Set-Valued Var. Anal
https://doi.org/10.1007/s11228-017-0452-5

Properties of Chance Constraints in Infinite Dimensions with an Application to PDE Constrained Optimization

M. H. Farshbaf-Shaker[1] · R. Henrion[1] · D. Hömberg[1,2]

Received: 1 October 2016 / Accepted: 2 October 2017

Abstract Chance constraints represent a popular tool for finding decisions that enforce the satisfaction of random inequality systems in terms of probability. They are widely used in optimization problems subject to uncertain parameters as they arise in many engineering applications. Most structural results of chance constraints (e.g., closedness, convexity, Lipschitz continuity, differentiability etc.) have been formulated in finite dimensions. The aim of this paper is to generalize some of these well-known semi-continuity and convexity properties as well as a stability result to an infinite dimensional setting. The abstract results are applied to a simple PDE constrained control problem subject to (uniform) state chance constraints.

Keywords Chance constraints · Probabilistic constraints · PDE constrained optimization

Mathematics Subject Classifications (2010) 90C15 · 49J20

1 Introduction

Many mathematical and engineering applications contain some considerable amount of uncertainty in their input data, e.g., unknown model coefficients, forcing terms and boundary conditions. Partial differential equations with uncertain coefficients play a central role

This research was partially carried out in the framework of MATHEON supported by the Einstein Foundation Berlin within the ECMath project SE13 as well as within project B04 of the Sonderforschungsbereich / Transregio 154 *Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks* funded by Deutsche Forschungsgemeinschaft.

✉ R. Henrion
Rene.Henrion@wias-berlin.de

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

[2] Department of Mathematical Sciences, NTNU, Alfred Getz vei 1, 7491 Trondheim, Norway

Published online: 11 October 2017

and are efficient tools for modeling randomness and uncertainty for the corresponding physical phenomena. Recently there is a growing interest and meanwhile a large amount of research literature for such PDEs, see e.g. [7–9, 22, 23] and references therein. Moreover, optimal control problems of such uncertain systems are of great practical importance. We mention here the works [10, 18, 24] and references therein. We note that the analysis of PDE constrained optimization with uncertain data is still in its beginning, in particular when uncertainty enters state constraints. The appropriate approach depends critically on the nature of uncertainty. If no statistical information is available, uncertainty cannot be modeled as a stochastic parameter but could be rather treated in a worst case or robust sense (e.g., [32]). On the other hand, if a (usually multivariate) statistical distribution can be approximated for the uncertain parameter, then a robust approach could turn out to be unnecessarily conservative and methods from stochastic optimization are to be preferred.

In [13, 16], the authors consider the minimization of different risk functionals (expected excess and excess probability) in the context of shape optimization, where the uncertainty is supposed to have a discrete distribution (finite number of load scenarios). In [4] an excess probability functional has been considered for a continuous multivariate (Gaussian) distribution. Randomness in constraints can be delt with by imposing a so-called chance constraint. To illustrate this, consider a random state constraint

$$y(x,\boldsymbol{\omega}) \le \bar{y}(x) \quad \forall x \in \boldsymbol{D},$$

where x, y refer to space and state variables, respectively, $\boldsymbol{\omega}$ is a random event, $\boldsymbol{D}$ is a given domain and $\bar{y}$ a given upper bounding function for the state. The associated *joint state chance constraint* then reads as

$$\mathbb{P}(y(x,\boldsymbol{\omega}) \le \bar{y}(x) \quad \forall x \in \boldsymbol{D}) \ge p,$$

where $\mathbb{P}$ is a probability measure and $p \in [0, 1]$ is a safety level, typically chosen close to, but different from one. The chance constraint expresses the fact that the state should uniformly stay below the given upper bound with high probability. In a problem of optimal control, the state chance constraint transforms into a (nonlinear) control constraint, thus defining an optimization problem with decisions which are robust in the sense of probability. This probabilistic interpretation of constraints has made them a popular tool first of all in engineering sciences (e.g., hydro reservoir control, mechanics, telecommunications etc.). We note that the state chance constraint above could be equivalently formulated as a constraint for the excess probability

$$\mathbb{P}(\mathscr{C}(y,\boldsymbol{\omega}) \ge 0) \ge p$$

of the random cost function

$$\mathscr{C}(y,\boldsymbol{\omega}) := \sup_{x\in\boldsymbol{D}} \{y(x,\boldsymbol{\omega}) - \bar{y}(x)\},$$

thus making a link to the papers discussed before. Note, however, that $\mathscr{C}$ is nondifferentiable in this case.

A mathematical theory treating PDE constrained optimization in combination with chance constraints is still in its infancy. The aim of this paper is to generalize semicontinuity and convexity properties of chance constraints, well-known in finite-dimensional optimization/operations research, to a setting of control problems subject to (uniform) state chance constraints. Although optimization problems with chance constraints (under continuous multivariate distributions of the random parameter) are considered to be difficult already in the finite-dimensional world, there exist a lot of structural results on, for instance, convexity (e.g., [20, 27, 28]), or differentiability (e.g., [26, 33]). For a numerical treatment in the framework of nonlinear optimization methods, efficient gradient formulae for probability functions have turned out to be very useful in the case of Gaussian or Gaussian-like distributions (e.g., [5, 21]). A classical monograph containing many basic theoretical results and numerous applications of chance constraints is [29]. A more modern presentation of the theory can be found in [31].

The paper is organized as follows: In Section 2, we provide some basic results on weak sequential semi-continuity properties of probability functions and on convexity of chance constraints in an abstract framework. Section 3 presents a stability result for optimal values and solutions to optimizatin problems with chance constraints under perturbations of the random distribution. In Section 4, these results will be applied to a specific PDE constrained optimisation problem with random state constraints.

2 Continuity Properties of Probability Functions

We consider the following probability function

$$h(u) := \mathbb{P}\left(g\left(u, \xi, x\right) \geq 0 \quad \forall x \in C\right) \quad (u \in U). \tag{1}$$

Here, U is a Banach space, C is an arbitrary index set, $g : U \times \mathbb{R}^s \times C \to \mathbb{R}$ is some constraint mapping and ξ is an s-dimensional random vector living on some probability space $\left(\Omega, \mathscr{F}, \mathbb{P}\right)$. Probability functions of this type figure prominently in stochastic optimization problems either in the form of chance constraints $h(u) \geq p$ or as an objective in reliability maximization problems. We are going to provide conditions for weak sequential upper semicontinuity of h first and, by adding appropriate assumptions, for weak sequential lower semicontinuity next. Throughout the paper we shall make use of the abbreviations w.s.u.s. for 'weakly sequentially upper semicontinuous' and w.s.l.s. for 'weakly sequentially lower semicontinuous'.

Proposition 1 *In (1), assume that the $g(u, \cdot, x)$ are Borel measurable for all $u \in U$ and $x \in C$ and that the $g\left(\cdot, z, x\right)$ are weakly sequentially upper semicontinuous (w.s.u.s.) for all $x \in C$ and $z \in \mathbb{R}^s$. Then, h defined in (1) is w.s.u.s.*

Proof Defining

$$\tilde{g}\left(u, z\right) := \inf_{x \in C} g\left(u, z, x\right) \quad \left(u \in U,\ z \in \mathbb{R}^s\right), \tag{2}$$

Equation (1) can be equivalently described as $h(u) = \mathbb{P}\left(\tilde{g}\left(u, \xi\right) \geq 0\right)$. By assumption on g, the function $\tilde{g}$ is Borel measurable in its second and w.s.u.s. in its first argument. Now, the assertion follows from Lemma 2 (applied to $\tilde{g}$) in the Appendix. □

The simple analogue of the previous Proposition, providing weak sequential lower semicontinuity of h under the condition that all functions $g(\cdot, \cdot, x) \quad (x \in C)$ are weakly sequentially lower semicontinuous (w.s.l.s.) cannot hold true even in a one-dimensional setting, where $g : \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ is defined as

$$g(u, z, x) := u - z \quad \forall x \in C := \mathbb{R}$$

and the distribution of ξ is the Dirac measure in zero. Then, clearly, g is even continuous but the probability function satisfies

$$h(u) = \begin{cases} 0 & \text{if } u < 0 \\ 1 & \text{if } u \geq 0 \end{cases}.$$

Hence, it fails to be lower semicontinuous at $\bar{u} := 0$.

The following proposition provides some missing conditions ensuring the weak sequential lower semicontinuity of h:

Proposition 2 *In* (1)*, assume that*

1. *C is a compact subset of $\mathbb{R}^d$.*
2. *g is w.s.l.s. (as function of all three variables simultaneously).*

Then h is w.s.l.s. at all $u \in U$ satisfying

$$\mathbb{P}\left(\tilde{g}\left(u, z\right) = 0\right) = 0, \tag{3}$$

where $\tilde{g}$ is defined in (2)*.*

Proof We show first that $\tilde{g}$ is w.s.l.s. Indeed, fix an arbitrary $(\bar{u}, \bar{z}) \in U \times \mathbb{R}^s$ and consider an arbitrary weakly convergent sequence $(u_k, z_k) \rightharpoonup (\bar{u}, \bar{z})$ and a realizing subsequence such that

$$\lim_{l} \tilde{g}(u_{k_l}, z_{k_l}) = \liminf_{k \to \infty} \tilde{g}(u_k, z_k).$$

By our assumptions 1. and 2., the infimum in (2) is attained. Hence, there exists a sequence $x_l \in C$ such that $\tilde{g}(u_{k_l}, z_{k_l}) = g(u_{k_l}, z_{k_l}, x_l)$. By compactness of C, we may assume that $x_{l_\alpha} \to_\alpha \bar{x}$ for some subsequence and some $\bar{x} \in C$. Exploiting 2. once more, we arrive at

$$\begin{aligned} \liminf_{k \to \infty} \tilde{g}(u_k, z_k) &= \lim_{\alpha} \tilde{g}(u_{k_{l_\alpha}}, z_{k_{l_\alpha}}) = \lim_{\alpha} g(u_{k_{l_\alpha}}, z_{k_{l_\alpha}}, x_{l_\alpha}) \\ &= \liminf_{\alpha} g(u_{k_{l_\alpha}}, z_{k_{l_\alpha}}, x_{l_\alpha}) \geq g(\bar{u}, \bar{z}, \bar{x}) \geq \tilde{g}(\bar{u}, \bar{z}). \end{aligned}$$

Consequently, $\tilde{g}$ is w.s.l.s. in both variables simultaneously. In particular, it is Borel measurable in the second one and, so, one may invoke Lemma 2 (applied to $\tilde{g}$) in the Appendix in order to derive that h is w.s.l.s. at all $u \in U$ satisfying (3). □

Remark 1 The result of Proposition 2 can be maintained by using the following alternative assumptions:

1. C is a finite subset of $\mathbb{R}^d$
2. $g(\cdot, \cdot, x)$ is w.s.l.s. for all $x \in C$

Note, that here we have strengthened the first assumption in favor of weakening the second one. The reason, why this is possible, is that $\tilde{g}$ in the proof of Proposition 2 happens to be w.s.l.s. as a <u>finite</u> minimum of w.s.l.s. functions.

We observe the following easy to check sufficient condition for (3) to hold:

Proposition 3 *In the setting of Proposition 2 assume that*

1. *the $g(u, \cdot, x)$ are concave for all $u \in U$ and $x \in C$.*
2. *for each $u \in U$ there exists some $\bar{z} \in \mathbb{R}^s$ such that $g(u, \bar{z}, x) > 0$ for all $x \in C$.*
3. *ξ has a density.*

Then, (3) *holds true at all $u \in U$.*

Proof Fix an arbitrary $u \in U$. Observe first that, as a consequence of 1., $\tilde{g}(u, \cdot)$ is a concave function. The assumptions of Proposition 2 ensure that the infimum in (2) is attained. Hence, by 2., there exists some $\bar{z} \in \mathbb{R}^s$ such that $\tilde{g}(u, \bar{z}) > 0$. Both observations entail that the set

$$E := \{z \in \mathbb{R}^s \mid \tilde{g}(u, \bar{z}) = 0\}$$

is a subset of the boundary of the convex set

$$\{z \in \mathbb{R}^s \mid \tilde{g}(u, \bar{z}) \geq 0\}.$$

Since the boundary of a convex set has Lebesgue measure zero, E itself has Lebesgue measure zero. By 3., the distribution of ξ is absolutely continuous with respect to the Lebesgue measure, whence $\mathbb{P}(\xi \in E) = 0$. This finally yields (3). □

We next address the question of convexity for a chance constraint $h(u) \geq p$ for h introduced in (1). To this aim, we recall that a function $\varphi : V \to \mathbb{R}$ (V a vectors space) is defined to be quasiconcave, if the following relation holds true:

$$\varphi(\lambda u + (1-\lambda)v) \geq \min\{\varphi(u), \varphi(v)\} \quad \forall u, v \in V; \quad \forall \lambda \in [0, 1]$$

The next proposition can be proven exactly in the same way as in [29, Theorem 10.2.1]. As this original proof has been given in an unnecessarily restricted setting (U finite dimensional, C a finite index set), we provide here a streamlined proof applicable to our setting in (1) for the readers convenience.

Proposition 4 *Let U be an arbitrary vector space and C be an arbitrary index set. Let the s-dimensional random vector ξ have a log-concave density (i.e., a density whose logarithm is a possibly extended-valued concave function). Finally, assume that the $g(\cdot, \cdot, x)$ are quasiconcave for all $x \in C$. Then, the set*

$$M := \{u \in U \mid h(u) \geq p\} \tag{4}$$

is convex for any $p \in [0, 1]$, where h refers to (1).

Proof Recall that, for $\tilde{g}$ defined in (2), we may write

$$h(u) = \mathbb{P}(\tilde{g}(u, \xi) \geq 0) \quad (u \in U). \tag{5}$$

We note that $\tilde{g}$ is quasiconcave. Indeed, fix an arbitrary pair of points

$$(u^1, z^1), (u^2, z^2) \in U \times \mathbb{R}^s$$

along with an arbitrary $\lambda \in [0, 1]$. Moreover, choose an arbitrary $\varepsilon > 0$. Then, there exists some $x \in C$ such that

$$\begin{aligned} \tilde{g}(\lambda(u^1, z^1) + (1-\lambda)(u^2, z^2)) &\geq g(\lambda(u^1, z^1) + (1-\lambda)(u^2, z^2), x) - \varepsilon \\ &\geq \min\{g(u^1, z^1, x), g(u^2, z^2, x)\} - \varepsilon \\ &\geq \min\{\tilde{g}(u^1, z^1), \tilde{g}(u^2, z^2)\} - \varepsilon. \end{aligned}$$

Here, in the second inequality, we exploit the quasiconcavity assumption on $g(\cdot, \cdot, x)$ for all $x \in C$. As $\varepsilon > 0$ was arbitrarily chosen, the claimed quasiconcavity of $\tilde{g}$ follows. Next, the assumption on ξ having a logconcave density implies by Prekopa's Theorem [29, Theorem 4.2.1] that ξ has a logconcave distribution. This means that

$$\mathbb{P}(\xi \in \lambda A + (1-\lambda)B) \geq [\mathbb{P}(\xi \in A)]^{\lambda}[\mathbb{P}(\xi \in B)]^{1-\lambda} \tag{6}$$

holds true for all convex subsets $A, B \in \mathbb{R}^s$ and all $\lambda \in [0, 1]$. In order to prove the claimed convexity of the set M in (4), let $u^1, u^2 \in M$ and $\lambda \in [0, 1]$ be arbitrarily given. Accordingly, $h(u^1), h(u^2) \geq p$. We have to show that $\lambda u^1 + (1-\lambda)u^2 \in M$. To this aim, define a multifunction $H : U \rightrightarrows \mathbb{R}^s$ by

$$H(u) := \{z \in \mathbb{R}^s \mid \tilde{g}(u, z) \geq 0\} \quad (u \in U).$$

Observe that $H(u^1)$ and $H(u^2)$ are convex sets as an immediate consequence of the quasiconcavity of $\tilde{g}$. We claim that

$$H(\lambda u^1 + (1-\lambda)u^2) \supseteq \lambda H(u^1) + (1-\lambda)H(u^2). \tag{7}$$

Indeed, selecting an arbitrary $z \in \lambda H(u^1) + (1-\lambda)H(u^2)$, we may find $z^1 \in H(u^1)$ and $z^2 \in H(u^2)$ such that $z = \lambda z^1 + (1-\lambda)z^2$. In particular,

$$\tilde{g}(u^1, z^1), \tilde{g}(u^2, z^2) \geq 0.$$

Exploiting the quasiconcavity of $\tilde{g}$ proven above, we arrive at

$$\begin{aligned}\tilde{g}(\lambda u^1 + (1-\lambda)u^2), z) &= \tilde{g}(\lambda(u^1, z^1) + (1-\lambda)(u^2, z^2)) \\ &\geq \min\{\tilde{g}(u^1, z^1), \tilde{g}(u^2, z^2)\} \geq 0.\end{aligned}$$

In other words, $z \in H(\lambda u^1 + (1-\lambda)u^2)$, which proves (7). Now, (5) along with (6) yields that

$$\begin{aligned}h(\lambda u^1 + (1-\lambda)u^2) &= \mathbb{P}(\xi \in H(\lambda u^1 + (1-\lambda)u^2)) \\ &\geq \mathbb{P}(\xi \in \lambda H(u^1) + (1-\lambda)H(u^2)) \\ &\geq [\mathbb{P}(\xi \in H(u^1))]^\lambda [\mathbb{P}(\xi \in H(u^2))]^{1-\lambda} \\ &= h^\lambda(u^1)h^{1-\lambda}(u^2) \geq p^\lambda p^{1-\lambda} = p.\end{aligned}$$

Consequently, $\lambda u^1 + (1-\lambda)u^2 \in M$ as desired. □

We note that in the previous convexity result the assumption of a log-concave density could be relaxed in the sense of generalized concavity properties (r-concavity), see [29]. We restrict ourselves here to log-concavity for the sake of simplicity and observe that many prominent multivariate distributions (including the Gaussian one) have a log-concave density.

3 A Stability Result for Chance Constrained Optimization Problems

In this section we establish a stability results for optimal solutions and optimal values of a chance constrained optimization problem in Banach spaces under perturbations of the distribution of the random vector. In this way, corresponding earlier finite-dimensional results in [19, 30] are substantially extended.

We consider the following (nominal) optimization problem with chance constraint:

$$\min_{u \in U_0} \{f(u) | \mathbb{P}(g(u, \xi, x) \geq 0 \quad \forall x \in C) \geq p\}. \tag{8}$$

Here, U_0 is a subset of the Banach space U and $f : U \to \mathbb{R}$ is some objective function, while g, ξ and C are as in (1). The scalar $p \in [0, 1]$ denotes a probability or safety level at which the random inequality system is supposed to be satisfied. We recall the set-valued mapping $H : U \rightrightarrows \mathbb{R}^s$ already introduced in the proof of Proposition 4 and defined by

$$H(u) := \{z \in \mathbb{R}^s \mid g(u, z, x) \geq 0 \quad \forall x \in C\} \quad (u \in U).$$

By $\mu := \mathbb{P} \circ \xi^{-1}$ we denote the distribution (the law) of our random vector ξ. Cleary, μ is the probability measure on $\mathbb{R}^s$ induced by ξ. By definition,

$$\mu\left(H\left(u\right)\right) = \mathbb{P}\left(g\left(u, \xi, x\right) \geq 0 \quad \forall x \in C\right) = h(u) \quad \left(u \in U\right), \tag{9}$$

where h is defined in (1). Then, problem (8) can be rewritten as

$$\min_{u \in U_0} \left\{f\left(u\right) | \mu\left(H\left(u\right)\right) \geq p\right\}. \tag{10}$$

The solution of this problem requires the distribution μ of the random vector ξ to be known. This, however, is rarely the case in practice and, more typically, one replaces the unknown μ by some approximating probability measure ν whose construction may be based on historical observations of ξ. This fact leads us to embed the nominal problem (10) into a family of optimization problems parameterized by the family $\mathscr{P}\left(\mathbb{R}^s\right)$ of all probability measures on $\mathbb{R}^s$:

$$\min\left\{f\left(u\right) | u \in \Phi\left(\nu\right)\right\} \quad \left(\nu \in \mathscr{P}\left(\mathbb{R}^s\right)\right). \tag{11}$$

Here, $\Phi : \mathscr{P}\left(\mathbb{R}^s\right) \rightrightarrows U$ is a mutlifunction representing the constraint set and being defined as

$$\Phi\left(\nu\right) := \left\{u \in U_0 | \nu\left(H(u)\right) \geq p\right\}.$$

Clearly, for $\nu = \mu$, problem (11) reduces to the nominal problem (10). In general, however, ν will be different from μ and so, the solution of (11) will differ from the theoretical solution of (10). Then, it comes as a natural question, under what conditions solutions and optimal values will behave in a stable way when perturbing the nominal measure μ. Does closeness of ν to μ (for instance, thanks to a large historical data base) imply closeness of solutions and optimal values of (11) to those of (10). In order to answer this question, we have to define first closeness of probability measures. To this aim, we introduce a so-called discrepancy distance:

$$\alpha\left(\nu_1, \nu_2\right) := \max\left\{\sup_{u \in U_0} \left|\nu_1\left(H(u)\right) - \nu_2\left(H(u)\right)\right|, \sup_{z \in \mathbb{R}^s} \left|\nu_1\left(z + \mathbb{R}^s_-\right) - \nu_2\left(z + \mathbb{R}^s_-\right)\right|\right\}$$
$$\left(\nu_1, \nu_2 \in \mathscr{P}\left(\mathbb{R}^s\right)\right) \tag{12}$$

We note that α is a metric on $\mathscr{P}\left(\mathbb{R}^s\right)$ by comparing ν_1 and ν_2 on all 'cells' $z + \mathbb{R}^s_-$. To emphasize this fact, we will write $\left(\mathscr{P}\left(\mathbb{R}^s\right), \alpha\right)$ for this metric space of probability measures. Moreover, by comparing ν_1 and ν_2 on all sets $H(u)$ with $u \in U_0$, this metric will turn out to be a suitable one for our stability analysis. Finally, we introduce the optimal value function $\phi : \mathscr{P}\left(\mathbb{R}^s\right) \to \mathbb{R}$ as well as the optimal solution mapping $\Psi : \mathscr{P}\left(\mathbb{R}^s\right) \rightrightarrows U$ for the parametric problem (11) as:

$$\phi\left(\nu\right) := \inf\left\{f\left(u\right) | u \in \Phi\left(\nu\right)\right\}; \quad \Psi\left(\nu\right) := \left\{u \in \Phi\left(\nu\right) | f\left(u\right) = \phi\left(\nu\right)\right\}. \tag{13}$$

Theorem 1 *In* (8), *let U be a reflexive Banach space and C an arbitrary index set. Let* $p \in (0, 1)$. *Assume the following conditions:*

1. *ξ has a log-concave density.*
2. *The $g(\cdot, \cdot, x)$ are quasiconcave for all $x \in C$.*
3. *The $g\left(\cdot, z, x\right)$ are w.s.u.s. for all $x \in C$ and $z \in \mathbb{R}^s$.*
4. *U_0 is bounded, closed and convex.*
5. *There exists some $\hat{u} \in U_0$ such that $\mathbb{P}\left(g\left(\hat{u}, \xi, x\right) \geq 0 \quad \forall x \in C\right) > p$*
6. *f is w.s.l.s.*

Then, there exists some $\varepsilon > 0$ *such that, with* μ *referring to the probability distribution of* ξ*,*

$$\Psi(\nu) \neq \emptyset \quad \forall \nu \in \mathscr{P}\left(\mathbb{R}^s\right) : \alpha(\mu, \nu) < \varepsilon. \tag{14}$$

Moreover, ϕ *is lower semicontinuous at* μ*. If additionally* f *is w.s.u.s., then* ϕ *is upper semicontinuous at* μ*. In other words, if* f *is weakly sequentially continuous, then* ϕ *is continuous at* μ*. Moreover, in this case,* Ψ *is weakly upper semicontinuous at* μ*, i.e., for every weakly open set* V *in* U *such that* $\Psi(\mu) \subseteq V$*, there exists some* $\varepsilon > 0$ *such that*

$$\Psi(\nu) \subseteq V \quad \forall \nu \in \mathscr{P}\left(\mathbb{R}^s\right) : \alpha(\mu, \nu) < \varepsilon. \tag{15}$$

Proof Define the multifunction $\tilde{M} : \mathbb{R} \rightrightarrows U$ by

$$\tilde{M}(t) := \{u \in U_0 | \mu(H(u)) \geq t\} \quad (t \in \mathbb{R}).$$

For every $\nu \in \mathscr{P}(\mathbb{R}^s)$ one has by (12) the implication

$$u \in \tilde{M}(p+\alpha(\mu, \nu)) \implies u \in U_0,\ \mu(H(u)) \geq p+\alpha(\mu, \nu) \geq p+\mu(H(u))-\nu(H(u)),$$

which entails the inclusion

$$\tilde{M}(p + \alpha(\mu, \nu)) \subseteq \Phi(\nu).$$

Taking this into account and recalling (9), Lemma 3 in the Appendix allows us to prove the existence of some $\varepsilon, \gamma > 0$ such that for all $u \in U_0$ and all $\nu \in \mathscr{P}(\mathbb{R}^s)$ with $\alpha(\mu, \nu) < \varepsilon$:

$$\begin{aligned} d(u, \Phi(\nu)) &\leq d(u, \tilde{M}(p + \alpha(\mu, \nu))) = d(u, \{u \in U_0 | \mu(H(u)) \geq p + \alpha(\mu, \nu)\}) \\ &= d(u, \{u \in U_0 | h(u) \geq p + \alpha(\mu, \nu)\} \\ &\leq \gamma \max\{\log(p + \alpha(\mu, \nu)) - \log h(u), 0\}. \end{aligned}$$

Assume now that $u \in \Phi(\mu)$ is arbitrarily given. In particular, $u \in U_0$ and $h(u) = \mu(H(u)) \geq p$. Exploiting the general relation $\log(c + d) - \log c \leq d/c$ for $c, d > 0$, we may continue the estimation above as

$$\begin{aligned} d(u, \Phi(\nu)) &\leq \gamma \max\{\log(\mu(H(u)) + \alpha(\mu, \nu)) - \log \mu(H(u)), 0\} \\ &\leq \gamma \max\{\alpha(\mu, \nu) / \mu(H(u)), 0\} \leq \gamma \max\{\alpha(\mu, \nu) / p, 0\} \\ &= L\alpha(\mu, \nu) \quad \forall u \in \Phi(\mu)\ \forall \nu \in \mathscr{P}\left(\mathbb{R}^s\right) : \alpha(\mu, \nu) < \varepsilon, \end{aligned} \tag{16}$$

for $L := \gamma/p$. By 5., we have that $\mu\left(H(\hat{u})\right) > p$, whence $\hat{u} \in \Phi(\mu)$. According to (16), we know that

$$d(\hat{u}, \Phi(\nu)) \leq L\alpha(\mu, \nu) \leq L\varepsilon < \infty \quad \forall \nu \in \mathscr{P}\left(\mathbb{R}^s\right) : \alpha(\mu, \nu) < \varepsilon.$$

In particular, the sets $\Phi(\nu)$ are nonempty and they are also weakly sequentially compact by Lemma 4 in the Appendix. As a consequence of 6., f attains its minimum over $\Phi(\nu)$. This proves (14).

Let now $\nu_n \in \mathscr{P}(\mathbb{R}^s)$ be a sequence with $\alpha(\mu, \nu_n) \to 0$ and

$$\liminf_{\alpha(\mu,\nu)\to 0} \phi(\nu) = \lim_{n\to\infty} \phi(\nu_n).$$

By the already proven relation (14), we may choose an associated sequence

$$w^{\nu_n} \in \Psi(\nu_n) \subseteq \Phi(\nu_n) \subseteq U_0.$$

By definition, $f(w^{\nu_n}) = \phi(\nu_n)$ for all n. Since w^{ν_n} is a bounded sequence in a reflexive Banach space, there exist a weakly converging subsequence $w^{\nu_{n_k}} \rightharpoonup_k \bar{w}$. Since also

$\alpha\left(\mu,\nu_{n_k}\right)\rightarrow_k 0$, it follows from Lemma 5 in the Appendix that $\bar{w}\in\Phi(\mu)$. Consequently, exploiting 6., one arrives at

$$\liminf_{\alpha(\mu,\nu)\rightarrow 0}\phi(\nu)=\lim_{k\rightarrow\infty}f\left(w^{\nu_{n_k}}\right)\geq f(\bar{w})\geq\phi(\mu).$$

This proves the asserted lower semicontinuity of ϕ at μ.

Next, assume that f is w.s.u.s. and let $\nu_n\in\mathscr{P}(\mathbb{R}^s)$ be a sequence with $\alpha(\mu,\nu_n)\rightarrow 0$ and

$$\limsup_{\alpha(\mu,\nu)\rightarrow 0}\phi(\nu)=\lim_{n\rightarrow\infty}\phi(\nu_n).$$

According to the already proven relation (14), we may select some $u^*\in\Psi(\mu)\subseteq\Phi(\mu)$. Then, $\phi(\mu)=f(u^*)$ and, by (16), we have that

$$d(u^*,\Phi(\nu_n))\leq L\alpha(\mu,\nu_n)$$

for n large enough. Since we have already seen that the $\Phi(\nu)$ are nonempty, whenever $\alpha(\mu,\nu)<\varepsilon$, we may select elements $u^{\nu_n}\in\Phi(\nu_n)$ satisfying the relation

$$\left\|u^*-u^{\nu_n}\right\|\leq d(u^*,\Phi(\nu_n))+n^{-1}\leq L\alpha(\mu,\nu_n)+n^{-1}.$$

Consequently, $u^{\nu_n}\rightarrow u^*$. Moreover $\varphi(\nu_n)\leq f(u^{\nu_n})$ and we conclude from f being w.s.u.s. (actually upper semicontinuity of f in the strong topology would be sufficient here) that

$$\limsup_{\alpha(\mu,\nu)\rightarrow 0}\phi(\nu)\leq\limsup_{n\rightarrow\infty}f(u^{\nu_n})\leq f(u^*)=\phi(\mu).$$

This proves the asserted upper semicontinuity of ϕ at μ.

Hence, we have shown so far that weak sequential continuity of f implies continuity of ϕ at μ. Having this result in mind, we finally prove the weak upper semicontinuity of Ψ at μ. If this didn't hold true, then there would exist a weakly open set V in U such that $\Psi(\mu)\subseteq V$ as well as a sequence $\nu_n\in\mathscr{P}(\mathbb{R}^s)$ such that $\alpha(\mu,\nu_n)\rightarrow 0$ and a sequence $u^{\nu_n}\in\Psi(\nu)\backslash V$. In particular, $f(u^{\nu_n})=\phi(\nu_n)$ for all n. Since $\Psi(\nu)\subseteq\Phi(\nu)\subseteq U_0$, where U_0 is bounded, there exists a weakly convergent subsequence $u^{\nu_{n_k}}\rightharpoonup_k\bar{u}\notin V$. From $u^{\nu_{n_k}}\in\Phi\left(\nu_{n_k}\right)$ and $\alpha\left(\mu,\nu_{n_k}\right)\rightarrow_k 0$ we derive with the help of Lemma 5 that $\bar{u}\in\Phi(\mu)$. On the other hand, the weak sequential continuity of f and the already proven in this case continuity of ϕ at μ provide

$$f(\bar{u})\leftharpoonup_k f(u^{\nu_{n_k}})=\phi\left(\nu_{n_k}\right)\rightarrow_k\phi(\mu).$$

The relations $\bar{u}\in\Phi(\mu)$ and $f(\bar{u})=\phi(\mu)$ now lead to the contradiction $\bar{u}\in\Psi(\mu)\subseteq V$. □

The following Corollary provides an important consequence of the weak upper semicontinuity of the solution set mapping at the nominal distribution $\mu\in\mathscr{P}(\mathbb{R}^s)$:

Corollary 1 *In addition to assumptions 1. - 5. in Theorem 1, let the objective f in (8) be weakly sequentially continuous and convex. Moreover, let $u_n\in U$ and $\nu_n\in\mathscr{P}(\mathbb{R}^s)$ be sequences such that $u_n\in\Psi(\nu_n)$ and $\alpha(\mu,\nu_n)\rightarrow_n 0$. Then, each weakly convergent subsequence u_{n_k} of u_n has a weak limit in $\Psi(\mu)$, i.e., a weak limit which is a solution of the nominal problem (8).*

Proof By assumption, (8) is a convex optimization problem (i.e., it has a convex objective and a convex constraint set $\Phi(\mu)$ as a consequence of assumption 4. in Theorem 1 and of Proposition 4). Hence, the set $\Psi(\mu)$ of optimal solutions to (8) is convex. Now, if $u_{n_k}\rightharpoonup_k$

$\bar{u} \in U$ is a weakly convergent subsequence of u_n and if we assumed that $\bar{u} \notin \Psi(\mu)$, then by the Hahn-Banach Theorem, one could find $u^* \in U^*$ and $\gamma \in \mathbb{R}$, such that

$$\langle u^*, u \rangle < \gamma < \langle u^*, \bar{u} \rangle \quad \forall u \in \Psi(\mu).$$

Since Ψ is weakly upper semicontinuous at μ by Theorem 1, there exists $\varepsilon > 0$ such that

$$\Psi(\nu) \subseteq V \quad \forall \nu \in \mathscr{P}(\mathbb{R}^s) : \alpha(\mu, \nu) < \varepsilon.$$

for the weakly open set

$$V := \{u \in U | \langle u^*, u \rangle < \gamma\}.$$

It follows that $\langle u^*, u_{n_k} \rangle < \gamma$ for k sufficiently large. Hence, u_{n_k} is not contained in the weakly open set

$$\tilde{V} := \{u \in U | \langle u^*, u \rangle > \gamma\}$$

containing $\bar{u}$. This contradicts $u_{n_k} \rightharpoonup_k \bar{u}$. □

4 Example from PDE Constrained Optimization

Chance constraints arise in many important engineering applications, where PDEs play a crucial role. The framework developed in Section 2 is used to treat simple linear PDE constrained optimization subject to such chance constraint. The solutions of linear PDEs depend linearly and continuously on the given data and this fact guarantees the weak sequentially semicontinuity of the function g in the chance constraint, which makes the framework developed in Section 2 applicable. More precisely, we consider the following simple PDE:

$$\begin{aligned} -\nabla_x \cdot (\kappa(x) \nabla_x y(x, \boldsymbol{\omega})) &= r(x, \boldsymbol{\omega}), \quad (x, \boldsymbol{\omega}) \in D \times \Omega \\ n \cdot (\kappa(x) \nabla_x y(x, \boldsymbol{\omega})) + \alpha\, y(x, \boldsymbol{\omega}) &= u(x) \quad (x, \boldsymbol{\omega}) \in \partial D \times \Omega, \end{aligned} \tag{17}$$

where $D \subset \mathbb{R}^d$, $d = 2, 3$, $\alpha > 0$ and ∇_x is the gradient operator with index x indicating that the gradient has to be build with respect to the spatial variable $x \in D$. Moreover $\boldsymbol{\omega}$ is the stochastic variable, which belongs to a complete probability space denoted by $(\Omega, \mathscr{F}, P)$. Here Ω is the set of outcomes, $\mathscr{F} \subset 2^{\Omega}$ is the σ -algebra of events, and $\mathbb{P} : \mathscr{F} \to [0, 1]$ is a probability measure. In (17) the function denoted by u will play the role of a deterministic control variable (boundary control), whereas the function r indicates an uncertain source function. Such PDEs appear for instance in shape optimization with stochastic loadings, see e.g. [16], or in induction heating problems in semiconductor single crystal growth processes, see e.g. [14]. For problems arising in the context of crystal growth of semiconductor single crystals optimizing the temperature - the state of the system - within a desirable range is one of important goals. In [14] a stationary heat equation is considered with a source term caused by an induction process. There, such an induction process generated by time-harmonic electromagnetic fields can not be realized exactly and exhibits uncertainty which consequently results in a random temperature field.

Remark 2 In order to make this section self-contained, we collect some well-known results concerning the well-posedness of (17), see [7, 8, 11, 18, 25], and highlight properties which are important for the applicability of the results of Section 2 in Section 4.2. We note that with the framework presented in Section 2, we are not able to treat PDE constrained optimization with a chance constraint involving nonlinear source terms in the PDE or even the case, where the coefficient κ in (17) is a random field $\kappa(x, \boldsymbol{\omega})$.

To ensure well-posedness of (17), we follow the lines in [7, 8, 11, 18] and assume that

$$D \in C^{1,1}, \quad \kappa \in C^{0,1}(D) \quad \text{and } \exists \kappa_0 > 0 : \kappa_0 \le \kappa(x)\, \forall x \in D. \tag{18}$$

4.1 Well-posedness of (17)

Throughout this paper, we use standard notations (e.g., see [3]) for the Sobolev spaces $H^m(D)$ for each real number m with norms $\|\cdot\|_{H^m(D)}$. We denote the inner product on H^m by $(\cdot,\cdot)_{H^m}$ and c a generic constant whose value may change with the context. Let ξ be an $\mathbb{R}^s$-valued random variable in a probability space $(\Omega, \mathscr{F}, P)$. If $\xi \in L^1_P(\Omega)$, we define $\mathbb{E}\xi = \int_\Omega \xi(\boldsymbol{\omega})\, dP(\boldsymbol{\omega})$ as its expected value. We now define the stochastic Sobolev spaces

$$L^2(\Omega; H^m(D)) = \{v : D \times \Omega \to \mathbb{R} \mid \|v\|_{L^2(\Omega;H^m(D))} < \infty\},$$

where

$$\|v\|^2_{L^2(\Omega;H^m(D))} = \int_\Omega \|v\|^2_{H^m(D)} dP(\boldsymbol{\omega}) = \mathbb{E}\|v\|^2_{H^m(D)}.$$

Note that the stochastic Sobolev space $L^2(\Omega; H^m(D))$ is a Hilbert space with the inner product

$$(u, v)_{L^2(\Omega;H^m(D))} = \mathbb{E}\int_D \nabla u \cdot \nabla v\, dx.$$

For simplicity, we use the following notation:

$$\mathscr{H}^m(D) = L^2(\Omega; H^m(D)).$$

For instance,

$$\mathscr{L}^2(D) = L^2(\Omega; L^2(D))$$

and

$$\mathscr{H}^1(D) = \{v \in \mathscr{L}^2(D) \mid \mathbb{E}\|v\|^2_{H^1(D)} < \infty\}.$$

Moreover we define

$$\mathscr{B}(\bar{D}) = L^2(\Omega; B(\bar{D})),$$

where by $B(\bar{D})$ we denote the space of continuous functions on $\bar{D}$.

We now state the well-posedness for (17).

Proposition 5 *Let (18) be fulfilled. Then for every $(r, u) \in \mathscr{L}^2(D) \times H^{\frac{1}{2}}(\partial D)$ there exists a unique solution $y \in \mathscr{H}^2(D)$ of (17) in the sense*

$$\begin{aligned} &\mathbb{E}\left(\int_D \kappa(x)\, \nabla_x y(x, \boldsymbol{\omega}) \cdot \nabla_x \rho(x, \boldsymbol{\omega})\, dx + \alpha \int_{\partial D} y(x, \boldsymbol{\omega})\, \rho(x, \boldsymbol{\omega})\, ds\right) \\ &= \mathbb{E}\left(\int_D r(x, \boldsymbol{\omega})\, \rho(x, \boldsymbol{\omega})\, dx + \int_{\partial D} u(x)\, \rho(x, \boldsymbol{\omega})\, ds\right), \quad \forall \rho \in \mathscr{H}^1(D) \end{aligned} \tag{19}$$

Moreover, the mapping

$$Y : \mathscr{L}^2(D) \times H^{\frac{1}{2}}(\partial D) \to \mathscr{H}^2(D), \quad (r, u) \mapsto y := Y(r, u)$$

is linear and continuous, i.e.

$$\|y\|_{\mathscr{H}^2(D)} \le c\left(\|r\|_{\mathscr{L}^2(D)} + \|u\|_{H^{\frac{1}{2}}(\partial D)}\right). \tag{20}$$

Proof This a conseuqence of the Lax-Milgram Lemma. □

Remark 3 For $\dim(D) = 3$, we know that the continuous embedding $\mathscr{H}^2(D) \hookrightarrow \mathscr{B}(\bar{D})$ is fulfilled. Hence, the solution y from Proposition 5 belongs to $\mathscr{B}(\bar{D})$ and we further obtain

$$\|y\|_{\mathscr{B}(\bar{D})} \leq c \left(\|r\|_{\mathscr{L}^2(D)} + \|u\|_{H^{\frac{1}{2}}(\partial D)} \right). \tag{21}$$

4.2 Optimization Problem

In preparation of the PDE constrained optimization problem we make the following assumptions:

(O1) Let $U := H^{\frac{1}{2}}(\partial D)$, $\bar{y}(\cdot) \in B(\bar{D})$ and a subset $C \subseteq D$ of the domain be given.
(O2) The admissible set U_{ad} is a bounded, closed and convex subset of U.
(O3) The cost functional $L : \mathscr{H}^2(D) \times H^{\frac{1}{2}}(\partial D) \to \mathbb{R}$ is weakly sequentially lower semi-continuous and bounded from below by zero.

Now, our overall optimization problem reads as

$$(P) \qquad \begin{cases} \min & \mathbb{E}(L(y(x,\boldsymbol{\omega}), u(x))) \\ \text{over} & \mathscr{H}^2(D) \times U_{ad} \\ \text{s.t.} & (19) \text{ is satisfied} \\ & \mathbb{P}(\boldsymbol{\omega} \in \Omega \mid y(x,\boldsymbol{\omega}) \leq \bar{y}(x), \forall x \in C) \geq p, \quad p \in (0,1) \end{cases}$$

Remark 4 As indicated in the beginning of this section for problems arising in the context of crystal growth of semiconductor single crystals optimizing the temperature - the state of the system - within a desirable range is one important goal. In application this is an important issue since engineers are interested to prevent damage in semiconductor single crystals which are caused by high temperature distributions. But as one has to deal with uncertain time-harmonic electromagnetic fields. The temperature field is consequently random, too. In this case it is reasonable to request that the temperature as state variable stays with high probability in some prescribed domain.

4.3 Finite Sum Expansion

For the source function r in (17) we make the ansatz of a finite (truncated) sum expansion extensively used in the literature:

$$r(x,\boldsymbol{\omega}) := \sum_{k=1}^{s} \beta_k(x)\, \xi_k(\boldsymbol{\omega}), \tag{22}$$

which enables us to approximate the infinite dimensional stochastic field by a finite dimensional (s-dimensional) random variable. For a discussion of this ansatz, we refer to [17] or [4, Section 2.4]. With

$$\beta(x) := (\beta_1(x), \ldots, \beta_s(x))^T; \quad \xi(\boldsymbol{\omega}) := (\xi_1(\boldsymbol{\omega}), \ldots, \xi_s(\boldsymbol{\omega}))^T,$$

we define

$$\tilde{r}(x,\xi) := \beta(x) \cdot \xi(\boldsymbol{\omega}), \tag{23}$$

where ξ is is an $\mathbb{R}^s$-valued random variable. Using the solution operator Y and (23) we define

$$g : U \times \mathbb{R}^s \times D \to \mathbb{R}, \quad g(u, \xi, x) := \bar{y}(x) - Y(\tilde{r}(x, \xi), u(x)). \tag{24}$$

Lemma 1 *The function $g(\cdot, \cdot, x)$, defined in (24), is weakly sequentially continuous and quasiconcave for all $x \in D$.*

Proof Using Proposition 5, under the assumption (22), we obtain from (20) the estimate

$$\|y\|_{\mathscr{H}^2(D)} \leq c\left((\|\beta\|_{[L^2(D)]^s} \cdot \|\xi\|_{[L^2(\Omega)]^s}) + \|u\|_U\right). \tag{25}$$

which means that y is depending linearly and continuously on the data (ξ, u) for fixed $x \in D$. Linearity in combination with continuity provides weak sequential continuity and quasiconcavity. Consequently the assertions of the lemma immediately follow. □

4.4 Properties of the Reduced Problem

Defining the reduced cost functional by

$$f(u(\cdot)) := \mathbb{E}(L(Y(\tilde{r}(\cdot, \xi), u(\cdot)), u(\cdot))) \tag{26}$$

and using the definition

$$h(u) := \mathbb{P}(g(u, \xi, x) \geq 0, \forall x \in C), \tag{27}$$

with g, defined in (24), and ξ, defined in (23), the chance constraint in (P) can be formulated as

$$M := \{u \in U \mid h(u) \geq p\}. \tag{28}$$

Then the reduced optimal control problem reads as

$$(P) \qquad \min_{u \in U_{ad} \cap M} f(u). \tag{29}$$

The aim of the following Theorem is to establish the existence of a solution to (P).

Theorem 2 *Assume (O1)-(O3). Then, the problem (P) admits a solution.*

Proof By Lemma 1, the function $g(\cdot, \cdot, x)$ is weakly sequentially continuous for all $x \in D$. Then, Proposition 1 yields that h is weakly sequentially upper semicontinuous, whence M in (28) is weakly sequentially closed. Consequently, by (O2) $U_{ad} \cap M$ is weakly sequentially closed, too. Moreover, the reduced cost function f, defined in (26) as a composition of three operators $\mathbb{E}$, L and Y, is weakly sequentially lower semicontinuous. This is true, because $\mathbb{E}$ and Y are linear and continuous and $\mathbb{E}$ additionally monotonous. Hence, (O3) provides the desired property of f. Now, the existence of a solution to (P) follows by the direct method in the calculus of variations. □

In the previous theorem, one of the main ingredients in proving the existence result was to establish the weak sequential upper semicontinuity of the function h. This was done by using Lemma 1 and Proposition 1. In the following theorem we will refine this upper semicontinuity result to a semicontinuity result by additionally taking into account a lower semicontinuity property. The theorem will then ensure weak sequential continuity of the fuction h.

Theorem 3 *Let C be a finite subset of $\mathbb{R}^d$ and the random variable ξ, defined in* (23)*, have a density. Moreover, assume that for each $u \in U$ there exists some $\bar{z} \in \mathbb{R}^s$ such that*

$$Y(\tilde{r}(x, \bar{z}), u(x)) < \bar{y}(x) \qquad \forall x \in C. \tag{30}$$

Then the function h, defined in (27)*, is weakly sequentially continuous.*

Proof Using once again Lemma 1, it follows that $g(\cdot, \cdot, x)$ is weakly sequentially continuous for all $x \in C \subseteq D$. Then, h is w.s.u.s. by Proposition 1. Moreover, it is obvious that $g(u, \cdot, x)$ is linear for all $u \in U$ and $x \in C$, and consequently concave. This is assumption 1. in Proposition 3, while the existence of a density for ξ required here, corresponds to assumption 3. of the same Proposition. Finally, (30) translates by (24) to assumption 2. of Proposition 3. Now this Proposition guarantees via Remark 1 that h is w.l.s.c. $\square$

We observe that we could not derive the result of the last Theorem for general compact sets $C \subseteq D$ by referring to Proposition 2. The reason is that we are not able in Lemma 1 to establish the required weak sequential lower semicontinuity of g in all three variables simultaneaously. Therefore we benefit from the alternative result mentioned in Remark 1 and using weak sequential lower semicontinuity of g in the first two variables only. This, of course, comes at the price of reducing C to a finite set.

The condition given by (30) can be interpreted as a Slater's condition. It means that for every given control u there must exists a realization $\bar{z}$ of the random variable ξ such that the state y has to be uniformly strictly smaller than the given state $\bar{y}$. If this condition is not fulfilled then the upper limit function $\bar{y}$ was chosen too restrictively.

An instance for the use of Theorem 3 is the consideration of random state constraints in disjunctive form which would lead to the following state chance constraint:

$$\mathbb{P}(\boldsymbol{\omega} \in \Omega \mid \exists x \in C : y(x, \boldsymbol{\omega}) > \bar{y}(x)) \geq p.$$

Here, in contrast to the previous setting in problem (P) one is interested in the complementary situation, namely that with high probability the random state exceeds some given threshold at least somewhere on the domain. Turning this state chance constraint into a control constraint as before and using the functions g, h defined in (24) and (27), respectively, we arrive at the condition

$$\begin{aligned}\mathbb{P}(\boldsymbol{\omega} \in \Omega \mid \exists x \in C : y(x, \boldsymbol{\omega}) > \bar{y}(x)) &= \mathbb{P}(\boldsymbol{\omega} \in \Omega \mid \exists x \in C : g(u, \xi, x) < 0) \\ &= 1 - h(u) \geq p.\end{aligned}$$

So, instead of (28) the chance constraint would be defined by $M := \{u \mid h(u) \leq 1 - p\}$. In order to prove an existence result similar to that of Theorem 2, one would now need the weak sequential lower (rather than upper) semicontinuity of h. This would come as a consequence of Theorem 3.

In the following theorem we are going to establish a condition such that (P) becomes a convex optimization problem.

Theorem 4 *Assume (O1)-(O3). Let the random variable ξ, defined in* (23)*, have a density whose logarithm is a (possibly extended-valued) concave function. Moreover, assume that the objective function L is convex. Then problem (P) is a convex optimization problem.*

Proof The convexity of L and the linearity of the solution operator Y, see (25), yield that the mapping

$$u(\cdot) \mapsto L(Y(\tilde{r}(\cdot, \xi), u(\cdot)), u(\cdot))$$

is convex. Then by the linearity of the expectation $\mathbb{E}$, we obtain that $u \mapsto f(u)$ is convex. Moreover, Lemma 1(b) provides that $g(\cdot, \cdot, x)$ is quasiconcave for all $x \in C$. Then, it follows from Proposition 4 that M is convex. By assumption U_{ad} is convex and consequently the intersection $M \cap U_{ad}$ is convex, too. Hence, the assertion of the theorem follows. □

Remark 5 Numerous multivariate distributions have log-concave densities, e.g. normal distribution, Student's t-distribution, uniform distribution on compact and convex sets, see e.g. [29]. Hence, the assumption about the logconcave densities is fairly general. Often in PDE constrained optimization the objective functional L has the form $L(y, u) = L_1(y) + L_2(u)$ where L_1 and L_2 are separately convex and are defined by $L_1 : \mathscr{H}^2(D) \ni y \mapsto L_1(y) \in \mathbb{R}$ and $L_2 : U \ni u \mapsto L_2(u) \in \mathbb{R}$.

As indicated in Remark 4 one has to deal with uncertain time-harmonic electromagnetic fields which result in uncertain temperature fields. In practice the distribution of such uncertain time-harmonic electromagnetic fields are unknown and engineers work instead with some approximating probability measure whose construction in most cases is based on historical observations of the uncertain electromagnetic fields. Now the stability results in Section 3 guarantee stability of solutions and optimal values of (P) to those with approximating probability measure.

In preparation of a stability result for our optimization problem with respect to perturbations of the random distribution, we denote by $\mu := \mathbb{P} \circ \xi^{-1}$ the distribution of our random vector ξ, defined in (23). We adapt the notation of Section 3 to our concrete optimization problem (29). We define the multifunction

$$H(u) := \{z \in \mathbb{R}^s | Y(\tilde{r}(x, z), u(x)) \leq \bar{y}(x) \quad \forall x \in C\} \quad (u \in U).$$

Moreover, with each probability measure $\nu \in \mathscr{P}(\mathbb{R}^s)$ we associate the feasible set

$$\Phi(\nu) := \{u \in U_{ad} | \nu(H(u)) \geq p\}.$$

This allows us to embed our given problem (29) into a family of problems

$$(P_\nu) \qquad \min\{f(u) | u \in \Phi(\nu)\} \quad \left(\nu \in \mathscr{P}\left(\mathbb{R}^s\right)\right). \tag{31}$$

parameterized by all probability measures. We observe that for $\nu := \mu = \mathbb{P} \circ \xi^{-1}$ we recover our nominal problem (29) with the given distribution of the random vector ξ: Indeed, by definition,

$$\begin{aligned}\Phi(\mu) &= \{u \in U_{ad} | \mu(H(u)) \geq p\} = \{u \in U_{ad} | \mathbb{P}(\xi \in H(u)) \geq p\} \\ &= \{u \in U_{ad} | \mathbb{P}(Y(\tilde{r}(x, \xi), u(x)) \leq \bar{y}(x) \quad \forall x \in C) \geq p\} = U_{ad} \cap M.\end{aligned}$$

Consequently, problems (P_ν) can be considered as perturbations of the nominal problem and it is of interest, whether optimal solutions and optimal values of (P_ν) behave stable under small perturbations. Here, closeness between ν and μ will be measured by the discrepancy distance α introduced in (12). Finally, we recall the defintions of the parameter-dependent optimal value function ϕ and optimal solution mapping Ψ defined in (13) and associated with the family of problems (P_ν) in (31). Now, we have paved the way for the desired stability result:

Theorem 5 *Consider the optimization problem (P) introduced in* Section 4.2 *and corresponding to* (29)*. Assume (O1)-(O3) with an arbitrary index set C. Let the random variable ξ, defined in* (23)*, have a density whose logarithm is a (possibly extended-valued) concave function. Moreover, assume that there exists some $\hat{u} \in U_{ad}$ such that*

$$\mathbb{P}\left(Y(\tilde{r}(x,\xi), \hat{u}(x)) < \bar{y}(x) \quad \forall x \in C\right) > p \tag{32}$$

Then, there exists some $\varepsilon > 0$ such that, with μ referring to the probability distribution of ξ,

$$\Psi(\nu) \neq \emptyset \quad \forall \nu \in \mathscr{P}\left(\mathbb{R}^s\right) : \alpha(\mu, \nu) < \varepsilon. \tag{33}$$

Moreover, ϕ is lower semicontinuous at μ. If additionally the cost function L in our optimization problem (P) is weakly sequentially upper semicontinuous, then ϕ is upper semicontinuous at μ. In other words, if L is weakly sequentially continuous, then ϕ is continuous at μ. Moreover, in this case, Ψ is weakly upper semicontinuous at μ, i.e., for every weakly open set V in U such that $\Psi(\mu) \subseteq V$, there exists some $\varepsilon > 0$ such that

$$\Psi(\nu) \subseteq V \quad \forall \nu \in \mathscr{P}\left(\mathbb{R}^s\right) : \alpha(\mu, \nu) < \varepsilon. \tag{34}$$

Proof To prove Theorem 5 we have to check the assumptions of Theorem 1, where the first one is evident by being directly imposed here. Next, Lemma 1 guarantees the assumption 2. and 3. of Theorem 1. Defining U_0 as U_{ad}, the assumption (O2) provides assumption 4. in Theorem 1. Moreover, since $H^{\frac{1}{2}}(\partial\Omega)$ as a Hilbert space is reflexive and arguing as in the proof of Theorem 2 that the reduced cost function f, defined in (26) as a composition of three operators $\mathbb{E}$, L and Y, is weakly sequentially lower semicontinuous gives assumption 6. in Theorem 1. Clearly, (32) corresponds to assumption 5. in Theorem 1. Therefore, we get the first part of the assertion of our Theorem. Arguing as in the proof of Theorem 2 that the reduced cost function f, defined in (26) as a composition of three operators $\mathbb{E}$, L and Y, is weakly sequentially upper semicontinuous, the second part of our Theorem 5 is proved by the second part of Theorem 1. □

Corollary 2 *In addition to the assumptions in Theorem 5, let the objective L in (O3) be weakly sequentially continuous and convex. Moreover, let $u_n \in U$ and $\nu_n \in \mathscr{P}(\mathbb{R}^s)$ be sequences such that $u_n \in \Psi(\nu_n)$ and $\alpha(\mu, \nu_n) \to_n 0$. Then, each weakly convergent subsequence u_{n_k} of u_n has a weak limit in $\Psi(\mu)$. In other words: each weakly convergent subsequence of solutions to the approximating problems has a weak limit which is a solution of the nominal problem (P).*

Proof Arguing as in the proof of Theorem 2 that the reduced cost function f, defined in (26) as a composition of three operators $\mathbb{E}$, L and Y, is weakly sequentially semicontinuous, Corollary 2 provides the assertion. □

5 Conclusions

We have shown that and how certain basic structural and stability properties of chance constraints, which are well established in finite dimensions, can be carried over to and verified in a Banach space setting. So far, a simple class of PDE constrained optimization problems could be demonstrated to be a good candidate for ensuring these properties and, in particular, for deriving the existence of solutions and their stable dependence on perturbations of the underlying probability distribution. Figuring out more complex problem classes will be a major future challenge. Moreover, passing to more interesting structural properties like Lipschitz continuity and differentiability of the probability functions or convexity of the feasible set under less restrictive assumptions will be in the focus of future research. The ultimate goal of such analysis - which is nontrivial already in finite dimensions (see [1, 2, 20, 21]) - would be the efficient numerical solution of state chance constrained in PDE constrained optimization.

Acknowledgments The authors express their gratitude to two anonymous referees whose very careful reading and critical comments led to a substantially improved presentation of this paper.

Appendix

Lemma 2 *Let X be a Banach space and $g : X \times \mathbb{R}^m \to \mathbb{R}$, be Borel measurable in the second argument. Further, let ξ be an m-dimensional random vector defined on a probability space $\left(\Omega, \mathscr{A}, \mathbb{P}\right)$. Then, the probability function*

$$\varphi(x) := \mathbb{P}\left(g\left(x, \xi\right) \geq 0\right) \quad (x \in X)$$

is well-defined and if g is weakly sequentially upper semicontinuous (w.s.u.s.) in the first argument, then φ is w.s.u.s. too. If, conversely, g is weakly sequentially lower semicontinuous (w.s.l.s.) in the first argument, then φ is w.s.l.s. too in those arguments $\bar{x} \in X$ satisfying the relation

$$\mathbb{P}\left(g\left(\bar{x}, \xi\right) = 0\right) = 0. \tag{35}$$

Proof Observe first, that φ is well defined by Borel measurability of g in the second argument. Fix an arbitrary $\bar{x}$ and let $x_n \rightharpoonup \bar{x}$ be an arbitrary weakly convergent sequence. Denote by x_{n_l} a subsequence such that

$$\limsup_{n\to\infty} \varphi(x_n) = \lim_{l\to\infty} \varphi(x_{n_l}). \tag{36}$$

Define the sets

$$A := \{\omega \in \Omega | g\left(\bar{x}, \xi\left(\omega\right)\right) \geq 0\}; \quad A_n := \{\omega \in \Omega | g\left(x_n, \xi\left(\omega\right)\right) \geq 0\} \quad (n \in \mathbb{N}).$$

Then, by g being w.s.u.s. in the first argument, we have that

$$\limsup_{n\to\infty} g(x_n, \xi\left(\omega\right)) \leq g\left(\bar{x}, \xi\left(\omega\right)\right) < 0 \quad \forall \omega \in \Omega \backslash A.$$

Consequently, $g(x_n, \xi\left(\omega\right)) < 0$ for all $\omega \in \Omega\backslash A$ and all $n \geq n_0(\omega)$. Denoting by χ_C the characteristic function of a set C, this entails that $\chi_{A_n}\left(\omega\right) \to_{n\to\infty} 0$ for all $\omega \in \Omega\backslash A$. By the dominance convergence theorem,

$$\int_{\Omega\backslash A} \chi_{A_n}\left(\omega\right) \mathbb{P}\left(d\omega\right) \to_{n\to\infty} 0 \quad \forall \omega \in \Omega\backslash A.$$

On the other hand, $\chi_{A_n}(\omega) \leq \chi_A(\omega) = 1$ for $\omega \in A$, whence

$$\begin{aligned}
\lim_{l\to\infty} \varphi(x_{n_l}) &= \lim_{l\to\infty} \mathbb{P}\left(g\left(x_{n_l},\xi\right) \geq 0\right) = \lim_{l\to\infty} \int_\Omega \chi_{A_{n_l}}(\omega)\,\mathbb{P}(d\omega) \\
&= \lim_{l\to\infty} \left(\int_{\Omega\backslash A} \chi_{A_{n_l}}(\omega)\,\mathbb{P}(d\omega) + \int_A \chi_{A_{n_l}}(\omega)\,\mathbb{P}(d\omega) \right) \\
&\leq \limsup_{l\to\infty} \int_{\Omega\backslash A} \chi_{A_{n_l}}(\omega)\,\mathbb{P}(d\omega) + \limsup_{l\to\infty} \int_A \chi_{A_{n_l}}(\omega)\,\mathbb{P}(d\omega) \\
&= \limsup_{l\to\infty} \int_A \chi_{A_{n_l}}(\omega)\,\mathbb{P}(d\omega) \\
&\leq \limsup_{l\to\infty} \int_A \mathbb{P}(d\omega) = \mathbb{P}(A) = \mathbb{P}(g(\bar{x},\xi) \geq 0) \\
&= \varphi(\bar{x}).
\end{aligned}$$

Combining this with (36) yields that φ is w.s.u.s. in $\bar{x}$.

Next, let $\bar{x} \in X$ be arbitrary such that (35) is fulfilled. Let $x_n \rightharpoonup \bar{x}$ be an arbitrary weakly convergent sequence. Define the sets

$$A := \{\omega \in \Omega | g(\bar{x}, \xi(\omega)) > 0\}; \quad A_n := \{\omega \in \Omega | g(x_n, \xi(\omega)) \geq 0\} \quad (n \in \mathbb{N}).$$

Then, with g being w.s.l.s. in the first argument, we have that

$$\liminf_{n\to\infty} g(x_n, \xi(\omega)) \geq g(\bar{x}, \xi(\omega)) > 0 \quad \forall \omega \in A.$$

Hence, $\chi_{A_n}(\omega) \to_{n\to\infty} \chi_A(\omega) = 1$ for all $\omega \in A$, whereas $\chi_{A_n}(\omega) \geq \chi_A(\omega) = 0$ for all $\omega \in \Omega\backslash A$. Now, Fatou's Lemma combined with (35) yields that

$$\begin{aligned}
\liminf_{n\to\infty} \varphi(x_n) &= \liminf_{n\to\infty} \int_\Omega \chi_{A_n}(\omega)\,\mathbb{P}(d\omega) \geq \int_\Omega \liminf_{n\to\infty} \chi_{A_n}(\omega)\,\mathbb{P}(d\omega) \\
&\geq \int_A \chi_A(\omega)\,\mathbb{P}(d\omega) = \mathbb{P}(g(\bar{x},\xi) > 0) \\
&= \mathbb{P}(g(\bar{x},\xi) \geq 0) = \varphi(\bar{x}).
\end{aligned}$$

Hence, φ is w.s.l.s. in $\bar{x} \in X$. □

Lemma 3 *Under the assumptions of Theorem 1, there are constants $\varepsilon, \gamma > 0$ such that (with d referring to the point-to-set distance)*

$$d(u, \{u \in U_0 | h(u) \geq \tau\}) \leq \gamma \max\{\log \tau - \log h(u), 0\} \quad \forall u \in U_0\ \forall \tau \in [p-\varepsilon, p+\varepsilon].$$

Proof By definition of h in (1), the inequality $\mu(H(u)) \geq p$ is equivalent with $h(u) \geq p$. At the end of the proof of Proposition 4 (which to invoke is justified by 1. and 2. in Theorem 1), we have shown that, for arbitrary $u^1, u^2 \in U$ and $\lambda \in [0, 1]$ the inequality

$$h(\lambda u^1 + (1-\lambda)u^2) \geq h^\lambda(u^1) h^{1-\lambda}(u^2)$$

holds true. This means that $\log h$ is concave and, hence, the inequality $\mu(H(u)) \geq p$ is equivalent with $\tilde{h}(u) \leq -\log p$, where $\tilde{h} := -\log h$ is a convex function. By 3. in Theorem 1 and Proposition 1, h is w.s.u.s. and, hence, $\tilde{h}$ is w.s.l.s. Define the multifunction $M : U \rightrightarrows \mathbb{R}$ by

$$M(u) := \begin{cases} [\tilde{h}(u), \infty) & \text{if } u \in U_0 \\ \emptyset & \text{else} \end{cases}.$$

We claim that M has a closed and convex graph. To this aim, consider an arbitrary sequence $(u_n, t_n) \to (\bar{u}, \bar{t})$ with $t_n \in M(u_n)$. Then, $u_n \in U_0$ and, hence, $\bar{u} \in U_0$ by closedness of U_0 (see 4. in Theorem 1). Moreover, $\tilde{h}(u_n) \leq t_n$. Since $\tilde{h}$ is w.s.l.s., we derive that

$$\tilde{h}(\bar{u}) \leq \liminf_{n\to\infty} \tilde{h}(u_n) \leq \liminf_{n\to\infty} t_n = \bar{t}.$$

Consequently, $\bar{t} \in M(\bar{u})$ implying that the graph of M is closed. To show its convexity, let $t^1 \in M(u^1)$, $t^2 \in M(u^2)$ and $\lambda \in [0,1]$ be arbitrarily given. Then, first, $u^1, u^2 \in U_0$, whence $\lambda u^1 + (1-\lambda)\, u^2 \in U_0$ by convexity of U_0 (see 4. in Theorem 1). Second, we have that $\tilde{h}(u^1) \leq t^1$ and $\tilde{h}(u^2) \leq t^2$. Then, convexity of $\tilde{h}$ yields that

$$\tilde{h}(\lambda u^1 + (1-\lambda)\, u^2) \leq \lambda t^1 + (1-\lambda)\, t^2.$$

In other words,

$$\lambda t^1 + (1-\lambda)\, t^2 \in M(\lambda u^1 + (1-\lambda)\, u^2),$$

proving that the graph of M is also convex.

Finally, observe that 5. in Theorem 1 implies $h(\hat{u}) > p$, whence $\tilde{h}(\hat{u}) < -\log p$. It follows that $-\log p \in \operatorname{int} M(\hat{u})$. Altogether, the previously shown properties allow us to invoke the Robinson-Ursescu Theorem [6, Chapter 3, Theorem 1] in order to derive the existence of some $\varepsilon > 0$ such that

$$d(u, M^{-1}(t)) \leq \frac{1}{\varepsilon} d(t, M(u))(1 + \|u - \hat{u}\|) \quad \forall u \in U_0\ \forall t \in [-\log p - \varepsilon, -\log p + \varepsilon].$$

Here, d represents the point to set distance and M^{-1} refers to the inverse multifunction corresponding to M. This is easily identified to be

$$M^{-1}(t) = \{u \in U_0 | \tilde{h}(u) \leq t\}.$$

Since U_0 is bounded, there exists some $\tilde{L} > 0$ with $\|u - \hat{u}\| \leq \tilde{L}$ for all $u \in U_0$. Hence, with $L := \tilde{L} + 1$, we get the estimate

$$\begin{array}{c} d(u, \{u \in U_0 | \tilde{h}(u) \leq t\}) \leq \frac{L}{\varepsilon} \max\{\tilde{h}(u) - t, 0\} \\ \forall u \in U_0\ \forall t \in [-\log p - \varepsilon, -\log p + \varepsilon] \end{array}.$$

which can further be developed to

$$\begin{array}{c} d(u, \{u \in U_0 | h(u) \geq e^{-t}\}) \leq \frac{L}{\varepsilon} \max\left\{\log \frac{e^{-t}}{h(u)}, 0\right\} \\ \forall u \in U_0\ \forall t \in [-\log p - \varepsilon, -\log p + \varepsilon] \end{array}.$$

and, finally, to

$$d\,(u, \{u \in U_0 | h(u) \geq \tau\}) \leq \frac{L}{\varepsilon} \max\{\log \tau - \log h(u), 0\} \quad \forall u \in U_0\ \forall \tau \in [pe^{-\varepsilon}, pe^{\varepsilon}].$$

Observing that $pe^{-\varepsilon} < p < pe^{\varepsilon}$, the assertion follows. □

Lemma 4 *Under the assumptions of Theorem 1, one has that for every $\nu \in \mathscr{P}(\mathbb{R}^s)$ the function $\nu\,(H(u))$ is w.s.u.s. and the set $\Phi\,(\nu)$ is weakly sequentially compact.*

Proof Let η be an s-dimensional random vector having distribution ν. Then, by definition

$$\nu\,(H(u)) = \mathbb{P}\,(\eta \in H(u)) = \mathbb{P}\,(g\,(u, \eta, x) \geq 0 \quad \forall x \in C)\,.$$

Replacing ξ by η in the definition of the probability function (1), Proposition 1 implies via 3. in Theorem 1 that the function $\nu(H(\cdot))$ is w.s.u.s. Hence, the set

$$\{u \in U | \nu(H(u)) \geq p\}$$

is weakly sequentially closed. On the other hand, U_0 is weakly sequentially compact by being bounded, convex and closed in a reflexive Banach space [12, p. 217] (see assumptions of Theorem 1). It follows that $\Phi(\nu)$, as an intersection of a weakly sequentially compact with a weakly sequentially closed set is weakly sequentially compact again. □

Lemma 5 *Let the assumptions of Theorem 1 hold true. Then, for every sequence* $(\nu_n, u_n) \in \mathscr{P}(\mathbb{R}^s) \times U$ *and every* $(\bar{\nu}, \bar{u}) \in \mathscr{P}(\mathbb{R}^s) \times U$ *satisfying the relations*

$$\alpha(\nu_n, \bar{\nu}) \to 0, u_n \rightharpoonup \bar{u}, u_n \in \Phi(\nu_n),$$

it follows that $\bar{u} \in \Phi(\bar{\nu})$.

Proof Clearly, $\bar{u} \in U_0$ due to $u_n \in U_0$ and by U_0 being weakly sequentially closed. By definition,

$$\nu_n(H(u_n)) \geq p \quad \forall n \in \mathbb{N}.$$

Next, let $\varepsilon > 0$ be arbitrarily given and let η be an s -dimensional random vector having distribution $\bar{\nu}$. Then, by definition

$$\bar{\nu}(H(u)) = \mathbb{P}(\eta \in H(u)) = \mathbb{P}(g(u, \eta, x) \geq 0 \quad \forall x \in C).$$

Replacing ξ by η in the definition of the probability function (1), Proposition 1 implies that the function $\bar{\nu}(H(\cdot))$ is w.s.u.s. Consequently, for n large enough, one has that

$$\bar{\nu}(H(u_n)) \leq \bar{\nu}(H(\bar{u})) + \varepsilon/2.$$

Since also $\alpha(\nu_n, \bar{\nu}) \leq \varepsilon/2$ for n large enough, we infer that

$$\begin{aligned}\bar{\nu}(H(\bar{u})) &\geq \bar{\nu}(H(\bar{u})) - \bar{\nu}(H(u_n)) - |\bar{\nu}(H(u_n)) - \nu_n(H(u_n))| + \nu_n(H(u_n)) \\ &\geq -\varepsilon/2 - \alpha(\nu_n, \bar{\nu}) + p \geq p - \varepsilon.\end{aligned}$$

As $\varepsilon > 0$ was chosen arbitrarily, it follows that $\bar{\nu}(H(\bar{u})) \geq p$ which entails that $\bar{u} \in \Phi(\bar{\nu})$. □

References

1. Van Ackooij, W., de Oliveira, W.: Convexity and optimization with copulae structured probabilistic constraints. Optimization **65**, 1349–1376 (2016)
2. Van Ackooij, W., Henrion, R.: (Sub-) Gradient formulae for probability functions of random inequality systems under Gaussian distribution. SIAM/ASA J. Uncertain. Quantif. **5**, 63–87 (2017)
3. Adams, R.A., Fournier, J.J.F.: Sobolev Spaces. Academic Press, Cambridge (2003)
4. Allaire, G., Dapogny, C.: A deterministic approximation method in shape optimization under random uncertainties. SMAI J. Comput. Math. **1**, 83–143 (2015)
5. Van Ackooij, W., Henrion, R.: Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions. SIAM J. Optim. **24**, 1864–1889 (2014)
6. Aubin, J.P., Cellina, A.: Differential Inclusions. Set-Valued Maps and Viability Theory. Springer, Berlin (1984)
7. Babuska, I., Nobile, F., Tempone, R.: A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data. SIAM J. Numer. Anal. **45**, 1005–1034 (2007)
8. Babuska, I., Chatzipantelidis, P.: On solving elliptic stochastic partial differential equations. Comput. Methods Appl. Mech. Engrg. **191**, 4093–4122 (2002)

9. Babuska, I., Liu, K., Tempone, R.: Solving stochastic partial differential equations based on the experimental data. Math. Models Methods Appl. Sci. **13**, 415–444 (2003)
10. Borzi, A.: Multigrid and sparse-grid schemes for elliptic control problems with random coefficients. Comput. Vis. Sci. **13**, 153–160 (2010)
11. Brenner, S.C., Scott, R.: The Mathematical Theory of Finite Element Methods, Texts in Applied Mathematics (2011)
12. Chang, K.-C.: Methods in Nonlinear Analysis. Springer, Berlin (2005)
13. Conti, S., Held, H., Pach, M., Rumpf, M., Schultz, R.: Risk averse shape optimization. SIAM J. Control Optim. **49**, 927–947 (2011)
14. Druet, P.E., Klein, O., Sprekels, J., Tröltzsch, F., Yousept, I.: Optimal control of 3D state-constrained induction heating problems with nonlocal radiation effects. SIAM J. Control Optim. **49**, 1707–1736 (2011)
15. Evans, L.C.: Partial Differential Equations, Graduate Studies in Mathematics Volume: 19 (2010)
16. Geihe, B., Lenz, M., Rumpf, M., Schultz, R.: Risk averse elastic shape optimization with parametrized fine scale geometry. Math. Program. **141**, 383–403 (2013)
17. Ghanem, R., Spanos, P.: Stochastic finite element expansion for random media. J. Eng. Mech. **115**, 1035–1053 (1989)
18. Gunzburger, M.D., Lee, H., Lee, J.: Error Estimates of Stochastic Optimal Neumann Boundary Control Problems. SIAM J. Numer. Anal. **49**, 1532–1552 (2011)
19. Henrion, R.: Qualitative stability of convex programs with probabilistic constraints. In: Nguyen, V.H., Strodiot, J.-J., Tossings, P. (eds.) Optimization, Lecture Notes in Economics and Mathematical Systems, vol. 481, pp. 164–180. Springer, Berlin (2000)
20. Henrion, R., Strugarek, C.: Convexity of Chance Constraints with Independent Random Variables. Comput. Optim. Appl. **41**, 263–276 (2008)
21. Henrion, R., Möller, A.: A gradient formula for linear chance constraints under Gaussian distribution. Math. Oper. Res. **37**, 475–488 (2012)
22. 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**, A1847—A1879 (2013)
23. Kouri, D.P., Surowiec, T.M.: Risk averse PDE-constrained optimization using the conditional value-at-risk. SIAM J. Optim. **26**, 365–396 (2016)
24. Kunoth, A., Schwab, C.: Analytic regularity and GPC approximation for control problems constrained by linear parametric elliptic and parabolic PDEs. SIAM J. Control Optim. **51**, 2442–2471 (2013)
25. Lord, G.J., Powell, C.E., Sahrdlow, T.: An introduction to computational stochastic PDEs. Cambridge University Press, New York (2014)
26. Marti, K.: Differentiation of probability functions: the transformation method. Comput. Math. Appl. **30**, 361–382 (1995)
27. Prėkopa, A.: On logarithmic concave measures and functions. Acta Sci. Math. (Szeged) **34**, 335–343 (1973)
28. Prėkopa, A.: Programming under probabilistic constraints with a random technology matrix. Optimization **5**, 109–116 (1974)
29. Prėkopa, A.: Stochastic Programming. Kluwer, Dordrecht (1995)
30. Römisch, W., Schultz, R.: Stability analysis for stochastic programs. Ann. Oper. Res. **30**, 241–266 (1991)
31. Shapiro, A., Dentcheva, D., Ruszczyński, A.: Lectures on Stochastic Programming, MPS-SIAM series on optimization 9 (2009)
32. Sichau, A., Ulbrich, S.: A second order approximation technique for robust shape optimization. Appl. Mech. Mater. **104**, 13–22 (2012)
33. Uryasev, S.: Derivatives of probability functions and some applications. Ann. Oper. Res. **56**, 287–311 (1995)