^{-4}$. In Fig. 4 (top left), different optimal controls are displayed for increasing values $\gamma = 10^k$ and increasing sample sizes 3^k.

So far, we have considered samples of the distribution in order to characterize solutions of the almost sure model. Sometimes, however, one can do better than sampling the distribution. As observed in [22, Example 3.3], problems with almost sure constraints are not equivalent, in general, to problems with worst-case constraints on the support of the random parameter [22, Lemma 3.4]. They are, however, if the underlying random inequality is lower semicontinuous with respect to the random parameter.
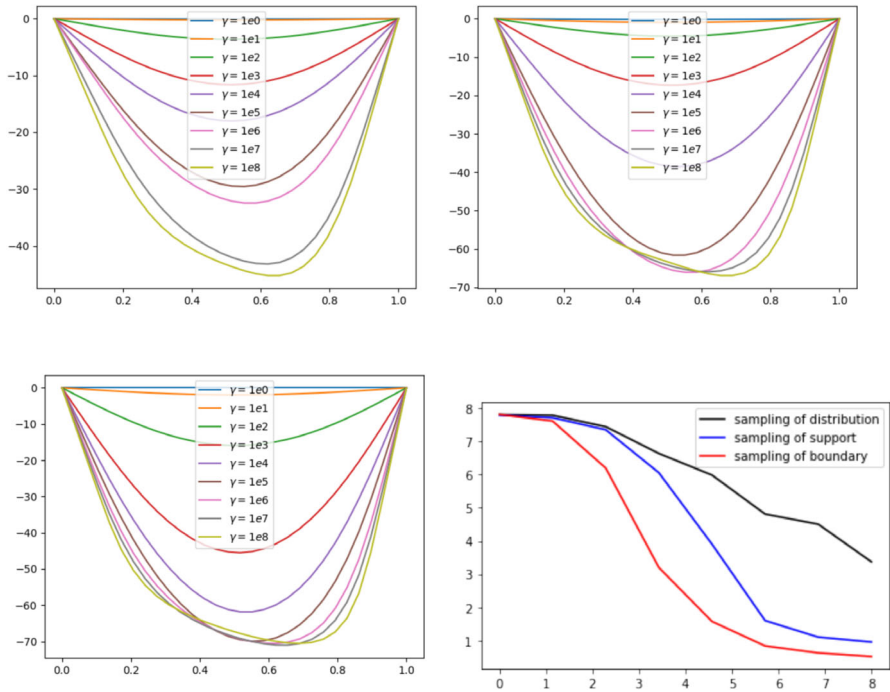
Fig. 4 Solutions u^γ obtained for increasing values $\gamma = 10^k$ and increasing sample sizes 3^k using sampling of the distribution (top left), sampling of the support (top right), and sampling on the boundary of the support (bottom left). The decrease of constraint violation as a function of k for the three sampling approaches is illustrated in the diagram at the bottom right

---

Algorithm 2 SAA Gradient descent with Moreau–Yosida penalization

---

Input: $D \subset \mathbb{R}^d$ (domain), K (number of penalty updates), $\{z_1, \ldots, z_{N_K}\}$ (set of samples), u_1 (initial control), $(\gamma_k)_{k \in \{1, \ldots, K\}}$ (sequence of penalties), $(t_j)_j$ (step-size rule)
for $k = 1, \ldots, K$ **do**
    for $j = 1, 2, \ldots$ **do**
        if $\|\nabla \hat{f}^{\gamma_k}(u_j)\| > \text{tol}$ **then**
            break
        end if
        for $\ell = 1, \ldots, N_k$ **do**
            $y_\ell \leftarrow$ solution to (27b)–(27c) with z_ℓ and u_j
            $p_\ell \leftarrow$ solution to (29) with y_ℓ and γ_k
        end for
        $u_{j+1} := u_j - t_j(\nabla F(u_j) - \frac{1}{N_k} \sum_{i=1}^{N_k} p_i)$
    end for
    $u_1 := u_j$
end for
Output: Solution u.

This is the case, for instance, in our problem. Therefore, the problem (22) with almost sure constraints is equivalent with the **robust** optimization problem

$$\min_{u \in L^2(D)} \quad F(u) \tag{30a}$$

$$\text{s.t.} \quad -\Delta \hat{y}(x, z) = u(x) + f(x, z), \quad (x, z) \in D \times \Xi, \tag{30b}$$

$$\hat{y}(x, z) = 0, \quad (x, z) \in \partial D \times \Xi, \tag{30c}$$

$$\hat{y}(x, z) \leq \alpha, (x, z) \in D \times \Xi, \tag{30d}$$

where the support Ξ of the random vector ξ plays the role of the uncertainty set with respect to which the worst case has to be taken into account. The parametrized solution $\hat{y} \colon D \times \Xi \to \mathbb{R}$ of the robust problem is distinguished from the solution $y \colon D \times \Omega \to \mathbb{R}$ introduced earlier. Since the support Ξ—contrary to our concrete simple ellipsoid $\mathcal{E}$— can be potentially complicated and hard to deal with, one approach might consist in *uniformly* sampling this support. Of course, our previous sampling of the distribution trivially provides samples on the support. These are, however, strongly concentrated around the mean 0 of the ellipsoid due to the underlying Gaussian distribution, while only few of these samples will be close to the boundary of the ellipsoid. However, points at the boundary are much more informative in defining the robust constraint (see discussion in Sect. 5.4 below). Therefore, it seems reasonable, if possible, to work with a uniform sample of the support or even just its boundary. In the case of our ellipsoid $\mathcal{E} := \{z \in \mathbb{R}^m \mid z^T \Sigma^{-1} z \leq 36\}$, a uniformly distributed sample can be created as follows: Let $\{v^{(k)}\}_{k=1}^K$ be a sample of the uniform distribution on the sphere $\mathbb{S}^{m-1}$ as in Sect. 4.1 and let $\{\tau^{(k)}\}_{k=1}^K$ be a sample of the uniform distribution on $[0, 1]$. Then, $\{6\tau^{(k)} \Sigma^{1/2} v^{(k)}\}_{k=1}^K$ is a sample of the uniform distribution on $\mathcal{E}$. When fixing $\tau^{(k)} \equiv 1$ instead, one rather obtains a uniform distribution on the boundary of $\mathcal{E}$. With this potentially more efficient uniform sampling of the support Ξ or its boundary, respectively, one may numerically proceed in the same way as with the previous sampling of the distribution. The corresponding results are illustrated in Fig. 4 (top right and bottom left). A plot showing the decrease in constraint violation for the three sampling strategies is illustrated in the diagram at the bottom right. More precisely, constraint violation refers to the maximum excess over the threshold α (with respect to the domain of the space variable x and to the support of the random variable ξ) of the random state under optimal control. All sampling schemes show convergence according to the theoretical results, but uniform sampling of the support and even more of just its boundary seem to exhibit much faster convergence than sampling of the distribution (see amplitudes of solutions in Fig. 4). This will be confirmed in the following section upon comparing results with the sample-free solution of the problem.

5.4 Sample-Free Solution of the Optimization Problem with Almost-Sure Constraints

So far, we have seen four alternative approaches for dealing with almost sure constraints: one by formulating a probabilistic constraint with probability level $p = 1$

and three sample average approaches based on Moreau–Yosida approximation with increasing penalty parameter. The latter three methods differed according to whether the distribution was sampled itself or rather its support or even just the boundary of the support. In this section, we present yet another alternative, namely the direct solution of the robust optimization problem (30) without sampling. This solution can be understood as the "true" solution of the almost sure problem. We emphasize that such a sample-free approach may not be possible in general, when the objective and the support are not simple enough to provide an analytical solution of the inner problem below.

We recall our assumption that the support $\Xi \subseteq \mathbb{R}^m$ of the random vector ξ is compact. Similar to the probabilistic setting, we may exploit the parametric control-to-state operator S there, in order to recast (30) as an infinite-dimensional optimization problem

$$\min_{u \in L^2(D)} F(u) \text{ s.t. } h(u) \leq 0, \quad h(u) := \max_{(x,z) \in \bar{D} \times \Xi} [S(u,z)](x) - \alpha. \quad (31)$$

In order to solve (31) numerically, we need to compute the function h and its (sub-) gradients. It is advantageous to rewrite h as an iterated maximum

$$h(u) = \max_{x \in \bar{D}} \max_{z \in \Xi} [S(u,z)](x) - \alpha. \quad (32)$$

Since S is affine linear in z, the inner maximization problem (for some fixed $u \in L^2(D)$ and $x \in \bar{D}$) can be solved analytically if the support Ξ is a simple set (e.g., rectangle, ellipsoid). Then, since the domain D is of low dimension, the outer maximization is easily carried out on a grid corresponding to the discretization of the state. For instance, if the support is given by an ellipsoid

$$\Xi = \{z \in \mathbb{R}^m \mid z^T B z \leq \gamma\}, \quad (33)$$

for some symmetric and positive definite matrix B and some $\gamma > 0$, then a linear function $c^T z$ realizes its maximum over Ξ at the point

$$z^* = \sqrt{\frac{\gamma}{c^T B^{-1} c}} B^{-1} c.$$

Now, in order to compare the sample-free robust solution with the previous three methodologies, assume that the support of Ξ of the truncated Gaussian random vector ξ is the ellipsoid $\mathcal{E} := \{z \in \mathbb{R}^m \mid z^T \Sigma^{-1} z \leq 36\}$ as in Sect. 4.2. For a fixed $u \in L^2(D)$, denote as in Theorem 3.3 by $\bar{y}_u := S(u,0)$ the mean state associated with u, i.e., the solution of the PDE (1c), (1d) with right-hand side $u(x) + f_0(x)$. Then, with (11), we infer that

$$[S(u,z)](x) = \bar{y}_{(u)}(x) + \sum_{i=1}^m z_i y^{(i)}(x) \quad \forall (x,z) \in \bar{D} \times \Xi.$$
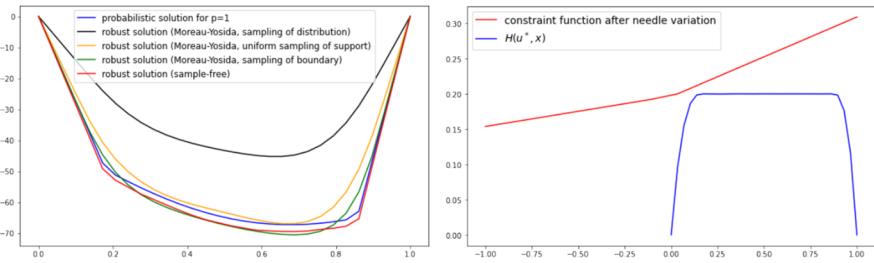
Fig. 5 Comparison of the probabilistic, three sampling-based the sample-free solutions to the worst-case problem (left). Illustration of the nonsmoothness of the robust constraint function h (right), details see text

Hence, for some fixed $u \in L^2(D)$ and $x \in \bar{D}$, the inner maximization problem mentioned above amounts to maximizing the linear function $c^T z$ with $c := (y^{(i)}(x))_{i=1}^m$ over the ellipsoid (33) with $B := \Sigma^{-1}$ and $\gamma := 36$, where the data for the support have been chosen as in the previous computations. Therefore, the maximum is realized for

$$z^* = \frac{6}{\sqrt{c^T \Sigma c}} \Sigma c.$$

Summarizing, the function value of h can be determined as

$$h(u) = \max_{x \in \bar{D}} \bar{y}_{(u)}(x) + \sum_{i=1}^m z_i^* y^{(i)}(x) - \alpha =: \max_{x \in \bar{D}} H(u, x) - \alpha,$$

which is easily approximated over the discretized state in low dimensions. The robust constraint function h will be typically nonsmooth (see discussion below). Its (convex) subdifferential is easily characterized as

$$\partial h(u) = \text{clco}\,\{\mathbf{u}_x \mid x \in M(u)\}; \quad M(u) := \{x \in \bar{D} \mid h(u) = H(u, x) - \alpha\},$$

where $\mathbf{u}_x$ is introduced below (13). Equipped with these tools, an appropriate algorithm for nonsmooth convex optimization can be employed in order to numerically solve problem (31). As this algorithmic aspect is not in the focus of our analysis here, we used for simplicity the previously mentioned SLSQP method without providing derivative information. Figure 5 (left) shows the comparison of the probabilistic solution in Fig. 1 (right) for $p = 1$ with the three sampled and the sample-free robust solutions on the same support. The sampled robust solutions are those illustrated in Fig. 4 with penalty parameter $\gamma = 10^8$ and sample size 3^8. Several conclusions can be drawn when understanding, as suggested above, the sample-free solution as the "true" reference: First, sampling the distribution is the least efficient method. This is not surprising, because in this sampling procedure a minimum amount of information is exploited. In particular, no knowledge about the support of the distribution is used. Sampling the support uniformly (and not according to the original random distribution, which might be strongly concentrated in the interior of the support) yields a much better

approximation. This can even be significantly improved when just sampling uniformly the boundary of the support. The explanation of this further improvement is that, thanks to the linearity of the objective in the inner maximization problem (32), all constraints are already induced by boundary points. This observation from Fig. 5 complements those from Fig. 4. Best performance in approximating the "true" solution is reached by the just-mentioned sampling of the boundary of the support combined with a Moreau–Yosida approximation on the one side and probabilistic constraint with level $p = 1$ on the other side. The difference between both results is that the former method better approximates the smooth pieces of the solution, whereas the latter one is more precise close to the nonsmooth kinks.

Remark 5.2 We note that the benefit of sampling the boundary of the support pertains to constraint functions $g(u, z)$ that are not necessarily linear but just convex in the random variable z. Then, the objective in the inner maximization of (32) is convex and, hence, assuming that the support Ξ is not just compact but also convex, the maximizers are still located on the boundary of Ξ.

It is interesting to observe that the robust constraint function h fails to be differentiable at the optimal solution u^* (here, we do not refer to the fact, evident from Fig. 5 (left), that this solution itself is a nonsmooth function). A first indication of this is that the function $H(u^*, \cdot)$ plotted in Fig. 5 (right) achieves its maximum over $\bar{D} = [0, 1]$ on a whole interval. Hence, the set $M(u^*)$ and the subdifferential $\partial h(u^*)$ do not shrink to a singleton and, consequently, h is nonsmooth. This fact can be alternatively illustrated by plotting h as a univariate function of a needle variation of u^* on a small interval $\Delta \subseteq D$ in the middle of the domain. More precisely, we define the function $\tilde{h}(t) := h(u_t)$, where

$$u_t(x) := \begin{cases} u^*(x) + t, & \text{if } x \in \Delta, \\ u^*(x), & \text{if } x \in D \setminus \Delta. \end{cases}$$

Evidently, if h was differentiable, then so would the univariate function $\tilde{h}$ be, too. This, however, is not the case as can be seen from Fig. 5 (right). Note that in the figure, we are simultaneously using the horizontal axis for the x-variable of the function $H(u^*, x)$ and for the t-variable of the function $\tilde{h}$.

A Appendix

In the following, we provide a proof of Theorem 3.2. It will follow from two lemmas below.

Lemma A.1 *Under the assumptions of Theorem 3.2, the mapping g defined there, satisfies the assumptions of Theorem 3.1.*

Proof We refer to the assumptions 1.-4. of Theorem 3.2. The continuity of h (see 1.) ensures that g can be written as a max rather than just sup and that $g(\bar{u}, 0) < 0$ by 3. Condition 2. implies that g is locally Lipschitzian. Indeed, the function $\gamma(u, z) :=$

$\max_{w \in K} \|D_{(u,z)}h(u, z, w)\|$ is continuous by 2., hence it is locally bounded around an arbitrary (u, z) by some constant C. This implies that all functions $h(\cdot, \cdot, w)$ for $w \in K$ are locally Lipschitz continuous around (u, z) with a common Lipschitz constant C. Hence, g is locally Lipschitz continuous around (u, z). The convexity of h in its second argument (see 1.) yields that g is convex in its second argument too. Therefore, our general assumption (GA) on g holds true. We show that 4. implies (4). To this aim, for $(u, d, z, t) \in U \times U \times \mathbb{R}^m \times \mathbb{R}$, denote by $w_{u,d,z,t} \in K$ some element with $g(u + td, z) = h(u + td, z, w_{u,d,z,t})$, which exists thanks to the compactness of K and the continuity of h. It follows that $g(u' + td, z) = h(u' + td, z, w_{u',d,z,t})$ and $-g(u', z) \leq -h(u', z, w_{u',d,z,t})$, which yields

$$
\begin{aligned}
g^\circ(\cdot, z)(u; d) &= \limsup_{u' \to u, t \downarrow 0} \frac{g(u' + td, z) - g(u', z)}{t} \\
&\leq \limsup_{u' \to u, t \downarrow 0} \frac{h(u' + td, z, w_{u',d,z,t}) - h(u', z, w_{u',d,z,t})}{t} \\
&\leq \limsup_{u'' \to u} \sup_{w \in K} \|D_u h(u'', z, w)\| \|d\| \\
&\leq c \exp(\|z\|) \|d\| \quad \forall u : \|u - \bar{u}\| \leq (2c)^{-1} \quad \forall z : \|z\| \geq c \quad \forall d \in U.
\end{aligned}
$$

Since, for some c_0,

$$
\exp(\|z\|) \leq \|z\|^{-m} \exp\left(\frac{\|z\|^2}{2\left\|\Sigma^{1/2}\right\|^2}\right) \quad \forall z : \|z\| \geq c_0,
$$

(4) holds true with $l := \max\{1, c_0, 2c\}$. $\qquad \square$

As a consequence of the previous lemma, Theorem 3.2 yields (7) via Theorem 3.1. The argument provided in [32] to infer (7) from the assumptions of Theorem 3.1, relies on the functions $e(\cdot, v)$ being uniformly (with respect to all $v \in \mathbb{S}^{m-1}$) Lipschitzian with a common modulus on a neighborhood of $\bar{u}$, see [32, Theorem 5, Corollary 2]. Since constants are integrable with respect to the uniform distribution on the sphere, one may invoke Clarke's theorem on subdifferentiation of integral functionals [14, Theorem 2.7.2] in order to derive (7). According to an addendum in Clarke's theorem, equality in (7) along with Clarke regularity of φ at $\bar{u}$ could be inferred if the partial functions $e(\cdot, v)$ were Clarke regular at $\bar{u}$. Unfortunately, they are not. On the other hand, with $e(\cdot, v)$ also the negative functions $-e(\cdot, v)$ are uniformly Lipschitzian with a common modulus on a neighborhood of $\bar{u}$. Hence, multiplying relation (5) by minus one, we may apply the same Clarke's theorem in order to derive the inclusion

$$
\partial^C(-\varphi)(\bar{u}) \subseteq \int_{\mathbb{S}^{m-1}} \partial_u^C(-e)(\bar{u}, v) \, d\mu_\zeta(v). \tag{34}
$$

Now, the addendum to Clarke's theorem mentioned above yields that equality in (34) and Clarke regularity of $-\varphi$ would hold true if the functions $-e(\cdot, v)$ were Clarke

regular at $\bar{u}$. Contrary to the functions $e(\cdot, v)$ themselves, this will hold true for $-e(\cdot, v)$, so that we can conclude equality instead of inclusion in (7) upon multiplying the relation (34) with minus one. Summarizing, the assertion of Theorem 3.2 will hold true as a consequence of the following

Lemma A.2 *Under the assumptions of Theorem 3.2, for each $v \in \mathbb{S}^{m-1}$, the functions $-e(\cdot, v)$ are Clarke regular at $\bar{u}$.*

Proof We define functions $e_w \colon U \times \mathbb{S}^{m-1} \to \mathbb{R}$ for $w \in K$ by $e_w(u, v) := \mu_\chi(A(u, v, w))$, where $A(u, v, w) := \{r \geq 0 \mid h(u, r\Sigma^{1/2}v, w) \leq 0\}$. By continuity of g and $g(\bar{u}, 0) < 0$, there exists a neighborhood $\mathcal{N}$ of $\bar{u}$ such that $g(u, 0) < 0$ or $h(u, 0, w) < 0$ for all $u \in \mathcal{N}$ and all $w \in K$. Then, the convexity of h in the second argument yields the existence of a positive (possibly infinite) number $\rho(u, v, w)$ such that $A(u, v, w) = [0, \rho(u, v, w)]$ for all $(u, v, w) \in \mathcal{N} \times \mathbb{S}^{m-1} \times K$. Now, the definitions of e and g provide that, for all $(u, v) \in \mathcal{N} \times \mathbb{S}^{m-1}$,

$$
e(u, v) = \mu_\chi \left(\bigcap_{w \in K} A(u, v, w) \right) = \mu_\chi \left(\bigcap_{w \in K} [0, \rho(u, v, w)] \right)
$$
$$
= \mu_\chi \left([0, \inf_{w \in K} \rho(u, v, w)] \right) = \inf_{w \in K} \mu_\chi([0, \rho(u, v, w)]) = \inf_{w \in K} e_w(u, v).
$$

Here, the penultimate identity follows from μ_χ being an absolutely continuous measure. Condition 4. of Theorem 3.2 ensures that $e_w(\cdot, v)$ is continuously differentiable on $\mathcal{N}$ for each $w \in K$ and each $v \in \mathbb{S}^{m-1}$ by [55, Corollary 3.2 and Example 3.1] and its derivative is given by (with ρ introduced above)

$$
D_u e_w(u, v) = \begin{cases} -f_\chi(\rho(u, v, w)) \dfrac{D_u h(u, \rho(u, v, w)\Sigma^{1/2}v, w)}{\langle D_z h(u, \rho(u, v, w)\Sigma^{1/2}v, w), \Sigma^{1/2}v \rangle} & \text{if } \rho(u, v, w) < \infty, \\ 0 & \text{if } \rho(u, v, w) = \infty. \end{cases}
$$
$$(35)$$

Here, f_χ denotes the density of the Chi-distribution μ_χ with m degrees of freedom. The Clarke regularity of $-e(\cdot, v) = \sup_{w \in K} -e_w(u, v)$ at $\bar{u}$ for every $v \in \mathbb{S}^{m-1}$ can be checked by applying [14, Theorem 2.8.2]. Indeed, according to this theorem and translated to our setting, it will be sufficient to show the following statements for each $v \in \mathbb{S}^{m-1}$ separately:

(a) The mapping $\mathcal{N} \times K \ni (u, w) \mapsto D_u(-e_w)(u, v)$ is continuous
(b) The functions $e_w(\cdot, v)$ with $w \in K$ are Lipschitz continuous on a neighborhood of $\bar{u}$ with some common modulus;
(c) The set $\{e_w(\bar{u}, v) \mid w \in K\}$ is bounded;
(d) The mapping $w \mapsto e_w(u, v) = \mu_\chi([0, \rho(u, v, w)])$ is continuous at each $u \in \mathcal{N}$.

In order to show (a), consider a sequence $(u_k, w_k) \to (u, w)$. On the one hand, if $\rho(u, v, w) < \infty$, we have that the continuity of ρ [56, Lemma 4.10 (ii)] and the continuity of $D_u h, D_z h$ by condition 2. of Theorem 3.2 imply via (35) that $D_u(-e_{w_k})(u_k, v) \to D_u(-e_w)(u, v)$. On the other hand, if $\rho(u, v, w) = \infty$, then

again by [56, Lemma 4.10 (ii)], we have that $\rho(u_k, v, w_k) \to \infty$. Now, adapting [55, Lemma 3.2] to our setting, we infer for k large enough the estimate

$$\|D_u e_{w_k}(u_k, v)\| \leq \frac{c}{|h(u, 0, w_k)|} f_\chi(\rho(u_k, v, w_k))\rho(u_k, v, w_k)$$
$$\exp\left(4\rho(u_k, v, w_k)\|\Sigma^{1/2}\|\right),$$

where c is the constant from condition 4. of Theorem 3.2. For $k \to \infty$, the expression $|h(u, 0, w_k)|$ tends to $|h(u, 0, w)|$, which is nonzero by definition of $\mathcal{N}$ (see above). Moreover, given the explicit formula for the density of the Chi-distribution with m degrees of freedom (including some normalizing constant $\tilde{c} > 0$), we get that

$$f_\chi(t)t \exp\left(4t\|\Sigma^{1/2}\|\right) = \tilde{c}t^m \exp(-t^2/2) \exp\left(4t\|\Sigma^{1/2}\|\right) \to 0 \quad (\text{as } t \to \infty).$$

Hence, by (35), $D_u e_w(u, v) = 0 = \lim_{k\to\infty} D_u e_{w_k}(u_k, v)$, which shows the continuity of $(u, w) \mapsto D_u(-e_w)(u, v)$. To proceed with items (b-d), observe that, from the already shown continuity of the mapping $(u, w) \mapsto D_u e_w(u, v)$, it follows that the function $u \mapsto \max_{w \in K} \|D_u e_w(u, v)\|$ is continuous on $\mathcal{N}$. Hence, there is a neighborhood $\mathcal{N}'$ of $\bar{u}$ and some $\kappa > 0$ such that $\|D_u e_w(u, v)\| \leq \kappa$ for all $w \in K$ and all $u \in \mathcal{N}'$. This implies (b) with the common Lipschitz modulus κ, while (c) is trivial because e_w is a probability. Concerning (d), fix an arbitrary $(u, v, w) \in \mathcal{N} \times \mathbb{S}^{m-1} \times K$ and a sequence $w_k \to w$ with $w_k \in K$. We repeat the case distinction on $\rho(u, v, w)$ from above: If $\rho(u, v, w) < \infty$, then $\rho(u, v, w_k) \to \rho(u, v, w)$ by the continuity of ρ already mentioned above. Moreover, $\rho(u, v, w_k) < \infty$ for k large enough and, hence, with F_χ denoting the continuous distribution function of the Chi-distribution with m degrees of freedom, we arrive at

$$e_{w_k}(u, v) = \mu_\chi([0, \rho(u, v, w_k)]) = F_\chi(\rho(u, v, w_k)) \to_k F_\chi(\rho(u, v, w))$$
$$= \mu_\chi([0, \rho(u, v, w)]) = e_w(u, v),$$

which is the desired continuity in the first case. Otherwise, if $\rho(u, v, w) = \infty$, then, as already seen above, we have that $\rho(u, v, w_k) \to \infty$. Let $\varepsilon > 0$ be arbitrary. Since $\lim_{t\to\infty} F_\chi(t) = 1$, it follows that $e_{w_k}(u, v) = F_\chi(\rho(u, v, w_k)) \geq 1 - \varepsilon$ whenever k is large enough and $\rho(u, v, w_k) < \infty$. If, in contrast, $\rho(u, v, w_k) = \infty$ for some k, then

$$e_{w_k}(u, v) = \mu_\chi([0, \infty)) = 1.$$

Hence, $e_{w_k}(u, v) \geq 1 - \varepsilon$ whenever k is large enough. Consequently,

$$e_{w_k}(u, v) \to_k 1 = \mu_\chi([0, \infty)) = \mu_\chi([0, \rho(u, v, w)]) = e_w(u, v).$$

This proves the desired continuity in the second case and, hence, (d). □

References

1. 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. SIAM/ASA J. Uncertain. Quantif. **5**(1), 1166–1192 (2017)

2. Ali, A.A., Ullmann, E., Hinze, M.: Multilevel Monte Carlo analysis for optimal control of elliptic PDEs with random coefficients. SIAM/ASA J. Uncertain. Quantif. **5**(1), 466–492 (2017)

3. Alla, A., Hinze, M., Kolvenbach, P., Lass, O., Ulbrich, S.: A certified model reduction approach for robust parameter optimization with PDE constraints. Adv. Comput. Math. **45**(3), 1221–1250 (2019)

4. Alnæs, M.S., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M.E., Wells, G.N.: The FEniCS project version 1.5. Arch. Numer. Softw. **3**(100), 9–23 (2015)

5. Alphonse, A., Geiersbach, C., Hintermüller, M., Surowiec, T. M.: Risk-averse optimal control of random elliptic variational inequalities. arXiv preprint arXiv:2210.03425 (2022)

6. Antil, H., Dolgov, S., Onwunta, A.: State-constrained optimization problems under uncertainty: a tensor train approach. arXiv preprint arXiv:2301.08684 (2023)

7. Benner, P., Onwunta, A., Stoll, M.: Block-diagonal preconditioning for optimal control problems constrained by PDEs with uncertain inputs. SIAM J. Matrix Anal. Appl. **37**(2), 491–518 (2016)

8. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer Science & Business Media, New York (2013)

9. Caillau, J.-B., Cerf, M., Sassi, A., Trélat, E., Zidani, H.: Solving chance constrained optimal control problems in aerospace via kernel density estimation. Optim. Control Appl. Methods **39**(5), 1833–1858 (2018)

10. Charnes, A., Cooper, W.W., Symonds, G.H.: Cost horizons and certainty equivalents: An approach to stochastic programming of heating oil. Manage. Sci. **4**(3), 235–263 (1958)

11. Chen, P., Ghattas, O.: Taylor approximation for chance constrained optimization problems governed by partial differential equations with high-dimensional random parameters. SIAM/ASA J. Uncertain. Quantif. **9**(4), 1381–1410 (2021)

12. 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)

13. Chen, P., Quarteroni, A., Rozza, G.: Multilevel and weighted reduced basis method for stochastic optimal control problems constrained by Stokes equations. Numer. Math. **133**(1), 67–102 (2016)

14. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley, New York (1983)

15. Conti, S., Held, H., Pach, M., Rumpf, M., Schultz, R.: Shape optimization under uncertainty-a stochastic programming perspective. SIAM J. Optim. **19**(4), 1610–1632 (2009)

16. Conti, S., Held, H., Pach, M., Rumpf, M., Schultz, R.: Risk averse shape optimization. SIAM J. Optim. **49**(3), 927–947 (2011)

17. Fang, K., Kotz, S.: Symmetric multivariate and related distributions. Chapman and Hall, London (1990)

18. 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**, 821–841 (2018)

19. Farshbaf-Shaker, M.H., Gugat, M., Heitsch, H., Henrion, R.: Optimal Neumann boundary control of a vibrating string with uncertain initial data and probabilistic terminal constraints. SIAM J. Optim. **58**(4), 2288–2311 (2020)

20. Gahururu, D.B., Hintermüller, M., Surowiec, T.M.: Risk-neutral PDE-constrained generalized Nash equilibrium problems. Math. Program. **198**(2), 1287–1337 (2023)

21. Garreis, S., Ulbrich, M.: Constrained optimization with low-rank tensors and applications to parametric problems with PDEs. SIAM J. Sci. Comput. **39**(1), A25–A54 (2017)

22. Geiersbach, C., Henrion, R.: Optimality conditions in control problems with random state constraints in probabilistic or almost sure form. Math. Oper. Res. (2024). https://doi.org/10.1287/moor.2023.0177

23. Geiersbach, C., Hintermüller, M.: Optimality conditions and Moreau–Yosida regularization for almost sure state constraints. *ESAIM: COCV*, 28, 80 (2022)

24. Geiersbach, Caroline, Pflug, Georg Ch: Projected stochastic gradients for convex constrained problems in Hilbert spaces. SIAM J. Optim. **29**(3), 2079–2099 (2019)

25. Geiersbach, C., Scarinci, T.: A stochastic gradient method for a class of nonlinear PDE-constrained optimal control problems under uncertainty. J. Diff. Eq. **364**, 635–666 (2023)

26. Geiersbach, C., Wollner, W.: Optimality conditions for convex stochastic optimization problems in Banach spaces with almost sure state constraints. SIAM J. Optim. **4**(31), 2455–2480 (2021)

27. Geletu, A., Hoffmann, A., Schmidt, P., Li, P.: Chance constrained optimization of elliptic PDE systems with a smoothing convex approximation. ESAIM Control Optim. Calc. Var. **26**, 70 (2020)
28. Genz, A., Bretz, F.: Computation of Multivariate Normal and t Probabilities. Lecture Notes in Statistics, vol. 195. Springer-Verlag, New York, NY (2009)
29. Göttlich, S., Kolb, O., Lux, K.: Chance-constrained optimal inflow control in hyperbolic supply systems with uncertain demand. Optim. Control Appl. Methods **42**, 566–589 (2021)
30. Guth, P.A., Kaarnioja, V., Kuo, F.Y., Schillings, C., Sloan, I.H.: A quasi-Monte Carlo method for optimal control under uncertainty. SIAM/ASA J. Uncertain. Quantif. **9**(2), 354–383 (2021)
31. Haber, E., Chung, M., Herrmann, F.: An effective method for parameter estimation with PDE constraints with multiple right-hand sides. SIAM J. Optim. **22**(3), 739–757 (2012)
32. Hantoute, A., Henrion, R., Pérez-Aros, P.: Subdifferential characterization of probability functions under Gaussian distribution. Math. Program. **174**(1–2), 167–194 (2019)
33. Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications, vol. 88 of Pure and Applied Mathematics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, (1980)
34. Kolvenbach, P., Lass, O., Ulbrich, S.: An approach for robust PDE-constrained optimization with application to shape optimization of electrical engines and of dynamic elastic structures under uncertainty. Optim. Eng. **19**(3), 697–731 (2018)
35. 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)
36. 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)
37. Kouri, D., Surowiec, T.: Risk-averse optimal control of semilinear elliptic PDEs. ESAIM Control Optim. Calc. Var. **26**, 53 (2019)
38. Kouri, D.P., Staudigl, M., Surowiec, T.M.: A relaxation-based probabilistic approach for PDE-constrained optimization under uncertainty with pointwise state constraints. Comput. Optim. Appl. **85**(2), 441–478 (2023)
39. Kouri, D., Surowiec, T.M.: Existence and optimality conditions for risk-averse PDE-constrained optimization. SIAM-ASA J. Uncertain. Quantif. **6**(2), 787–815 (2018)
40. Martin, M., Krumscheid, S., Nobile, F.: Complexity analysis of stochastic gradient methods for PDE-constrained optimal control problems with uncertain parameters. *ESAIM: Math. Modell. Numer. Anal.*, 55(4), 1599–1633 (2021)
41. Matthieu, M., Fabio, N.: PDE-constrained optimal control problems with uncertain parameters using SAGA. SIAM/ASA J. Uncertain. Quantifi. **9**(3), 979–1012 (2021)
42. Milz, J.: Consistency of Monte Carlo estimators for risk-neutral PDE-constrained optimization. Appl. Math. Optim. **87**(3), 57 (2023)
43. Ouanes, N., González Grandón, T., Heitsch, H., Henrion, R.: Optimizing the economic dispatch of weakly-connected mini-grids under uncertainty using joint chance constraints. Preprint No. 3069 (2023) Weierstrass Institute Berlin
44. Pérez-Aros, P., Quiñinao, C., Tejo, M.: Control in probability for SDE models of growth population. Appl. Math. Optim. **86**(3), 44 (2022)
45. Prékopa, A.: Stochastic Programming. Kluwer Academic Publishers, Dordrecht (1995)
46. Rockafellar, R.T., Royset, J.O.: Engineering decisions under risk averseness. ASCE-ASME J. Risk Uncertain. Eng. Syst., Part A: Civ. Eng. **1**(2), 04015003 (2015)
47. Schuster, M., Strauch, E., Gugat, M., Lang, J.: Probabilistic constrained optimization on flow networks. Optim. Eng. **23**(2), 1–50 (2022)
48. Shapiro, A., Dentcheva, D., Ruszczyński, A.: Lectures on Stochastic Programming—Modeling and Theory, volume 28 of MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, (2021). Third edition
49. Teka, K., Geletu, A., Li, P.: Structural properties and convergence approach for chance-constrained optimization of boundary-value elliptic partial differential equation systems. In: Yang, B., Stipanovic, D. (eds.) Nonlinear Systems, chapter 7. IntechOpen, Rijeka (2023)
50. van Ackooij, W., Henrion, R.: Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions. SIAM J. Optim. **24**(4), 1864–1889 (2014)

51. van Ackooij, W., Pérez-Aros, P.: Generalized differentiation of probability functions: parameter dependent sets given by intersections of convex sets and complements of convex sets. Appl. Math. Optim. **85**(1), 2 (2022)

52. van Ackooij, W.: A discussion of probability functions and constraints from a variational perspective. Set-Valued Var. Anal **28**(4), 585–609 (2020)

53. van Ackooij, W., Henrion, R.: (Sub-)gradient formulae for probability functions of random inequality systems under Gaussian distribution. SIAM/ASA J. Uncertain. Quantif. **5**(1), 63–87 (2017)

54. van Ackooij, W., Henrion, R., Pérez-Aros, P.: Generalized gradients for probabilistic/robust (probust) constraints. Optimization **69**, 1451–1479 (2020)

55. van Ackooij, W., Henrion, R., Zidani, H.: Pontryagin's principle for some probabilistic control problems. Appl. Math. Optim. **90**(5), 1–36 (2024)

56. van Ackooij, W., Pérez-Aros, P.: Generalized differentiation of probability functions acting on an infinite system of constraints. SIAM J. Optim. **29**(3), 2179–2210 (2019)

57. Van Barel, A., Vandewalle, S.: Robust Optimization of PDEs with random coefficients using a multilevel monte Carlo method. SIAM/ASA J. Uncertain. Quantif. **7**(1), 174–202 (2019)

58. Zahr, M.J., Carlberg, K.T., Kouri, D.P.: An efficient, globally convergent Method for optimization under uncertainty using adaptive model reduction and sparse grids. SIAM/ASA J. Uncertain. Quantif. **7**(3), 877–912 (2019)

Authors and Affiliations

Caroline Geiersbach[1] · René Henrion[1] · Pedro Pérez-Aros[2]

✉ Pedro Pérez-Aros
   pperez@dim.uchile.cl

   Caroline Geiersbach
   geiersbach@wias-berlin.de

   René Henrion
   henrion@wias-berlin.de

[1] Weierstrass Institute, Mohrenstrasse 39, 10117 Berlin, Germany

[2] Departamento de Ingeniería Matemática and Centro de Modelamiento Matemático (CNRS UMI 2807), Universidad de Chile, Santiago, Chile