Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Chance constraints in PDE constrained optimization

M. Hassan Farshbaf-Shaker[1], René Henrion[1], Dietmar Hömberg[1,2]

submitted: November 25, 2016

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: farshbafshaker@wias-berlin.de
rene.henrion@wias-berlin.de
dietmar.hoemberg@wias-berlin.de

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

No. 2338

Berlin 2016

2000 *Mathematics Subject Classification.* 90C15, 49J20.

Key words and phrases. Chance constraints, PDE constrained optimization.

This research was partially carried out in the framework of MATHEON supported by the Einstein Foundation Berlin within the ECMath project SE13.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

Chance constraints represent a popular tool for finding decisions that enforce a robust 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 a finite-dimensional setting. The aim of this paper is to generalize some of these well-known semi-continuity and convexity properties to a setting of control problems subject to (uniform) state chance constraints.

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 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. [4], [5], [6] and references therein. Moreover, optimal control problems of such uncertain systems are of great practical importance. We mention here the works [7], [12], [15] 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., [22]). 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 unnessecarily conservative and methods from stochastic optimization are to be preferred.

In [11], [8], 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 [2] 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

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

where x, y refer to space and state variables, respectively, ω is a random event, 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

$$\mathbb{P}(y(x,\omega) \leq \bar{y}(x) \quad \forall x \in D) \geq p,$$

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 contro, the state chance constraint transforms into a (nonlinear) control constraint, thus defining an optimization problem with robust in the sense of probability decisions. 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

$$\mathbb{P}(C(y,\omega) \geq 0) \geq p$$

of the random cost function

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

thus making a link to the papers discussed before. Note, however, that 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 semi-continuity 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., [17], [18], [13]), or differentiability (e.g., [16], [23]). 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., [14], [3]). A classical monograph containing many basic theoretical results and numerous applications of chance constraints is [19]. A more modern presentation of the theory can be found in [21].

The paper is organized as follows: In Section 2, we provide some basic results on weak sequential closedness and convexity of chance constraints as well as weak sequential semi-continuity properties of probability functions in an abstract framework. In Section 3, 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 $(\Omega, \mathcal{F}, \mathbb{P})$.

Probability functions of this type figure prominently in stochastic optimization problems either in the form of probabilistic constraints $h(x) \geq p$ 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. The following proposition follows the idea of the proof of Prop. 3.1 in [20] which was given in a finite dimensional setting .

Proposition 1 *In (1), let U have a separable dual U^* and assume that the $g(\cdot,\cdot,x)$ are weakly sequentially upper semicontinuous for all $x \in C$. Then,*

1 $g(u,\cdot,x)$ is Borel measurable for each $u \in U$ and $x \in C$.

2 $M := \{u \in U \mid h(u) \geq p\}$ is weakly sequentially closed for each $p \in \mathbb{R}$.

3 h is weakly sequentially upper semicontinuous.

Proof. By assumption, the functions $g(u,\cdot,x) : \mathbb{R}^s \to \mathbb{R}$ are upper semicontinuous for each $u \in U$ and $x \in C$. Consequently, the sets $\{z \in \mathbb{R}^s | g(u,z,x) \geq 0\}$ are closed for all $u \in U$ and $x \in C$, which implies 1. and justifies to talk about probabilities of events as in (1). 3. is an immediate consequence of 2. Hence, in order to prove 2., let $p \in \mathbb{R}$ be arbitrary and consider any weakly convergent sequence $u_n \rightharpoonup_n \bar{u} \in U$ with $u_n \in M$ for all n. We have to show that $\bar{u} \in M$. We define

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

From $u_n \in M$, it follows that

$$\mathbb{P}(\xi \in H(u_n)) \geq p \quad \forall n \in \mathbb{N}. \tag{2}$$

Boundedness of u_n (by weak convergence) implies that there is some closed ball $\mathbb{B}$ with sufficiently large radius such that $\bar{u} \in \mathbb{B}$ and $u_n \in \mathbb{B}$ for all n. By U^* being separable, the weak topology on $\mathbb{B}$ is metrizable by some metric d. We put

$$A_k := \overline{\bigcup\{H(u) \mid u \in \mathbb{B},\, d(u,\bar{u}) \leq k^{-1}\}} \quad (k \in \mathbb{N}).$$

It holds that

$$H(\bar{u}) = \bigcap_{k \in \mathbb{N}} A_k. \tag{3}$$

Indeed, the inclusion '$\subseteq$' being trivial, let $z \in \bigcap_{k\in\mathbb{N}} A_k$ be arbitrary. By definition, there exist sequences $z_k \in \mathbb{R}^s$ and $w_k \in \mathbb{B}$ such that

$$\|z_k - z\| \leq k^{-1},\ d(w_k,\bar{u}) \leq k^{-1},\ z_k \in H(w_k) \quad \forall k \in \mathbb{N}.$$

Hence $z_k \to_k z$ and $w_k \rightharpoonup_k \bar{u}$, where the latter convergence follows from the fact that d metrizes the weak topology on $\mathbb{B}$. In particular, $g(w_k,z_k,x) \geq 0$ for all $x \in C$ and all $k \in \mathbb{N}$. Now, the weak sequential upper semicontinuity of $g(\cdot,\cdot,x)$ for all $x \in C$ yields that

$$g(\bar{u},z,x) \geq \limsup_k g(w_k,z_k,x) \geq 0 \quad \forall x \in C.$$

Hence, $z \in H(\bar{u})$ which shows the reverse inclusion of (3).

Clearly, $A_{k+1} \subseteq A_k$ for all $k \in \mathbb{N}$ which along with (3) entails that

$$\mathbb{P}\left(\xi \in A_k\right) \to_k \mathbb{P}\left(\xi \in H(\bar{u})\right).$$

Accordingly, for any arbitrarily fixed $\varepsilon > 0$ there is some $k' \in \mathbb{N}$ with

$$\mathbb{P}\left(\xi \in H(\bar{u})\right) - \mathbb{P}\left(\xi \in A_{k'}\right) \geq -\varepsilon.$$

Moreover, by $u_n \rightharpoonup_n \bar{u}$ there exists some $n^* \in \mathbb{N}$ with $d\left(u_{n^*}, \bar{u}\right) \leq \left(k'\right)^{-1}$. It follows that $H\left(u_{n^*}\right) \subseteq A_{k'}$, whence altogether

$$\mathbb{P}\left(\xi \in H(\bar{u})\right) - \mathbb{P}\left(\xi \in H\left(u_{n^*}\right)\right) \geq \mathbb{P}\left(\xi \in H(\bar{u})\right) - \mathbb{P}\left(\xi \in A_{k'}\right) \geq -\varepsilon.$$

Now, (2) provides that $\mathbb{P}\left(\xi \in H(\bar{u})\right) \geq p - \varepsilon$. Since, $\varepsilon > 0$ was chosen arbitrarily, we infer that $\mathbb{P}\left(\xi \in H(\bar{u})\right) \geq p$ or $\bar{u} \in M$ as was to be shown. □

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 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), let U have a separable dual U^*. Assume that*

1 *C is a compact subset of $\mathbb{R}^d$.*

2 *the g are weakly sequentially lower semicontinuous (as functions of all three variables simultaneously).*

3 *the $g(u,\cdot,x)$ are concave for all $u \in U$ and $x \in C$.*

4 *For each $u \in U$ there exists some $\bar{z} \in \mathbb{R}^s$ such that $g(u,\bar{z},x) > 0$ for all $x \in C$.*

5 *ξ has a density.*

Then h is weakly sequentially lower semicontinuous.

Proof. We introduce the function

$$g^{min}(u,z) := \min_{x\in C} g(u,z,x) \quad (u\in U;\ z\in\mathbb{R}^s)$$

Note first that this definition is justified by weak sequential lower semicontinuity of g in x and by compactness of C. We show next that, as a consequence of our assumptions 1. and 2., g^{min} is weakly sequentially lower semicontinuous. Indeed, fixing an arbitrary $(\bar{u},\bar{z})\in U\times\mathbb{R}^s$, we may select a weakly convergent sequence $(u_k,z_k)\rightharpoonup(\bar{u},\bar{z})$ realizing the inferior limit:

$$\lim_k g^{min}(u_k,z_k) = \liminf_{(u,z)\rightharpoonup(\bar{u},\bar{z})} g^{min}(u,z).$$

By definition, there exists a sequence $x_k\in C$ such that $g^{min}(u_k,z_k)=g(u_k,z_k,x_k)$. Due to the compactness of C, we may assume that $x_{k_l}\to_l \bar{x}$ for some subsequence and some $\bar{x}\in C$. Summarizing, exploiting the assumed weak lower semicontinuity of g in all three variables simultaneously, we get the desired weak lower semicontinuity of g^{min}:

$$\begin{aligned}\liminf_{(u,z)\rightharpoonup(\bar{u},\bar{z})} g^{min}(u,z) &= \lim_l g^{min}(u_{k_l},z_{k_l}) = \lim_l g(u_{k_l},z_{k_l},x_{k_l})\\ &= \liminf_l g(u_{k_l},z_{k_l},x_{k_l}) \geq g(\bar{u},\bar{z},\bar{x}) \geq g^{min}(\bar{u},\bar{z}).\end{aligned}$$

As a consequence, $-g^{min}$ is weakly sequentially upper semicontinuous. We may apply now Proposition 1 in order to derive the weak sequential upper semicontinuity of the probability function

$$\tilde{h}(u) := \mathbb{P}(-g^{min}(u,\xi)\geq 0)$$

(by formally choosing the set C as a singleton so that the dependence on x disappears). Accordingly, for an arbitrarily fixed $\bar{u}\in U$ we have that

$$\limsup_{u\rightharpoonup\bar{u}} \mathbb{P}(-g^{min}(u,\xi)\geq 0) \leq \mathbb{P}(-g^{min}(\bar{u},\xi)\geq 0). \tag{4}$$

Then, it follows:

$$\begin{aligned}\liminf_{u\rightharpoonup\bar{u}} h(u) &= \liminf_{u\rightharpoonup\bar{u}} \mathbb{P}\left(g\left(u,\xi,x\right)\geq 0 \quad \forall x\in C\right) & (5)\\ &= \liminf_{u\rightharpoonup\bar{u}} \mathbb{P}\left(g^{min}\left(u,\xi\right)\geq 0\right)\\ &\geq \liminf_{u\rightharpoonup\bar{u}} \mathbb{P}\left(g^{min}\left(u,\xi\right)> 0\right) = -\limsup_{u\rightharpoonup\bar{u}} -\mathbb{P}\left(g^{min}\left(u,\xi\right)> 0\right)\\ &= -\limsup_{u\rightharpoonup\bar{u}} \left(\mathbb{P}\left(g^{min}\left(u,\xi\right)\leq 0\right)-1\right)\\ &= 1-\limsup_{u\rightharpoonup\bar{u}} \mathbb{P}\left(g^{min}\left(u,\xi\right)\leq 0\right)\\ &\geq 1-\mathbb{P}\left(g^{min}\left(\bar{u},\xi\right)\leq 0\right) & (6)\\ &= \mathbb{P}\left(g^{min}\left(\bar{u},\xi\right)> 0\right) = \mathbb{P}\left(g^{min}\left(\bar{u},\xi\right)\geq 0\right) & (7)\\ &= \mathbb{P}\left(g\left(\bar{u},\xi,x\right)\geq 0 \quad \forall x\in C\right) = h(\bar{u}).\end{aligned}$$

Here, (6) follows from (4) and it remains to justify (25): Observe first that, as a consequence of our assumption 3., $g^{min}(\bar{u},\cdot)$ is a concave function. Moreover, by our assumption 4., there exists some $\bar{z}\in\mathbb{R}^s$ such that $g^{min}(\bar{u},\bar{z})>0$. Both observations entail that the set

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

is a subset of the boundary of the convex set

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

Since the boundary of a convex set has Lebesgue measure zero, E itself has Lebesgue measure zero. By our assumption 5., the distribution of ξ is absolutely continuous with respect to the Lebesgue measure, whence $\mathbb{P}(\xi \in E) = 0$. This finally yields (25), so that the chain of relations above proves the weak sequential lower semicontinuity of h. □

Remark 1 *Assumptions 1. and 2. in the previous proposition were needed in order to show the weak sequential lower semicontinuity of the minimum function g^{min}. If the set C happens to be just a finite one, then of course the same property of g^{min} can be derived from the substantially weaker (compared with 2.) assumption that $g(\cdot, \cdot, x)$ is weakly sequentially lower semicontinuous for each $x \in C$ because the finite minimum of lower semicontinuous functions also happens to be so.*

Corollary 1 *In (1), let U have a separable dual U^*. Assume that*

1 C is a compact subset of $\mathbb{R}^d$.

2 the g are weakly sequentially continuous (as functions of all three variables simultaneously).

3 the $g(u, \cdot, x)$ are concave for all $u \in U$ and $x \in C$.

4 For each $u \in U$ there exists some $\bar{z} \in \mathbb{R}^s$ such that $g(u, \bar{z}, x) > 0$ for all $x \in C$.

5 ξ has a density.

Then h is weakly sequentially continuous. Moreover, owing to Remark 1, the same result can be derived in the case that C happens to be a finite set upon replacing 2. by the weaker assumption that the $g(\cdot, \cdot, x)$ are weakly sequentially continuous for all $x \in C$.

We finally 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 $p : V \to \mathbb{R}$ (V a vectors space) is defined to be quasiconcave, if the following relation holds true:

$$p(\lambda x + (1 - \lambda)y) \geq \min\{p(x), p(y)\} \quad \forall x, y \in V; \quad \forall \lambda \in [0, 1]$$

The next proposition can be proven exactly in the same way as in [19, 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 3 *In the setting of* (1)*, let U be an arbitrary vector space. Moreover, let the random vector ξ have 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*

$$M := \{u \in U \mid h(u) \geq p\} \tag{8}$$

is convex for any $p \in [0, 1]$.

Proof. Define the infimum function

$$g^{inf}(u,z) := \inf_{x\in C} g(u,z,x) \quad (u \in U;\ z \in \mathbb{R}^s)$$

and observe that, according to (1),

$$h(u) = \mathbb{P}(g^{inf}(u,\xi) \geq 0) \quad (u \in U). \tag{9}$$

We note that g^{inf} is quasiconcave. Indeed, fix an arbitrary couple 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

$$\begin{array}{rcl} g^{inf}(\lambda(u^1,z^1)+(1-\lambda)(u^2,z^2)) & \geq & g(\lambda(u^1,z^1)+(1-\lambda)(u^2,z^2),x)-\varepsilon \\ & \geq & \min\{g(u^1,z^1,x),g(u^2,z^2,x)\}-\varepsilon \\ & \geq & \min\{g^{inf}(u^1,z^1),g^{inf}(u^2,z^2)\}-\varepsilon \end{array}$$

Here, in the second inequality, we exploited our assumption on $g(\cdot,\cdot,x)$ being quasiconcave for all $x \in C$. As $\varepsilon > 0$ was arbitrarily chosen, the claimed quasiconcavity of g^{inf} follows. Next, the assumption on ξ having a logconcave density implies by Prekopa's Theorem [19, Theorem 4.2.1] that ξ has a logconcave distribution. This means that

$$\mathbb{P}(\xi \in \lambda A + (1-\lambda)B) \geq [\mathbb{P}(\xi \in A)]^\lambda [\mathbb{P}(\xi \in B)]^{1-\lambda} \tag{10}$$

holds true for all convex subsets $A, B \in \mathbb{R}^s$ and all $\lambda \in [0,1]$. I order to prove the claimed convexity of the set M in (8), 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

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

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

$$H(\lambda u^1 + (1-\lambda)u^2) \supseteq \lambda H(u^1) + (1-\lambda)H(u^2). \tag{11}$$

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,

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

Exploiting the quasiconcavity of g^{inf} proven above, we arrive at

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

In other words, $z \in H(\lambda u^1 + (1-\lambda)u^2)$, which proves (11). Now, (9) along with (10) 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. □

3 Example from PDE constrained optimization

In this part we are going to apply the method developed in section 2 to a simple PDE constrained optimization with chance constraint. To this end, we consider a fairly general PDE

$$\begin{aligned} -\nabla_x \cdot (\kappa(x)\, \nabla_x y(x,\omega)) = r(x,\omega), \quad (x,\omega) \in D \times \Omega \\ n \cdot (\kappa(x)\, \nabla_x y(x,\omega)) + \alpha\, y(x,\omega) = u(x) \quad (x,\omega) \in \partial D \times \Omega, \end{aligned} \tag{12}$$

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 ω is the stochastic variable, which as in Section 2 belongs to a complete probability space denoted by $(\Omega, \mathcal{F}, P)$. Here Ω is the set of outcomes, $\mathcal{F} \subset 2^\Omega$ is the σ-algebra of events, and $\mathbb{P} : \mathcal{F} \to [0,1]$ is a probability measure. In (12) 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. [11], or in induction heating problems in semiconductor single crystal growth processes, see e.g. [9]. 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 [9] 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 temperatur field.

To ensure well-posedness of (12), we assume that

$$D \in C^{1,1}, \quad \kappa \in C^{0,1}(D) \quad \text{and } \exists \kappa_0 > 0 \,:\, \kappa_0 \leq \kappa(x)\, \forall x \in D. \tag{13}$$

3.1 Well-posedness of (12)

Throughout this paper, we use standard notations (e.g., see [1]) 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 context. Let ξ be an $\mathbb{R}^s$-valued random variable in a probability space $(\Omega, \mathcal{F}, P)$. If $\xi \in L^1_P(\Omega)$, we define $\mathbb{E}\xi = \int_\Omega \xi(\omega)\, dP(\omega)$ 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

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

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:

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

For instance,

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

and

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

Moreover we define

$$\mathcal{C}(\bar{D}) = L^2(\Omega; C(\bar{D})).$$

We now state the well-posedness for (12).

Proposition 4 *Let (13) be fulfilled. Then for every $(r, u) \in \mathcal{L}^2(D) \times H^{\frac{1}{2}}(\partial D)$ there exists a unique solution $y \in \mathcal{H}^2(D)$ of (12) in the sense*

$$\begin{aligned} &\mathbb{E}\left(\int_D \kappa(x)\, \nabla_x y(x,\omega) \cdot \nabla_x \varphi(x,\omega)\, dx + \alpha \int_{\partial D} y(x,\omega)\, \varphi(x,\omega)\, ds\right) \\ &= \mathbb{E}\left(\int_D r(x,\omega)\, \varphi(x,\omega)\, dx + \int_{\partial D} u(x)\, \varphi(x,\omega)\, ds\right), \quad \forall \varphi \in \mathcal{H}^1(D) \end{aligned} \tag{14}$$

Moreover, the mapping

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

is linear and continuous, i.e.

$$\|y\|_{\mathcal{H}^2(D)} \le C \left(\|r\|_{\mathcal{L}^2(D)} + \|u\|_{H^{\frac{1}{2}}(\partial D)}\right). \tag{15}$$

Proof. Use the Lax-Milgram lemma and [10]. □

Remark 2 *For $n = 3$, we know that the continuous embedding $\mathcal{H}^2(D) \hookrightarrow \mathcal{C}(\bar{D})$ is fulfilled. Hence, the solution y from Proposition 4 belongs to $\mathcal{C}(\bar{D})$ and we further obtain*

$$\|y\|_{\mathcal{C}(\bar{D})} \le C \left(\|r\|_{\mathcal{L}^2(D)} + \|u\|_{H^{\frac{1}{2}}(\partial D)}\right). \tag{16}$$

3.2 Optimization problem

Let U_{ad} be a bounded, closed and convex subset of $U := H^{\frac{1}{2}}(\partial D)$ and $\bar{y}(x) \in C(\bar{D})$ a given function. Moreover, we will work in different cases with a subset C of D. The first case ist $C = D$, whereas the second case is $C \subset D$ with finite C. Given a weakly sequentially lower semi-continuous cost functional $L : \mathcal{H}^2(D) \times H^{\frac{1}{2}}(\partial D) \to \mathbb{R}$ which is additionally bounded from below, our overall optimization problem reads as

$$(P) \quad \begin{cases} \text{min} & \mathbb{E}(L(y(x,\omega), u(x))) \\ \text{over} & \mathcal{H}^2(D) \times U_{ad} \\ \text{s.t.} & (12) \text{ is satisfied} \\ & \mathbb{P}(\omega \in \Omega \mid y(x,\omega) \leq \bar{y}(x), \forall x \in C) \geq p, \quad p \in (0,1) \end{cases}$$

Remark 3 *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 of important goals. In application this is an important issue since engineers are interested to prevent damage in semiconductor single crystals which are caused by high temperatur distributions. But as one has to deal with uncertain time-harmonic electromagnetic fields the temperatur field is consequently random, too. In this case it is reasonable to claim that the temperatur as state variable stays with high probability in some prescribed domain.*

3.3 Finite sum expansion

For the source function r in (12) we make the ansatz of a finite (truncated) sum expansion

$$r(x,\omega) := \sum_{k=1}^{s} \phi_k(x)\, \xi_k(\omega), \tag{17}$$

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 [2, Section 2.4]. With

$$\phi(x) := (\phi_1(x), \ldots, \phi_s(x))^T; \quad \xi(\omega) := (\xi_1(\omega), \ldots, \xi_s(\omega))^T,$$

we define

$$\tilde{r}(x,\xi) := \phi(x) \cdot \xi(\omega), \tag{18}$$

where ξ is is an $\mathbb{R}^s$-valued random variable. Using the solution operator Y and (18) we define

$$g : U \times \mathbb{R}^s \times D \to \mathbb{R}, \quad g(u,\xi,x) := \bar{y}(x) - Y(\tilde{r}(x,\xi), u(x)). \tag{19}$$

Lemma 5 *Let $C = D$ in (P). Then $g(\cdot,\cdot,x)$, defined in (19), is weakly sequentially continuous for all $x \in C$.*

Proof. Using Proposition 4, in case of (17), we obtain from (15) the estimate

$$\|y\|_{\mathcal{H}^2(D)} \leq C\left((\|\phi\|_{[L^2(D)]^s} \cdot \|\xi\|_{[L^2(\Omega)]^s}) + \|u\|_U\right). \tag{20}$$

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 weakly sequentially continuity. Consequently the assertion of the lemma immediately follows. □

3.4 Properties of the reduced problem

Defining the reduced cost functional by

$$f(u(\cdot)) := \mathbb{E}(L(Y(\tilde{r}(\cdot, \xi), u(\cdot)), u(\cdot))) \tag{21}$$

and using the definition

$$h(u) := \mathbb{P}(g(u(x), \tilde{r}(x, \xi), x) \geq 0, \forall x \in C), \tag{22}$$

the chance constraint in (P) can be formulated as

$$M := \{u \in U \,|\, h(u) \geq p\}. \tag{23}$$

Then the reduced optimal control problem reads as

$$(P) \qquad \min_{u \in U_{ad} \cap M} f(u). \tag{24}$$

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

Theorem 6 *The problem (P) admits a solution $u \in U_{ad} \cap M$.*

Proof. As a Hilbert space $U := H^{\frac{1}{2}}(\partial D)$ has a separable dual and moreover g is weakly sequentially continuous by Lemma 5. Consequently Proposition 1 from section 2 yields that M is weakly sequentially closed. Hence, by assumptions on U_{ad} it is obvious that $U_{ad} \cap M$ is weakly sequentially closed, too. Taking into account the assumptions on the cost functional 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 weakly sequentially upper semicontinuity of the function h. This was done by using Lemma 5 and Proposition 1. In the following theorem we will refine this upper semicontinuity result to a semicontinuity result by taking into account additional assumptions. The theorem will then ensure weakly sequentially continuity of the function h.

Theorem 7 *Let C be finite and the random variable ξ, defined in (18), have a density. Moreover, assume that for each $u \in U$ there exists some $\bar{z} \in \mathbb{R}^s$ such that*

$$Y(\tilde{r}(x, \bar{z}), u(x)) < \bar{y}(x) \qquad \forall x \in C. \tag{25}$$

Then the function h, defined in (22), is weakly sequentially continuous.

Proof. Using once again Lemma 5, it follows that g is weakly sequentially continuous as a function of all three variables simultaneously. Moreover, it is obvious that $g(u, \cdot, x)$ is linear for all $u \in U$ and $x \in C$, and consequently concave. Then the third assumption of Corollary 1 is fulfilled. Hence, the assertion of the Theorem follows from Corollary 1. $\square$

The condition given by (25) 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.

To provide an instance for the use of Theorem 7 is the consideration of random state constraints in disjunctive form which would lead to the following state chance constraint:

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

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 (19) and (22), respectively, we arrive at the condition

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

So, instead of (23) 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 6, one would now need the weak sequential lower (rather than upper) semicontinuity of h. This would come as a consequence of Theorem 7.

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

Theorem 8 *Let the random variable ξ, defined in (18), 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 (20), 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, by the linearity of $g(\cdot, \cdot, x)$, it follows that it is quasiconcave for all $x \in C$. Then, it follows from Proposition 3 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. $\square$

Remark 4 *Numerous multivariate distributions do have logconcave densities, e.g. normal distribution, Student's t-distribution, uniform distribution on compact and convex sets, see e.g. [19].*

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 $L_1 : \mathcal{H}^2(D) \ni y \mapsto L_1(y) \in \mathbb{R}$ and $L_2 : U \ni u \mapsto L_2(u) \in \mathbb{R}$.

References

[1] R. A. Adams, J. J. F. Fournier, Sobolev Spaces, Academic Press, 2003.

[2] G. Allaire, C. Dapogny, A deterministic approximation method in shape optimization under random uncertainties, SMAI J. Comput. Math., 1 (2015), 83-143.

[3] W. Van Ackooij, and R. Henrion, Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions. SIAM J. Optim., 24 (2014) 1864-1889.

[4] I. Babuska, F. Nobile, and R. Tempone, A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data, SIAM J. Numer. Anal., 45 (2007), 1005-1034.

[5] I. Babuska and P. Chatzipantelidis, On solving elliptic stochastic partial differential equations, Comput. Methods Appl. Mech. Engrg., 191 (2002), 4093-4122.

[6] I. Babuska, K. Liu, and R. Tempone, Solving stochastic partial differential equations based on the experimental data, Math. Models Methods Appl. Sci., 13 (2003), 415-444.

[7] A. Borzi, Multigrid and sparse-grid schemes for elliptic control problems with random coefficients, Comput. Vis. Sci., 13 (2010), 153-160.

[8] S. Conti, H. Held, M. Pach, M. Rumpf, and R. Schultz, Risk averse shape optimization, SIAM J. Control Optim., 49 (2011), 927-947.

[9] P. E. Druet, O. Klein, J. Sprekels, F. Tröltzsch and I. Yousept, Optimal control of 3D state-constrained induction heating problems with nonlocal radiation effects, SIAM J. Control Optim., 49 (2011), 1707-1736.

[10] L. C. Evans, Partial Differential Equations, Graduate Studies in Mathematics Volume: 19; 2010.

[11] B. Geihe, M. Lenz, M. Rumpf, R. Schultz, Risk averse elastic shape optimization with parametrized fine scale geometry. Math. Program., 141 (2013), 383-403.

[12] M. D. Gunzburger, H. Lee, and J. Lee, Error Estimates of Stochastic Optimal Neumann Boundary Control Problems, SIAM J. Numer. Anal., 49 (2011), 1532-1552.

[13] R. Henrion, and C. Strugarek, Convexity of Chance Constraints with Independent Random Variables, Comput. Optim. Appl., 41 (2008), 263-276.

[14] R. Henrion, and A. Möller, A gradient formula for linear chance constraints under Gaussian distribution, Math. Oper. Res., 37 (2012), 475-488.

[15] A. Kunoth and C. Schwab, Analytic Regularity and GPC Approximation for Control Problems Constrained by Linear Parametric Elliptic and Parabolic PDEs, SIAM J. Control Optim., 51 (2013), 2442-2471.

[16] K. Marti, Differentiation of probability functions: the transformation method, Comput. Math. Appl., 30 (1995), 361-382.

[17] A. Prékopa, On logarithmic concave measures and functions, Acta Sci. Math. (Szeged), 34 (1973), 335-343.

[18] A. Prékopa, Programming under probabilistic constraints with a random technology matrix, Optimization, 5 (1974), 109-116.

[19] A. Prékopa, *Stochastic Programming*, Kluwer, Dordrecht, 1995.

[20] W. Römisch and R. Schultz, Stability analysis for stochastic programs, Ann. Oper. Res., 30 (1991), 241-266.

[21] A. Shapiro, D. Dentcheva, and A. Ruszczyński, Lectures on Stochastic Programming, MPS-SIAM series on optimization 9, 2009.

[22] A. Sichau, and S. Ulbrich, A second order approximation technique for robust shape optimization, Appl. Mech. Mater., 104 (2012), 13-22.

[23] S. Uryasev, Derivatives of Probability Functions and some Applications, Ann. Oper. Res., 56 (1995), 287-311.