SIAM J. CONTROL OPTIM.
Vol. 58, No. 4, pp. 2288–2311

OPTIMAL NEUMANN BOUNDARY CONTROL OF A VIBRATING STRING WITH UNCERTAIN INITIAL DATA AND PROBABILISTIC TERMINAL CONSTRAINTS*

M. HASSAN FARSHBAF-SHAKER†, MARTIN GUGAT‡, HOLGER HEITSCH‡, AND RENÉ HENRION‡

Abstract. In optimal control problems, often initial data are required that are not known exactly in practice. In order to take into account this uncertainty, we consider optimal control problems for a system with an uncertain initial state. A finite terminal time is given. On account of the uncertainty of the initial state, it is not possible to prescribe an exact terminal state. Instead, we are looking for a control that steers the system into a given neighborhood of the desired terminal state with sufficiently high probability. This neighborhood is described in terms of an inequality for the terminal energy. The probabilistic constraint in the considered optimal control problem leads to optimal controls that are robust against the inevitable uncertainties of the initial state. Numerical examples with optimal Neumann control of the wave equation are presented.

Key words. terminal constraint, uncertain initial data, probabilistic constraint, optimal control, boundary control, wave equation

AMS subject classifications. 49J20, 49J55, 49M37

DOI. 10.1137/19M1269944

1. Introduction.

Many applications in engineering sciences are modeled by initial boundary value problems with a hyperbolic system; see, for example, [3, 2, 6]. In applications in engineering, often some data are uncertain. In order to provide analytical insights for such a situation, in this paper we consider a system that is governed by a wave equation with uncertain initial data that are modeled by random Fourier series. The corresponding Cauchy problems have been analyzed in [4].

For the situation with uncertain initial data, we consider an optimal control problem with conditions on the terminal state. For a given control, the terminal state is also uncertain. Optimal control problems with terminal constraints have been studied before in the classical deterministic setting where the desired terminal state is prescribed exactly; see, for example, [24, 31, 18]. This is possible since the system is exactly controllable; that is, for a known initial state and a sufficiently large control time, there exists a control such that the desired terminal state is reached exactly (see [26]). This exact terminal condition is equivalent to a sequence of moment equations for the terminal state; see, for example, [15]. Control to a position of rest can also be characterized by the requirement that the terminal energy be equal to zero.

In our probabilistic setting, since the initial state is uncertain for a fixed control function, it is impossible to predict the terminal state exactly. Therefore, instead of

*Received by the editors June 24, 2019; accepted for publication (in revised form) June 12, 2020; published electronically August 5, 2020.

https://doi.org/10.1137/19M1269944

Funding: This work was supported by DFG in the Collaborative Research Centre CRC/Transregio 154, Mathematical Modelling, Simulation and Optimization Using the Example of Gas Networks, Project B04, Project C03, and Project C05 and by the FMJH Program Gaspard Monge in optimization and operations research including support to this program by EDF.

†Weierstrass-Institut, Berlin 10117, Germany (shaker@wias-berlin.de, heitsch@wias-berlin.de, henrion@wias-berlin.de).

‡Mathematik, Lehrstuhl für Angewandte Analysis (Alexander von Humboldt-Professur), Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen 91058, Germany (martin.gugat@fau.de).

a terminal constraint, we require that the terminal energy of the system state be less than or equal to a given upper bound ε at least with a given probability. So, similar to [17], we present a relaxation of the exact terminal constraint that increases the robustness of the optimal controls.

In this way, we obtain optimal controls for which the probability that the terminal position is contained in a certain ε-neighborhood of the desired terminal position is greater than or equal to a prescribed parameter $p \in (0, 1)$. So, starting from the uncertain initial state, the optimal controls generate a terminal state that satisfies a *probabilistic constraint*. The solution of this optimal control problem can be approximated by a suitable numerical method. By truncating the infinite series that represents the terminal energy, we obtain a sequence of auxiliary problems whose solutions converge to the solution of the optimal control problem with the probabilistic constraint on the terminal energy. In this way, we can compute approximations for a control that satisfies the probabilistic constraints and minimizes the objective function of the optimal control problem. In this paper, we consider the L^2-norm of the control as the objective function. In fact, if for some $p > 2$ the L^p-norm is chosen as the objective function, for the deterministic problem with terminal constraints the situation is quite similar as for $p = 2$; see, for example, [18] for the case $p \in (2, \infty)$ and [17] for the case $p = \infty$. We expect that also for the problems with probabilistic terminal constraints, a generalization to these cases is possible; however, in order to keep the presentation concise, here we only consider the case $p = 2$.

The problem that we consider falls into the class of risk-averse PDE-constrained optimization problems. This topic has gained much interest recently. One direction of approaching such problems is the consideration of the conditional value at risk in the objective or constraints (see, e.g., [23], and see [20] for the approximation of the conditional value at risk). Another perspective, the one taken in this paper, is the use of probabilistic constraints, which are very popular in engineering problems but traditionally applied in finite dimension. As stated in a recent monograph [27, p. 89], their application in the environment of PDE-constrained optimization is still in its infancy (e.g., [7, 11, 29]). It can be observed that ideas from various backgrounds (e.g., numerical solution of PDEs, stochastic optimization, etc.) are coming to the fore right now. For instance, in [27, 29], the basic approach consists in smoothing the Heaviside function and using some sampling scheme of the given random vector in order to approximate the values and gradients of the probability function involved. This "mollifying" idea can be traced back in the stochastic optimization literature at least to the 1980s (see, e.g., [10, 25]) and has received recent attention again in numerical methods for probabilistic programming [12, 30].

Our approach presented here follows the spirit of the "Hungarian school" represented by A. Prékopa and his students (see [28]) in that it aims at exploiting specific information about the problem, for instance, a concrete probability distribution (or a class thereof). More precisely, we assume a multivariate Gaussian distribution of the random vector (which could be replaced by any elliptically symmetric or Gaussian-like distribution, such as Student, lognormal, or truncated Gaussian) and apply the spheric-radial decomposition method, which allows for an efficient approximation of the (original, not smoothed) probability function and its gradient [1] even in larger dimensions of the random vector (say, a few hundred). Moreover, we benefit from an explicit description via Fourier series of the control-to-state operator in order to derive structural results (convexity, existence of solutions, convergence with increasing number of Fourier coefficients, etc.).

The paper has the following structure. In section 2 we discuss the deterministic initial boundary value problem and the corresponding problem with uncertain initial data. In section 3 we discuss the corresponding problem of norm-minimal Neumann control for the system with uncertain initial state and a probabilistic constraint for the terminal energy. Also, an approximation of the energy with finite sums is introduced.

In section 4 it is shown that the probabilistic constraint in the optimal control problem with approximated energy is convex. Based on this fact, methods for the computation of the probabilities and the corresponding gradients with respect to the control are presented in section 5. The existence of optimal controls for the original problem and the problems with the approximated energy constraint are shown in section 6. In section 7 numerical examples are illustrated and discussed.

2. System states for Neumann boundary of the wave equation control with uncertain initial data. We consider a vibrating string of length $L > 0$ with homogeneous Dirichlet boundary conditions at one end and Neumann-boundary control at the other end. To study this problem, we use the Sobolev space $H^1(0, L) = \{f \in L^2(0, L)$: The derivative f' in the sense of distributions is in $L^2(0, L)\}$. Let $y_0 \in H^1(0, L)$ and $y_1 \in L^2(0, L)$ be given. Let $T > 0$ denote a given terminal time, and let $c > 0$ denote the wave speed. For a given control $u \in L^2(0, T)$, the deterministic initial boundary value problem

$$(P_{\mathrm{det}}) \begin{cases} y(0,x) = y_0(x), & x \in (0,L) \\ y_t(0,x) = y_1(x), & x \in (0,L) \\ y_{tt}(t,x) = c^2\, y_{xx}(t,x), & (t,x) \in (0,T) \times (0,L) \\ y(t,0) = 0, & t \in (0,T) \\ y_x(t,L) = u(t), & t \in (0,T) \end{cases}$$

has been analyzed, for example, in [16, Thm. 2.3]. In engineering applications, often the initial data are uncertain. As a model for the uncertain initial data, we use random Fourier series as studied in [22]. For $n \in \{0, 1, 2, \ldots\}$, define the complete orthonormal series

$$\varphi_n(x) := \frac{\sqrt{2}}{\sqrt{L}}\ \sin\left(\left(\frac{\pi}{2} + n\pi\right)\frac{x}{L}\right)$$

and the corresponding coefficients $\alpha_n^0, \alpha_n^1 \in \mathbb{R}$ with

$$\alpha_n^0 := \int_0^L y_0(x)\,\varphi_n(x)\,dx, \quad \alpha_n^1 := \int_0^L y_1(x)\,\varphi_n(x)\,dx.$$

Then we have the series representations

$$y_0(x) = \sum_{n=0}^{\infty} \alpha_n^0\,\varphi_n(x), \quad y_1(x) = \sum_{n=0}^{\infty} \alpha_n^1\,\varphi_n(x). \tag{2.1}$$

Unless otherwise stated, we shall make the following standing assumption for our model of uncertain initial data.

(A) Assume that for all $n \in \{0, 1, 2, \ldots\}$, identically distributed random variables a_n, b_n are given on a complete probability space $(\Omega, \mathcal{A}, \mathcal{P})$.

Consider the random initial data

$$y_0^\omega(x) = \sum_{n=0}^{\infty} a_n^\omega\,\alpha_n^0\,\varphi_n(x), \quad y_1^\omega(x) = \sum_{n=0}^{\infty} b_n^\omega\,\alpha_n^1\,\varphi_n(x), \tag{2.2}$$

where the superscript ω indicates the evaluation of the random variable at this out-

come $\omega \in \Omega$. Then our assumption (**A**) along with an argument related to the Paley–Zygmund theorem implies that for these random series, almost surely we have $y_0^\omega \in H^1(0,L)$, $y_1^\omega \in L^2(0,L)$ (see [21, Lem. 4.3], [4, 8]). The corresponding probabilistic initial boundary value problem with uncertain initial data is given by

$$(P_{\mathsf{prob}}) \begin{cases} y^\omega(0,x) = y_0^\omega(x), & x \in (0,L) \\ y_t^\omega(0,x) = y_1^\omega(x), & x \in (0,L) \\ y_{tt}^\omega(t,x) = c^2\, y_{xx}^\omega(t,x), & (t,x) \in (0,T)\times(0,L) \\ y^\omega(t,0) = 0, & t \in (0,T) \\ y_x^\omega(t,L) = u(t), & t \in (0,T), \end{cases}$$

where $\omega \in \Omega$. The solution of (P_{prob}) is almost surely at least as well behaved as the solution of (P_{det}). In fact, we have again a series representation that is presented in the following theorem.

THEOREM 2.1. *Assume that $y_0 \in H^1(0,L)$ with $y_0(0) = 0$, $y_1 \in L^2(0,L)$, and $u \in L^2(0,T)$. For $n \in \{0,1,2,\ldots\}$, define*

$$\lambda_n := \frac{1}{L^2}\left(\frac{\pi}{2} + n\pi\right)^2$$

and the random variables

$$\begin{aligned}\alpha_n^\omega(t) := a_n^\omega\, \alpha_n^0\, \cos\left(\sqrt{\lambda_n}\, c\, t\right) + b_n^\omega\, \alpha_n^1\, \tfrac{1}{\sqrt{\lambda_n}\, c} \sin\left(\sqrt{\lambda_n}\, c\, t\right) \\ + c^2\, \varphi_n(L)\, \tfrac{1}{\sqrt{\lambda_n}\, c} \int_0^t u(s) \sin\left(\sqrt{\lambda_n}\, c\, (t-s)\right)\, ds.\end{aligned}$$

Then

$$y^\omega(t,x) = \sum_{n=0}^{\infty} \alpha_n^\omega(t)\, \varphi_n(x) \tag{2.3}$$

is almost surely the unique solution of (P_{prob}). For all $t \in (0,T)$, we have almost surely $y^\omega(t,\cdot) \in L^2(0,L)$ and

$$\int_0^L y^\omega(t,x)^2\, dx = \sum_{n=0}^{\infty} \left(\alpha_n^\omega(t)\right)^2.$$

Moreover, we have almost surely $y^\omega \in C((0,T),\, L^2(0,L))$.

Proof. The functions φ_n are the solutions of the eigenvalue problem

$$\varphi_{xx}(x) = -\lambda\, \varphi(x),\ x \in [0,L],\ \varphi(0) = 0,\ \varphi_x(L) = 0$$

with the eigenvalues λ_n and the normalization

$$\int_0^L \varphi_n(x)^2\, dx = 1,\ n \in \{0,1,2,\ldots\}.$$

The sequence of functions $(\varphi_n)_{n=0}^\infty$ is a complete orthonormal system in the Hilbert space $L^2(0,L)$. The definition of the functions $\alpha_n^\omega(t)$ implies that almost surely $\alpha_n^\omega(0) = a_n^\omega\, \alpha_n^0$ and almost surely $(\alpha_n^\omega)'(0) = b_n^\omega\, \alpha_n^1$. Hence, we have almost surely $\sum_{n=0}^\infty \alpha_n^\omega(0)\, \varphi_n(x) = y_0^\omega(x)$ and $\sum_{n=0}^\infty (\alpha_n^\omega)'(0)\, \varphi_n(x) = y_1^\omega(x)$ almost everywhere on $(0,\, L)$. Thus, we have almost surely $y^\omega(0,\, x) = y_0^\omega(x)$ and $y_t^\omega(0,\, x) = y_1^\omega(x)$ almost everywhere on $(0,\, L)$; that is, the initial conditions of (P_{prob}) are satisfied. Now we consider the boundary traces. Almost everywhere on $(0,\, T)$, we have almost surely $y^\omega(t,\, 0) = \sum_{n=0}^\infty \alpha_n^\omega(t)\, \varphi_n(0) = 0$.

The definition of $\alpha_n^\omega(t)$ implies that for all $n \in \{0, 1, 2, \ldots\}$, we have almost surely

$$(\alpha_n^\omega)''(t) = -\frac{c^2}{L^2}\left(\frac{\pi}{2} + n\pi\right)^2 \alpha_n^\omega(t) + (-1)^n c^2 \frac{\sqrt{2}}{\sqrt{L}} u(t).$$

For a test function $\varphi \in \mathcal{D}([0, T] \times [0, L])$, we have

$$\varphi(t, x) = \sum_{n=0}^{\infty} \int_0^L \varphi(t, s)\, \varphi_n(s)\, ds\, \varphi_n(x).$$

Hence, $0 = \varphi(t, L) = \sum_{n=0}^{\infty} \int_0^L \varphi(t, s)\, \varphi_n(s)\, ds\, \varphi_n(L)$. Thus, the definition of the distributional derivative implies almost surely

$$\begin{aligned}
&\int_0^T \int_0^L \left[y_{tt}^\omega(t,x) - c^2\, y_{xx}^\omega(t,x)\right]\, \varphi(t, x)\, dx\, dt \\
&= \int_0^T \int_0^L y^\omega(t,x)\, \varphi_{tt}(t, x) - c^2\, y^\omega(t,x)\, \varphi_{xx}(t, x)\, dx\, dt \\
&= \int_0^T \int_0^L \sum_{n=0}^{\infty} \alpha_n^\omega(t)\, \varphi_n(x)\, \varphi_{tt}(t, x) - c^2\, \alpha_n^\omega(t)\, \varphi_n(x)\, \varphi_{xx}(t, x)\, dx\, dt \\
&= \sum_{n=0}^{\infty} \int_0^T \int_0^L \varphi_n(x)\, (\alpha_n^\omega)''^2\, \alpha_n^\omega(t)\, \varphi_n''(x)\, \varphi(t, x)\, dx\, dt \\
&= \sum_{n=0}^{\infty} \int_0^T \int_0^L \varphi_n(x)\, \alpha_n^\omega(t)\, \left[-c^2\, \lambda_n\, + c^2 \lambda_n\right]\, \varphi(t, x) \\
&\quad + \varphi_n(x)\, c^2 \varphi_n(L)\, u(t)\, \varphi(t, x)\, dx\, dt \\
&= \int_0^T c^2\, u(t) \sum_{n=0}^{\infty} \int_0^L \varphi(t, s)\, \varphi_n(s)\, ds\, \varphi_n(L)\, dt \\
&= \int_0^T c^2\, u(t)\, \varphi(t, L)\, dt = 0.
\end{aligned}$$

Thus, almost surely y^ω satisfies the wave equation in the sense of distributions.

Now we choose a test function $\varphi(x) \in C^2([0, L])$ such that $\varphi(0) = 0$ and $\varphi_x(L) = 0$. Then we have $\varphi(L) = \sum_{n=0}^{\infty} \int_0^L \varphi(x)\, \varphi_n(x)\, dx\, \varphi_n(L)$. Hence, we have almost surely for $t \in (0, T)$ almost everywhere

$$\begin{aligned}
&\int_0^L \left[\tfrac{1}{c^2}\, y_{tt}^\omega(t, x) - y_{xx}^\omega(t, x)\right]\, \varphi(x)\, dx \\
&= \int_0^L \tfrac{1}{c^2}\, y_{tt}^\omega(t, x)\, \varphi(x) - y^\omega(t, x)\, \varphi_{xx}(x)\, dx \\
&\quad - y_x^\omega(t, x)\, \varphi(x)|_{x=0}^L + y^\omega(t, x)\, \varphi_x(x)|_{x=0}^L \\
&= \sum_{n=0}^{\infty} \left(\tfrac{1}{c^2}\, \alpha_n''(t) + \lambda_n \alpha_n(t)\right) \int_0^L \varphi(x)\, \varphi_n(x)\, dx - y_x^\omega(t, L)\, \varphi(L) \\
&= \sum_{n=0}^{\infty} \varphi_n(L)\, u(t) \int_0^L \varphi(x)\, \varphi_n(x)\, dx - y_x^\omega(t, L)\, \varphi(L) \\
&= u(t)\, \varphi(L) - y_x^\omega(t, L)\, \varphi(L) \\
&= [u(t) - y_x^\omega(t, L)]\, \varphi(L).
\end{aligned}$$

In this weak sense, y^ω satisfies the Neumann boundary condition at $x = L$.

Hence, the series (2.3) almost surely solves (P_{prob}). Consider the classical energy

$$E^\omega(u,t) := \int_0^L y_x^\omega(t,x)^2 + \frac{1}{c^2}\,\left(y_t^\omega(t,x)\right)^2\,dx.$$

Since the energy decays almost surely for $u(t) = 0$, the uniqueness of the solution follows. The continuity of the solution with respect to time follows from the regularity of the $\alpha_n(t)$. Thus, Theorem 2.1 is proved. ∎

Remark 2.1. Note that the functions $(\frac{1}{\sqrt{\lambda_n}}\,\varphi_n')_{n=0}^\infty$ also form a complete orthonormal system in $L^2(0,L)$.

3. Optimal Neumann control. In this section, we look at the problem of optimal control from an uncertain initial state to a desired terminal position in a probabilistic sense in a given finite time for a system that is governed by the initial boundary value problem (P_{prob}). We assume for simplicity that $L = 1$. Let $T \geq 2$ be given. Let an expected initial position $y_0 \in H^1(0,1)$ with $y_0(0) = 0$ and an expected initial velocity $y_1 \in L^2(0,1)$ be given. For the convenience of the reader, we first restate the deterministic optimal control problem (N_{det}) for the case of initial states without uncertainty with an exact terminal condition that has been studied in [15]:

$$(N_{\mathsf{det}})\begin{cases} \min_{u\in L^2(0,T)}\ \|u\|^2_{L^2(0,T)} \text{ subject to} \\ \\ y(0,x) = y_0(x),\ y_t(0,x) = y_1(x),\ x \in (0,1) \\ \\ y(t,0) = 0,\ y_x(t,1) = u(t),\ t \in (0,T) \\ \\ y_{tt}(t,x) = c^2\,y_{xx}(t,x),\ (t,x) \in (0,T)\times(0,1) \\ \\ y(T,x) = 0,\ y_t(T,x) = 0,\ x \in (0,1). \end{cases}$$

The terminal constraints can also be replaced by the requirement that the terminal energy be equal to zero. For the convenience of the reader, we include the representation of the optimal control for the deterministic case (N_{det}) with wave speed $c = 1$.

THEOREM 3.1 (representation of the optimal Neumann control; see [15]).
Let $T \geq 2$, $k := \max\{n \in \mathbb{N} : 2n \leq T\}$, and $\Delta := T - 2k$. For $t \in [0,2)$, let

$$d(t) := \begin{cases} k+1, & t \in (0,\Delta], \\ k, & t \in (\Delta, 2). \end{cases} \tag{3.1}$$

Then the optimal control u_0 that solves (N_{det}) is 4-periodic, with

$$u_0(t) = \begin{cases} \frac{1}{2d(t)}\left[y_0'(1-t) - y_1(1-t)\right], & t \in (0,1), \\ \\ \frac{1}{2d(t)}\left[y_0'(t-1) + y_1(t-1)\right], & t \in (1,2). \end{cases}$$

For $t \in (0,2)$, $l \in \{0,1,\dots,k\}$ with $t + 2l \leq T$, we have $u_0(t+2l) = (-1)^l\,u_0(t)$.

Now we present a problem that is suitable for uncertain initial data. Let an accuracy parameter $\varepsilon > 0$ and a probability threshold $p \in (0,1)$ be given. We

consider the problem of optimal exact control and uncertain initial data

$$(N_{\mathsf{prob}})(\varepsilon,\, p)\begin{cases} \min_{u\in L^2(0,T)} \|u\|^2_{L^2(0,T)} \text{ subject to} \\ y^\omega(0,x) = y_0^\omega(x),\ y_t^\omega(0,x) = y_1^\omega(x),\ x\in(0,1) \\ y^\omega(t,0) = 0,\ y_x^\omega(t,1) = u(t),\, t\in(0,T) \\ y_{tt}^\omega(t,x) = c^2\, y_{xx}^\omega(t,x),\ (t,x)\in(0,T)\times(0,1) \\ \mathbb{P}\left(\ E^\omega(u,T) \le \varepsilon\ \right) \ge p. \end{cases}$$

The parameter ε is an upper bound for the terminal energy that is valid at least with the given probability p. For the terminal energy, we have

$$\begin{aligned} E^\omega(u,T) &= \int_0^L y_x^\omega(T,x)^2 + \frac{1}{c^2}\, y_t^\omega(T,x)^2 dx \\ &= \sum_{n=0}^{\infty} \left| \int_0^L y_x^\omega(T,x)\, \tfrac{1}{\sqrt{\lambda_n}}\, \varphi_n'(x)\, dx \right|^2 + \frac{1}{c^2} \left| \int_0^L y_t^\omega(T,x)\, \varphi_n(x)\, dx \right|^2 dx \\ &= \sum_{n=0}^{\infty} \lambda_n\, \left(\alpha_n^\omega(T)\right)^2 + \frac{1}{c^2}\, \left((\alpha_n^\omega)'(T)\right)^2. \end{aligned}$$

For $n\in\{1,2,3,\ldots\}$, define

$$\sigma_n^\omega(t) := a_n^\omega\, \alpha_n^0\, \cos\left(\sqrt{\lambda_n}\, c\, t\right) + b_n^\omega\, \alpha_n^1\, \tfrac{1}{\sqrt{\lambda_n}\, c} \sin\left(\sqrt{\lambda_n}\, c\, t\right)$$

and

$$c_n^{(1)}(u) := c^2\, \varphi_n(1)\, \tfrac{1}{\sqrt{\lambda_n}\, c} \int_0^T u(s) \sin\left(\sqrt{\lambda_n}\, c\, (T-s)\right)\, ds.$$

With the explicit representation of $\alpha_n^\omega(T)$ from Theorem 2.1, we obtain

$$\alpha_n^\omega(T) = c_n^{(1)}(u) + \sigma_n^\omega(T).$$

We have

$$(\sigma_n^\omega)'(t) = -a_n^\omega\, \alpha_n^0\, \sqrt{\lambda_n}\, c\, \sin\left(\sqrt{\lambda_n}\, c\, t\right) + b_n^\omega\, \alpha_n^1\, \cos\left(\sqrt{\lambda_n}\, c\, t\right).$$

Define

$$c_n^{(2)}(u) := c^2\, \varphi_n(1)\, \int_0^T u(s) \cos\left(\sqrt{\lambda_n}\, c\, (T-s)\right)\, ds.$$

Then

$$(\alpha_n^\omega)'(T) = c_n^{(2)}(u) + (\sigma_n^\omega)'(T).$$

For the terminal energy, this yields

$$E^\omega(u,\, T) = \sum_{n=0}^{\infty} \lambda_n\, \left(c_n^{(1)}(u) + \sigma_n^\omega(T)\right)^2 + \frac{1}{c^2}\, \left(c_n^{(2)}(u) + (\sigma_n^\omega)'(T)\right)^2.$$

Thus, the inequality constraint in $(N_{\mathsf{prob}})(\varepsilon,\, p)$ has the form

$$\mathbb{P}\left(\sum_{n=0}^{\infty} \lambda_n\, \left(c_n^{(1)}(u) + \sigma_n^\omega(T)\right)^2 + \frac{1}{c^2}\, \left(c_n^{(2)}(u) + (\sigma_n^\omega)'(T)\right)^2 \le \varepsilon \right) \ge p. \tag{3.2}$$

In this way, the probabilistic constraint in the definition of $(N_{\mathsf{prob}})(\varepsilon,\, p)$ demands that the probability that the control is successful in a reduction of the energy such that $E^\omega(u,\, T) \le \varepsilon$ be greater than or equal to p.

In order to make the energy accessible to numerical computations, the infinite series that defines $E^\omega(u,\, T)$ is approximated by the finite sum of the first N terms with $N \in \{1,\, 2,\, 3, \ldots\}$. Define

$$E_N^\omega(u,\, T) := \sum_{n=0}^{N} \lambda_n \left(c_n^{(1)}(u) + \sigma_n^\omega(T)\right)^2 + \frac{1}{c^2} \left(c_n^{(2)}(u) + (\sigma_n^\omega)'(T)\right)^2. \tag{3.3}$$

We obtain the problem

$$(N_{\mathsf{prob}})(\varepsilon,\, p,\, N) \begin{cases} \min_{u \in L^2(0,T)} \|u\|^2_{L^2(0,T)} \ \text{ subject to} \\ \\ \mathbb{P}\left(\, E_N^\omega(u,\, T) \le \varepsilon \,\right) \ge p. \end{cases}$$

4. Convexity of the probabilistic constraint in $(N_{\mathsf{prob}})(\varepsilon,\, p,\, N)$. In this and the following section, we shall be dealing with finite-dimensional random vectors in the context of problem $(N_{\mathsf{prob}})(\varepsilon,\, p,\, N)$ with cutoff N-term energy 3.3. Therefore, we will not have to impose our standing assumption (**A**) on identical distributions of all of its components. This assumption will become critical again when showing in section 6 the existence of solutions to the original problem $(N_{\mathsf{prob}})(\varepsilon,\, p)$, where the whole sequence of random variables comes into play.

In this section, we show that the probabilistic constraint in $(N_{\mathsf{prob}})(\varepsilon,\, p,\, N)$ defines a convex set of admissible controls u. To this aim, let ξ be an m-dimensional random vector, U a vector space, and $\tilde{g} : U \times \mathbb{R}^m \to \mathbb{R}^k$ a given mapping. The following theorem is due to Prékopa (in the original formulation, U was supposed to be finite-dimensional, but this restrictive property is not exploited in the proof of the result; see also [11, Prop. 4]):

THEOREM 4.1 ([28]). *Let ξ have a log-concave density, i.e., a density whose log is (possibly extended-valued) concave. If all components $\tilde{g}_i$ of $\tilde{g}$, $(i = 1, \ldots, k)$, are quasi-convex, then the probability function $\tilde{\phi} : U \to \mathbb{R}$ defined by*

$$\tilde{\phi}(u) := \mathbb{P}\left(\tilde{g}_i(u, \xi) \le 0 \quad (i = 1, \ldots, k)\right) \tag{4.1}$$

is log-concave.

We introduce the probability functions $\phi_N : L^2(0, T) \to \mathbb{R}$ for $N \in \mathbb{N}$ associated with the inequality constraints in $(N_{\mathsf{prob}})(\varepsilon,\, p,\, N)$ as

$$\phi_N(u) := \mathbb{P}\left(\, E_N^\omega(u, T) \le \varepsilon \,\right)$$

or, more explicitly, (see (3.3))

$$\phi_N(u) := \mathbb{P}\left(\sum_{n=0}^{N} \lambda_n \left(c_n^{(1)}(u) + \sigma_n^\omega(T)\right)^2 + \frac{1}{c^2} \left(c_n^{(2)}(u) + (\sigma_n^\omega)'(T)\right)^2 \le \varepsilon\right). \tag{4.2}$$

PROPOSITION 4.2. *Let the random vector $\xi := ((a_n, b_n))_{n=1}^N$ have a log-concave density. Then ϕ_N defined in* (4.2) *is also log-concave.*

Proof. By definition, ϕ_N can be written as $\phi_N(u) = \mathbb{P}(g(u,\xi) \leq 0)$, where, for $n = 1, \ldots, N$,

$$
\begin{aligned}
g(u,z) &:= \sum_{n=1}^{N} g_n(u,z) - \varepsilon \\
g_n(u,z) &:= \lambda_n \left(c_n^{(1)}(u) + \langle A_n, z \rangle \right)^2 + \frac{1}{c^2} \left(c_n^{(2)}(u) + \langle B_n, z \rangle \right)^2 \\
A_n &:= \Bigg((0,0), \ldots, (0,0), \\
&\qquad \left(\alpha_n^0 \cos\left(\sqrt{\lambda_n} cT \right), \frac{\alpha_n^1}{\sqrt{\lambda_n} c} \sin\left(\sqrt{\lambda_n} cT \right) \right), (0,0), \ldots, (0,0) \Bigg) \\
B_n &:= \Bigg((0,0), \ldots, (0,0), \\
&\qquad \left(-\alpha_n^0 \sqrt{\lambda_n} c \sin\left(\sqrt{\lambda_n} cT \right), \alpha_n^1 \cos\left(\sqrt{\lambda_n} cT \right) \right), (0,0), \ldots, (0,0) \Bigg) \\
z &:= \left(\left(z_n^a, z_n^b \right) \right)_{n=1}^{N},
\end{aligned}
$$

where the nonzero expressions in A_n, B_n appear at position n. Since the functions $c_n^{(1)}(u)$ and $c_n^{(2)}(u)$ are linear in u, the functions g_n are convex as squares of linear functions in (u,z) jointly. Hence, g is also convex. Now the result follows from Theorem 4.1 upon putting $\tilde{\phi} := \phi_N$ and $\tilde{g} := g$. ∎

COROLLARY 4.3. *In the setting of Proposition* 4.2, *the set of feasible controls* u *defined by the probabilistic constraint in problem* $(N_{\mathsf{prob}})(\varepsilon, p, N)$ *is convex and can be equivalently represented by a convex inequality* $h_N(u) \leq 0$, *where*

$$
h_N(u) := -\log \phi_N(u) + \log p \tag{4.3}
$$

and ϕ_N *is defined in* (4.2).

Remark 4.1. We observe that many prominent multivariate distributions share the property of having log-concave densities, e.g., multivariate Gaussian, Dirichlet, Wishart, Gamma, or the uniform distribution on convex compact sets; see [28].

5. Algorithmic approach to the solution of problem $(N_{\mathsf{prob}})(\varepsilon, p, N)$. For the algorithmic solution of the optimization problem $(N_{\mathsf{prob}})(\varepsilon, p, N)$, the numerical evaluation/approximation of the probability function ϕ_N in (4.2) along with its sensitivity with respect to the control variable u is necessary. We will describe next an approach via the so-called *spheric-radial decomposition* of Gaussian random vectors (see, e.g., [9, p. 105], [13, p. 30]). Although the same idea applies to a whole class of distributions (e.g., Gaussian-like such as lognormal, truncated Gaussian, or elliptically symmetric distributions, such as Student), we will content ourselves here with the probably most prominent case of purely Gaussian distributions. The following result is well known (see [9, eq. (8)] or in its present form [14, Thm. 3.1]).

LEMMA 5.1. *Let* $\xi \sim \mathcal{N}(\mu, \Sigma)$ *be an* m*-dimensional Gaussian random vector having expectation* μ *and (nondegenerate) covariance matrix* Σ. *Then, for any Borel-measurable subset* $M \subseteq \mathbb{R}^m$, *its probability with respect to this distribution can be represented as*

$$
\mathbb{P}(\xi \in M) = \int_{\mathbb{S}^{m-1}} \nu_\chi \{ r \geq 0 | \mu + rLv \in M \} \, d\nu_\eta(v).
$$

Here, $\mathbb{S}^{m-1}$ is the $(m-1)$-dimensional unit sphere in $\mathbb{R}^m$, ν_η is the uniform distribution on $\mathbb{S}^{m-1}$, ν_χ denotes the one-dimensional χ-distribution with m degrees of freedom, and L is such that

$$\Sigma = LL^T \tag{5.1}$$

(e.g., Cholesky decomposition).

Next, we apply the previous lemma to the probability function $\tilde{\phi}(u)$ defined in Theorem 4.1 for $k = 1$ and consider

$$\tilde{\phi}(u) := \mathbb{P}\left(\tilde{g}\left(u, \xi\right) \leq 0\right), \tag{5.2}$$

where $\tilde{g} : U \times \mathbb{R}^m \to \mathbb{R}$ for some Banach space U and ξ is a Gaussian random vector as in Lemma 5.1.

COROLLARY 5.2. *Fix any $u \in U$ with the following properties:*

1. *$\tilde{\phi}(u) > 0.5$.*
2. *$\tilde{g}(u, \cdot)$ is convex.*
3. *The convex inequality $\tilde{g}(u, \cdot) \leq 0$ admits a Slater point; i.e., there is some $\bar{z}$ with $\tilde{g}(u, \bar{z}) < 0$.*

Then,

$$\tilde{\phi}(u) = \int_{\mathbb{S}^{m-1}} e\left(u, v\right) d\nu_\eta(v), \tag{5.3}$$

where $e(u, v) = F_{\nu_\chi}(\rho(u, v))$ (with F_{ν_χ} referring to the distribution function of ν_χ) if the equation

$$\tilde{g}\left(u, \mu + rLv\right) = 0$$

admits a (unique) nonnegative solution $r = \rho(u, v) \geq 0$ and $e(u, v) = 1$ else (i.e., $\tilde{g}(u, \mu + rLv) < 0$ for all $r \geq 0$).

Proof. Lemma 5.1 yields that

$$\tilde{\phi}(u) = \int_{\mathbb{S}^{m-1}} \nu_\chi\left\{r \geq 0 | \tilde{g}(u, \mu + rLv) \leq 0\right\} d\nu_\eta(v). \tag{5.4}$$

For fixed $u \in U$, define the set $M := \{z \in \mathbb{R}^m \mid \tilde{g}(u, z) \leq 0\}$. Clearly, M is convex and nonempty by assumption 3. From our assumptions, it follows that the mean vector μ of ξ satisfies the strict inequality $\tilde{g}(u, \mu) < 0$ (see [1, Prop. 3.11]). Now, by the convexity of M, one has either that $\tilde{g}(u, \mu + rLv) < 0$ for all $r \geq 0$ or that $\tilde{g}(u, \mu + rLv) = 0$ for exactly one $r = \rho(u, v) \geq 0$. In the first case,

$$\nu_\chi\left\{r \geq 0 | \tilde{g}(u, \mu + rLv) \leq 0\right\} = \nu_\chi(\mathbb{R}_+) = 1$$

since the support of the χ-distribution is the nonnegative reals. In the second case,

$$\begin{aligned}\nu_\chi\left\{r \geq 0 | \tilde{g}(u, \mu + rLv) \leq 0\right\} &= \nu_\chi([0, \rho\left(u, v\right)]) = F_{\nu_\chi}\left(\rho\left(u, v\right)\right) - F_{\nu_\chi}\left(0\right) \\ &= F_{\nu_\chi}\left(\rho\left(u, v\right)\right).\end{aligned}$$

Now the assertion follows from (5.4). ∎

A few words on the assumptions of Corollary 5.2 are in order: The convexity assumption 2 will hold true in the case of our problem $(N_{\mathsf{prob}})(\varepsilon, p, N)$ because the function g defined in the proof of Proposition 4.2 and playing the role of $\tilde{g}$ in Corollary 5.2 was shown to be convex (actually in both arguments simultaneously) in the mentioned proof. Requiring the value of the probability function to be larger than 0.5 as in assumption 1 is no practical restriction because in probabilistic constraints, one is typically dealing with probabilities close to one. Finally, it is well known that, without the Slater point assumption 3, the probability function $\tilde{\phi}$ may even fail to be continuous. In the case of our problem $(N_{\mathsf{prob}})(\varepsilon, p, N)$, this condition is always automatically satisfied because, given a control u, we may always find an initial state which is steered to zero terminal energy by that control. Hence, the terminal energy is strictly smaller than ε, which amounts to the existence of the desired Slater point for $\tilde{g} := g$ and with g defined in the proof of Proposition 4.2.

For numerical purposes, the spheric integral in Corollary 5.2 will be approximated by a finite sum based on a sample $v^1, \ldots, v^K \in \mathbb{S}^{m-1}$ from the uniform distribution on the sphere. One possibility to do so efficiently consists in generating a quasi–Monte Carlo sample $w^1, \ldots, w^K$ of the m-dimensional standard Gaussian distribution in $\mathbb{R}^m$ and to normalize it so that $v^k := w^k/\|w^k\|$ for $k = 1, \ldots, K$. Then the desired value of the probability function in (5.2) can be approximated as

$$\tilde{\phi}(u) \approx K^{-1} \sum_{k=1}^{K} e\left(u, v^k\right). \tag{5.5}$$

Of course, the larger K, the better this approximation. It remains to clarify how the function $e(u, v)$ in Corollary 5.2 can be evaluated in the context of our optimization problem $(N_{\mathsf{prob}})(\varepsilon, p, N)$. To this end, the general function $\tilde{g}$ defined in (5.2) will be specified now as the function g introduced in the proof of Proposition 4.2. Hence, we have

$$\begin{aligned}\tilde{g}\left(u, \mu + rLv\right) &:= g\left(u, \mu + rLv\right) \\ &= \sum_{n=1}^{N} \left(\lambda_n \left(c_n^{(1)}(u) + \langle A_n, \mu + rLv \rangle \right)^2 + \frac{1}{c^2} \left(c_n^{(2)}(u) + \langle B_n, \mu + rLv \rangle \right)^2 \right) - \varepsilon.\end{aligned}$$

Now, to solve the equation $\tilde{g}(u, \mu + rLv) = 0$ for given u and v in $r \geq 0$ as mentioned in Corollary 5.2, we can regroup and rewrite it as a quadratic equation in r:

$$\alpha\left(u, v\right) r^2 + \beta\left(u, v\right) r + \gamma\left(u, v\right) = 0, \tag{5.6}$$

where

$$\begin{aligned}\alpha\left(u, v\right) &:= \sum_{n=1}^{N} \lambda_n \left\langle A_n, Lv \right\rangle^2 + \frac{1}{c^2} \left\langle B_n, Lv \right\rangle^2, \\ \beta\left(u, v\right) &:= \sum_{n=1}^{N} 2\lambda_n \left\langle A_n, Lv \right\rangle \left(\langle A_n, \mu \rangle + c_n^{(1)}(u) \right) + \frac{2}{c^2} \left\langle B_n, Lv \right\rangle \left(\langle B_n, \mu \rangle + c_n^{(2)}(u) \right), \\ \gamma\left(u, v\right) &:= \sum_{n=1}^{N} \lambda_n \left(c_n^{(1)}(u) + \langle A_n, \mu \rangle \right)^2 + \frac{1}{c^2} \left(c_n^{(2)}(u) + \langle B_n, \mu \rangle \right)^2 - \varepsilon.\end{aligned}$$

Under the assumptions of Corollary 5.2, it follows that $\tilde{g}(u, \mu) < 0$ (see proof there), which means that $\gamma(u, v) < 0$. Since, moreover, $\alpha(u, v) \geq 0$, it follows that exactly one of the following cases may occur:

(a) If $\alpha(u,v)=0$ and $\beta(u,v)\le 0$, then (5.6) has no solution $r\ge 0$ at all, and hence $e(u,v)=1$ (see Corollary 5.2).
(b) If $\alpha(u,v)=0$ and $\beta(u,v)>0$, then the unique solution of (5.6) is

$$r=\rho(u,v):=-\frac{\gamma(u,v)}{\beta(u,v)}>0.$$

(c) If $\alpha(u,v)>0$, then (5.6) has exactly one solution $r\ge 0$ given by

$$r=\rho(u,v):=\frac{-\beta(u,v)+\sqrt{\beta^2(u,v)-4\gamma(u,v)\,\alpha(u,v)}}{2\alpha(u,v)}.$$

We note that the one-dimensional distribution function F_{ν_χ} of the χ-distribution is a direct built-in function of many mathematical software packages or can be easily derived from the corresponding distribution functions of the χ^2- or Γ-distributions (for instance, $F_{\nu_\chi}(t)=F_{\nu_{\chi^2}}(t^2)$).

After having described how to evaluate fairly precise values for the probability function $\tilde{\phi}(u)$ via (5.5) (or in the context of optimization problem $(N_{\mathsf{prob}})(\varepsilon,\,p,\,N)$ the probability function $\phi_N(u)$ defined in (4.2)), we address now the question of how to obtain its derivative $D\tilde{\phi}(u)$ in order to set up a gradient-based optimization procedure for solving problem $(N_{\mathsf{prob}})(\varepsilon,\,p,\,N)$. For the following result, we refer the reader to [19, Cor. 3, Rem. 1]. We note that the result in [19] has been presented for standardized Gaussian random vectors. We formulate it here for the case of general random vectors, which follows immediately from the previous one.

THEOREM 5.3. *Let $\tilde{g}:U\times\mathbb{R}^m\to\mathbb{R}$ in (5.2) be continuously differentiable, where U is some reflexive and separable Banach space, and let $\xi\sim\mathcal{N}(\mu,\Sigma)$ be an m-dimensional Gaussian random vector. Additionally to the assumptions of Corollary 5.2, suppose that the following growth condition is satisfied around $u\in U$:*

$$\exists l>0:\|\nabla_u\tilde{g}(w,z)\|\le l e^{\|\mu+\Delta z\|}\quad\forall w:\|w-u\|\le 1/l\ \forall z:\|z\|\ge l,$$

where Δ is a diagonal matrix with diagonal elements $\Delta_i:=\operatorname{diag}\sqrt{\Sigma_{i,i}}$. Then the probability function $\tilde{\phi}$ from (5.2) is (strictly) differentiable at u, and its gradient $\nabla\tilde{\phi}(u)$ is given as

$$\nabla\tilde{\phi}(u)=-\int_{v\in F(u)}\frac{\chi(\rho(u,v))\cdot\nabla_u\tilde{g}(u,\mu+\rho(u,v)\,Lv)}{\langle\nabla_z\tilde{g}(u,\mu+\rho(u,v)\,Lv),Lv\rangle}d\nu_\eta(v). \tag{5.7}$$

Here,

$$\chi(t):=\frac{t^{m-1}e^{-t^2/2}}{2^{m/2-1}\Gamma(k/2)}\quad(t\ge 0)$$

is the density of the one-dimensional χ-distribution with m degrees of freedom. Moreover, $F(u)$ refers to the set of directions $v\in\mathbb{S}^{m-1}$ such that the equation

$$\tilde{g}(u,\mu+rLv)=0$$

admits a (unique) nonnegative solution $r=\rho(u,v)\ge 0$ (see Cor. 5.2).

The preceding theorem can now be applied to the probability function of our problem $(N_{\mathsf{prob}})(\varepsilon,\,p,\,N)$:

THEOREM 5.4. *Let the random vector* $\xi := (a_n, b_n)_{n=1}^N$ *with coefficients introduced in* (A) *have a* $2N$*-dimensional Gaussian distribution* $\xi \sim \mathcal{N}(\mu, \Sigma)$. *Consider any* $u \in L^2(0,T)$ *which is feasible in* $(N_{\mathsf{prob}})(\varepsilon, p, N)$ *for some* $p > 0.5$. *Then the probability function* $\phi_N(u) := \mathbb{P}(\ E_N^\omega(u) \leq \varepsilon\ )$ *in* (4.2) *is (strictly) differentiable at* u, *and its gradient is given by*

$$\nabla\phi_N(u) = -\int_{v\in F(u)} \frac{\chi\left(\rho\left(u,v\right)\right)\cdot \nabla_u g\left(u, \mu+\rho\left(u,v\right)Lv\right)}{\left\langle \nabla_z g\left(u,\mu+\rho\left(u,v\right)Lv\right), Lv\right\rangle} d\nu_\eta(v), \tag{5.8}$$

where, with the definitions of A_n, B_n *in the proof of Proposition* 4.2,

$$\nabla_u g(u,z) = 2\sum_{n=1}^N \left\{ \lambda_n\left(c_n^{(1)}(u) + \langle A_n, z\rangle\right)\nabla c_n^{(1)}(u) + \frac{c_n^{(2)}(u) + \langle B_n, z\rangle}{c^2}\nabla c_n^{(2)}(u)\right\}, \tag{5.9}$$

$$\nabla_z g(u,z) = 2\sum_{n=1}^N \left\{ \lambda_n\left(c_n^{(1)}(u) + \langle A_n, z\rangle\right) A_n + \frac{c_n^{(2)}(u) + \langle B_n, z\rangle}{c^2} B_n\right\}. \tag{5.10}$$

Proof. Clearly, the function g introduced in Proposition 4.2 is continuously differentiable with partial gradients as in (5.9), (5.10). By virtue of our assumptions, we obtain that $\phi_N(u) \geq p > 0.5$, and hence, recalling the remarks below the proof of Corollary 5.2, all assumptions of that corollary are satisfied for $\tilde{g} := g$ and $\tilde{\phi} := \phi_N$. As for the growth condition in Theorem 5.3, we observe that

$$\begin{aligned}\|\nabla_u g(v,z)\| \leq 2\sum_{n=1}^N \Bigg\{ &\lambda_n \left\|\nabla c_n^{(1)}(v)\right\| \left(\left|c_n^{(1)}(v)\right| + \|A_n\|\,\|\mu + \Delta z\|\right) \\ &+ \frac{1}{c^2}\left\|\nabla c_n^{(2)}(v)\right\| \left(\left|c_n^{(2)}(v)\right| + \|B_n\|\,\|\mu+\Delta z\|\right)\Bigg\}.\end{aligned}$$

By linearity of $c_n^{(1)}$, $c_n^{(2)}$, the norms $\|\nabla c_n^{(1)}(v)\|$, $\|\nabla c_n^{(2)}(v)\|$ are constants, and moreover there exists some $\varkappa, \delta > 0$ such that

$$\max\left\{\left|c_n^{(1)}(v)\right|, \left|c_n^{(2)}(v)\right|\right\} \leq \varkappa \quad \forall v : \|v-u\| \leq \delta.$$

This allows us to derive an estimate of the type

$$\|\nabla_u g(v,z)\| \leq K_1 + K_2\,\|\mu + \Delta z\| \quad \forall v : \|v-u\| \leq \delta \quad \forall z,$$

which clearly entails the growth condition in Theorem 5.3. Summarizing, we observe that all assumptions of Theorem 5.3 (including those of Corollary 5.2) are fulfilled for $\tilde{g} := g$ and $\tilde{\phi} := \phi_N$. ∎

We recall that the functions $c_n^{(1)}, c_n^{(2)}$ occurring in the formulas (5.9), (5.10) are defined below problem $(N_{\mathsf{prob}})(\varepsilon, p)$. In particular, one calculates their gradients occurring in (5.9) as

$$\nabla c_n^{(1)}(u)\,(s) = \frac{c}{\sqrt{\lambda_n}}\,\varphi_n(1)\ \sin\left(\sqrt{\lambda_n}\,c\,(T-s)\right) \quad (s \in (0,T)), \tag{5.11}$$

$$\nabla c_n^{(2)}(u)\,(s) = c^2\,\varphi_n(1)\ \cos\left(\sqrt{\lambda_n}\,c\,(T-s)\right) \quad (s \in (0,T)). \tag{5.12}$$

The results presented in this section suggest an algorithmic scheme for determining approximations of the probability function $\phi_N(u) = \mathbb{P}(\, E_N^\omega(u) \le \varepsilon\,)$ and its gradients $\nabla\phi_N(u)$ in the iterative solution of our optimization problem $(N_{\mathsf{prob}})(\varepsilon, p, N)$ at a given iterate $u \in L^2(0,T)$. The idea is to approximate the gradient $\nabla\phi_N(u)$ by its values on a finite subset $\{t_1, \dots, t_M\} \subseteq (0,T)$. In this way, a gradient-based solution algorithm for $(N_{\mathsf{prob}})(\varepsilon, p, N)$ (e.g., projected gradients) is easily set up. Observe that, thanks to Corollary 5.2 and Theorem 5.4, both the value $\phi_N(u)$ and its gradient $\nabla\phi_N(u)$ are represented as spherical integrals (in the latter case, one reduces the sphere to its subset $v \in F(u)$ by a simple check of $\rho(u,v) < \infty$). Hence, the same sample $v \in \mathbb{S}^{m-1}$ can be used in order to update both $\phi_N(u)$ and $\nabla\phi_N(u)$. Here, one takes advantage of the fact that the value $\rho(u,v)$ has only to be determined once. The following algorithm assigns to a given iterate $u \in L^2(0,T)$ approximations for $\phi_N(u)$ and for $\nabla\phi_N(u)$ on a given grid $\{t_1, \dots, t_M\}$ under the given Gaussian distribution $(a_n, b_n)_{n=1}^N \sim \mathcal{N}(\mu, \Sigma)$ of the coefficients in (2.2).

ALGORITHM 5.1.

1. *Generate a (quasi–Monte Carlo) sample $v^1, \dots, v^K$ of the m-dimensional standard Gaussian distribution $\mathcal{N}(0, I)$.*
2. *Find a Cholesky decomposition $\Sigma = LL^T$ for the covariance matrix of the given distribution of coefficients.*
3. *Initialize the desired approximation s_v for $\phi_N(u)$ and s_g for $\nabla\phi_N(u)$ on the grid $\{t_1, \dots, t_M\}$ by $s_v := 0$ and $s_g(t_i) := 0$ for $i = 1, \dots, M$. Initialize the iteration counter for the sample in step 1 as $k := 1$.*
4. *For the given iterate u and the sampled direction $v := v^k \in \mathbb{S}$, check for and identify the solution of (5.6). Compute values $\alpha(u, v^k)$, $\beta(u, v^k)$, $\gamma(u, v^k)$ as indicated below (5.6). Update the contribution of sample v^k to the discretized versions of the spherical integrals (5.3) and (5.8) according to the case distinction made above:*
   (a) *If $\alpha(u, v^k) = 0$ and $\beta(u, v^k) \le 0$ (i.e., (5.6) has no solution, from which $e(u, v^k) = 1$ and $v^k \notin F(u)$), then update $s_v := s_v + 1$.*
   (b) *If $\alpha(u, v^k) = 0$ and $\beta(u, v^k) > 0$, then put*

$$\rho\left(u, v^k\right) := -\frac{\gamma\left(u, v^k\right)}{\beta\left(u, v^k\right)}.$$

   (c) *If $\alpha(u, v^k) > 0$, then put*

$$\rho\left(u, v^k\right) := \frac{-\beta\left(u, v^k\right) + \sqrt{\beta^2\left(u, v^k\right) - 4\gamma\left(u, v^k\right)\alpha\left(u, v^k\right)}}{2\alpha\left(u, v^k\right)}.$$

   In both cases (b) and (c), $v^k \in F(u)$. Update $s_v := s_v + F_{\nu_\chi}(\rho(u, v^k))$ and for $(i = 1, \dots, M)$ (see (5.8))

$$s_g\left(t_i\right) := s_g\left(t_i\right) + \frac{\chi\left(\rho\left(u, v^k\right)\right) \cdot \nabla_u g\left(u, \mu + \rho\left(u, v^k\right) Lv^k\right)\left(t_i\right)}{\left\langle \nabla_z g\left(u, \mu + \rho\left(u, v^k\right) Lv^k\right), Lv^k \right\rangle}.$$

   In this last formula, use the representations (5.9), (5.10). Referring to (5.11) and (5.12), we obtain, for instance, the following fully explicit representation

for $\nabla_u g$, required above:

$$\begin{aligned}
&\nabla_u g\left(u, \mu{+}\rho\left(u, v^k\right) Lv^k\right)(t_i) \\
&\quad = 2c \sum_{n=1}^{N} \left(\frac{c\varphi_n(1)}{\sqrt{\lambda_n}} \int_0^T u(s) \sin\left(\sqrt{\lambda_n}\, c\,(T-s)\right) ds \right. \\
&\qquad\qquad \left. + \left\langle A_n, \mu{+}\rho\left(u, v^k\right) Lv^k \right\rangle \right) \\
&\qquad \cdot \sqrt{\lambda_n}\, \varphi_n(1) \sin\left(\sqrt{\lambda_n}\, c\,(T-t_i)\right) \\
&\qquad + 2 \sum_{n=1}^{N} \left((c^2\, \varphi_n(1) \int_0^T u(s) \cos\left(\sqrt{\lambda_n}\, c\,(T-s)\right) ds \right. \\
&\qquad\qquad \left. + \left\langle B_n, \mu{+}\rho\left(u, v^k\right) Lv^k \right\rangle \right) \\
&\qquad \cdot \varphi_n(1) \cos\left(\sqrt{\lambda_n}\, c\,(T-t_i)\right).
\end{aligned}$$

5. *If $k < K$, then $k := k+1$ and go to step* 4.
6. *STOP with $\phi_N(u) \approx K^{-1} s_v$ and $\nabla \phi_N(u)(t_i) \approx K^{-1} s_g(t_i)$ for $i = 1, \ldots, M$.*

6. Existence of solutions. In this section, we show that for all natural numbers $N \in \{1, 2, 3, \ldots\}$ and $\varepsilon > 0$, a solution of $(N_{\mathsf{prob}})(\varepsilon, p, N)$ exists if p is sufficiently small. For a given $\varepsilon > 0$, we set

$$p_{\sup}(\varepsilon, N) := \sup\left\{ p \in [0,1] \mid \exists u \in L^2(0,T) \text{ s.t. } \mathbb{P}\left(E_N^{\omega}(u, T) \le \varepsilon \right) \ge p \right\},$$

which is decreasing with respect to N. Hence, for all $p < p_{\sup}(\varepsilon, N)$, there exists $u \in L^2(0, T)$ such that $\mathbb{P}(E_N^{\omega}(u, T) \le \varepsilon) \ge p$, which means that the feasible set of the optimal control problem $(N_{\mathsf{prob}})(\varepsilon, p, N)$ is nonempty. This fact is used to prove the following lemma, which ensures the existence of a unique solution to $(N_{\mathsf{prob}})(\varepsilon, p, N)$.

LEMMA 6.1. *Let $N \in \{1, 2, 3, \ldots\}$ be given. Assume that $p \in [0, p_{sup}(\varepsilon, N))$. Then $(N_{\mathsf{prob}})(\varepsilon, p, N)$ has a unique solution $u \in L^2(0, T)$.*

Proof. With the convex function h_N as defined in (4.3), we can state problem $(N_{\mathsf{prob}})(\varepsilon, p, N)$ in the form

$$\min_{u \in L^2(0,T)} \|u\|^2_{L^2(0,T)} \text{ subject to } h_N(u) \le 0.$$

Then the direct method of the calculus of variations yields the solution as the weak limit point of a minimizing sequence. The strong convexity of the objective function implies the uniqueness. ∎

In preparation of the following lemma, let $\nu(\varepsilon, p, N)$ and $\nu(\varepsilon, p)$ denote the optimal value of $(N_{\mathsf{prob}})(\varepsilon, p, N)$ and $(N_{\mathsf{prob}})(\varepsilon, p)$, respectively.

LEMMA 6.2. *Assume that $p \in [0, \inf_N p_{sup}(\varepsilon, N))$ and $\nu(\varepsilon, p) < \infty$. Then the sequence of solutions $u^*(\varepsilon, p, N)$ of $(N_{\mathsf{prob}})(\varepsilon, p, N)$ $(N \in \{1, 2, 3, \ldots\})$ contains a subsequence that converges strongly in $L^2(0, T)$.*

Proof. Since $E_N^{\omega}(u, T) \le E_{N+1}^{\omega}(u, T) \le \cdots \le E^{\omega}(u, T)$, we have

$$\nu(\varepsilon, p, N) \le \nu(\varepsilon, p) \text{ for all } N \in \{1, 2, 3, \ldots\}.$$

Moreover, the sequence $(\nu(\varepsilon, p, N))_{N\in\mathbb{N}}$ is increasing. Let $\tilde{L} := \lim_{N\to\infty} \nu(\varepsilon, p, N)$. Then $\tilde{L} \le \nu(\varepsilon, p)$.

The assumption $\nu(\varepsilon, p) < \infty$ implies that the sequence $(u^*(\varepsilon, p, N))_{N\in\mathbb{N}}$ is bounded. Hence, there exists a weakly convergent subsequence with a weak limit $\tilde{u}$. We have $\tilde{L} \ge \|\tilde{u}\|^2_{L^2(0,L)}$. For all $N \in \{1, 2, 3, \ldots\}$, we have, for all $M \ge N$,

$$h_N(u^*(\varepsilon, p, M)) \le 0.$$

Since h_N is sequentially weakly lower semicontinuous, see [11, Prop. 1], we obtain for all $N \in \{1, 2, 3, \ldots\}$

$$h_N(\tilde{u}) \le \liminf_{M\to\infty} h_N(u^*(\varepsilon, p, M)) \le 0.$$

This implies that $\tilde{u}$ is feasible for $(N_{\mathsf{prob}})(\varepsilon, p, N)$ for all $N \in \{1, 2, 3, \ldots\}$. But $\nu(\varepsilon, p, N) \le \|\tilde{u}\|^2_{L^2(0,T)}$, and hence $\tilde{L} \le \|\tilde{u}\|^2_{L^2(0,T)}$. Thus, we have $\tilde{L} = \|\tilde{u}\|^2_{L^2(0,T)}$. This implies in turn the strong convergence of the subsequence in $L^2(0, T)$ to $\tilde{u}$. ∎

In the following theorem, we show that each strong limit point $\tilde{u}$ of the sequence of solutions $(u^*(\varepsilon, p, N))_{N\in\mathbb{N}}$ of the problems $(N_{\mathsf{prob}})(\varepsilon, p, N)$ is feasible for $(N_{\mathsf{prob}})(\varepsilon, p)$.

THEOREM 6.3. *Under the assumptions of Lemma* 6.2 *and under our standing assumption* **(A)**, $(N_{\mathsf{prob}})(\varepsilon, p)$ *has a solution* $\tilde{u}$ *that is a strong accumulation point of the sequence* $(u^*(\varepsilon, p, N))_{N\in\mathbb{N}}$ *given by Lemma* 6.2. *Moreover, we have*

$$\lim_{N\to\infty} \nu(\varepsilon, p, N) = \nu(\varepsilon, p). \tag{6.1}$$

Proof. Let $\tilde{u}$ be as in the proof of Lemma 6.2. Note that we have $\|\tilde{u}\|^2_{L^2(0,T)} \le \nu(\varepsilon, p)$. Since $\tilde{u}$ is feasible for $(N_{\mathsf{prob}})(\varepsilon, p, N)$ for all $N \in \{1, 2, 3, \ldots\}$, we have, for all $N \in \{1, 2, 3, \ldots\}$,

$$\mathbb{P}\left(E^\omega_N(\tilde{u}, T) \le \varepsilon\right) \ge p. \tag{6.2}$$

Define the random variable $\delta^\omega_N := E^\omega(\tilde{u}, T) - E^\omega_N(\tilde{u}, T) \le 0$. We obtain that $\delta^\omega_{N+1} \le \delta^\omega_N$. In addition, as a consequence of our standing assumption **(A)**, we have almost surely $\lim_{N\to\infty} \delta^\omega_N = 0$. Moreover, we have

$$\mathbb{P}(E^\omega_N(\tilde{u}, T) \le \varepsilon) = \mathbb{P}(E^\omega(\tilde{u}) \le \varepsilon + \delta^\omega_N).$$

Consider the sets

$$S_N := \{\omega \in \Omega \mid E^\omega(\tilde{u}, T) \le \varepsilon + \delta^\omega_N\} \text{ and } \cap_{N=1}^\infty S_N = \{\omega \in \Omega \mid E^\omega(\tilde{u}) \le \varepsilon\}.$$

Then we have $S_{N+1} \subset S_N$, and due to (6.2), we have $\mathbb{P}(S_N) \ge p$. Define the set

$$S := \{\omega \in \Omega \mid E^\omega(\tilde{u}, T) \le \varepsilon\} \cup \left\{\omega \in \Omega \middle| \lim_{N\to\infty} \delta^\omega_N \ne 0\right\},$$

where $\mathbb{P}(\{\omega \in \Omega \mid \lim_{N\to\infty} \delta^\omega_N \ne 0\}) = 0$. Then $\mathbb{P}(S) = \lim_{N\to\infty} \mathbb{P}(S_N)$. Thus, we have shown that

$$\mathbb{P}(E^\omega(\tilde{u}) \le \varepsilon) \ge p.$$

Hence, $\tilde{u}$ is feasible for $(N_{\mathsf{prob}})(\varepsilon, p)$ and hence also a solution of $(N_{\mathsf{prob}})(\varepsilon, p)$. Since the arguments in the proof of Lemma 6.2 and Theorem 6.3 can be applied to any weak limit point of the sequence $(u^*(\varepsilon, p, N))_{N\in\mathbb{N}}$, this implies that in fact every accumulation point of the sequence $(u^*(\varepsilon, p, N))_{N\in\mathbb{N}}$ is a solution of $(N_{\mathsf{prob}})(\varepsilon, p)$. Moreover, we have (6.1). ∎

7. Numerical solution of two examples. In this section, we discuss the numerical solution of two examples with different expected initial conditions for problem $(N_{\mathsf{prob}})(\varepsilon,\, p)$. Referring to the explicit description of that problem as well as to (2.2) and (2.1), we consider the following problem data:

$$c := 1;\ \varepsilon := 0.1;\ T := 4;\ p \in \{0.10, 0.15, 0.20 \ldots, p^{\max}\}$$
$$a_n \sim \mathcal{N}(1, 0.2)\ (n \in \mathbb{N});\ a_n \text{ pairwise uncorrelated};\ b_n := 0\ (n \in \mathbb{N});$$
$$y_0(x) := x \text{ (example 1)};\ y_0(x) := \pi^{-1}\sin(\pi x) \text{ (example 2)};\ y_1(x) := 0 \quad (x \in (0,1)).$$

The coefficients α_n^0 and α_n^1 in (2.2) and (2.1), respectively, are obtained as the Fourier coefficients of the chosen functions $y_0(x), y_1(x)$. In particular, $\alpha_n^1 = 0$ for all n. The latter implies that the (formal) multiplicative random coefficients b_n for perturbing $y_1(x)$ can be chosen arbitrarily without any effect. As for the coefficients a_n, they all follow an identical Gaussian distribution (with mean 1 and standard deviation 0.2) in order to satisfy our standing assumption (**A**). This allows us to apply all the existence and convergence results of section 6 to our examples. However, as pointed out earlier, this assumption is not necessary for the numerical solution of the approximating problem $(N_{\mathsf{prob}})(\varepsilon,\, p)$. Moreover, assuming all coefficients to be pairwise uncorrelated is of absolutely no importance (recall Algorithm 5.1 allowing for correlated components of the Gaussian random vector) for either the theory or the numerical solution and is just due to a lack of significant information about correlations here. With the a_n having expectation 1, it is ensured that the expected initial value coincides with the nominal one, i.e., $\mathbb{E}y_0^\omega(x) = y_0(x)$ for all $x \in (0,1)$.

In order to deal with problem $(N_{\mathsf{prob}})(\varepsilon,\, p)$ numerically, one has to pass to finite-dimensional approximations on two sides simultaneously: First, the series expansion for the terminal energy has to be cut after N terms, which leads us to the consideration of problem $(N_{\mathsf{prob}})(\varepsilon,\, p,\, N)$. Second, we compute approximations for the optimal controls $u \in L^2(0,\, T)$ in an M-dimensional space of piecewise constant functions. Let a grid $0 = t_0 < t_1 < t_2 < \cdots < t_M = T$ be given. For $i \in \{1, \ldots, M\}$, let

$$v_j(t) := \begin{cases} 1 & \text{if } t \in [t_{j-1}, t_j), \\ 0 & \text{elsewhere,} \end{cases}$$

and define the finite-dimensional space $X_M(T)$ by

$$X_M(T) := \operatorname{span}\{v_j(\cdot) : j = 1, \ldots, M\}.$$

For any $u \in X_M(T)$, we use the representation

$$u(t) = \sum_{j=1}^{M} u(t_{j-1}) v_j(t) \qquad t \in [0, T).$$

Hence, we are finally led to solve the problem

$$(N_{\mathsf{prob}})(\varepsilon,\, p,\, N,\, M) \begin{cases} \min_{u \in X_M(T)} \sum_{j=1}^{M} (t_j - t_{j-1})\, u(t_j)^2 \text{ subject to} \\ \phi_N(u) \geq p. \end{cases}$$

The solution will be based on a projected gradient method using values and gradients of ϕ_N computed by means of Algorithm 5.1. Apart from explicit expressions occurring in these computations, we have to specify (see (5.8), (5.9), (5.10)) the integrals defining

the functions $c_n^1(u), c_n^2(u)$ introduced below problem $(N_{\mathsf{prob}})(\varepsilon, p)$. By elementary calculus, one obtains that for $u \in X_M(T)$ and for all $n \in \mathbb{N}$,

$$\int_0^T u(s) \sin\left(\sqrt{\lambda_n}(T-s)\right) ds$$
$$= \frac{1}{\sqrt{\lambda_n}} \left[-\cos\left(\sqrt{\lambda_n} T\right) u(t_0) + \sum_{j=1}^{M-1} (u(t_{j-1}) - u(t_j)) \cos\left(\sqrt{\lambda_n}(T-t_j)\right) + u(t_M) \right],$$
$$\int_0^T u(s) \cos\left(\sqrt{\lambda_n}(T-s)\right) ds$$
$$= \frac{1}{\sqrt{\lambda_n}} \left[\sin\left(\sqrt{\lambda_n} T\right) u(t_0) - \sum_{j=1}^{M-1} (u(t_{j-1}) - u(t_j)) \sin\left(\sqrt{\lambda_n}(T-t_j)\right) \right].$$

Before we proceed, we present a further analytical result in order to clarify the relation between the discretized problem $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$ and the original problem $(N_{\mathsf{prob}})(\varepsilon, p)$. In the analysis, the semidiscretized problem $(N_{\mathsf{prob}})(\varepsilon, p, N)$, where the control space is still infinite dimensional, appears as an intermediate auxiliary problem.

For our convergence analysis, we assume that for each $M \in \{1, 2, 3, \ldots\}$, the grids in the definition of the space $X_M(T)$ are chosen in such a way that for all $\tilde{u} \in L^2(0, T)$, there exists a sequence $(u_M)_M$ with $u_M \in X_M(T)$ for all $M \in \{1, 2, 3, \ldots\}$ and $\lim_{M\to\infty} \|u_M - \hat{u}\|_{L^2(0,T)} = 0$. Then we have the following lemma.

LEMMA 7.1. *Let $\nu(\varepsilon, p, N, M)$ denote the optimal value of $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$. Let the assumptions of Theorem* 6.3 *hold. Then we have, for all $M \in \{1, 2, 3, \ldots\}$,*

$$\liminf_{N\to\infty} \nu(\varepsilon, p, N, M) \geq \nu(\varepsilon, p). \tag{7.1}$$

Assume in addition that the Slater condition holds in the sense that for all sufficiently large $N \in \{1, 2, 3, \ldots\}$ (say for $N \geq N_0$), there exists a Slater point $u_S \in L^2(0, T)$ such that we have $h_N(u_S) < 0$. (This is the case if p is chosen sufficiently small.)

Then for all $N \geq N_0$ and all sufficiently large values of M, the discretized problem $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$ has a unique solution. Let $u^*(\varepsilon, p, N, M)$ denote the solution of $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$. Then for all $N \geq N_0$, we have

$$\limsup_{M\to\infty} \nu(\varepsilon, p, N, M) \leq \nu(\varepsilon, p). \tag{7.2}$$

For a sequence $(u^*(\varepsilon, p, N_k, M_k))_k$, where $\lim_{k\to\infty} M_k = \lim_{k\to\infty} N_k = \infty$, we have

$$\lim_{k\to\infty} \nu(\varepsilon, p, N_k, M_k) = \nu(\varepsilon, p), \tag{7.3}$$

and there exists a strong accumulation point of the sequence $(u^*(\varepsilon, p, N_k, M_k))_k$ that is a solution of $(N_{\mathsf{prob}})(\varepsilon, p)$.

Proof. Since $X_M(T) \subset L^2(0, T)$, we have $\nu(\varepsilon, p, N, M) \geq \nu(\varepsilon, p, N)$. Due to (6.1) for all $\delta > 0$, there exists a number N_1 such that for all $N \geq N_1$, we have $\nu(\varepsilon, p, N) \geq \nu(\varepsilon, p) - \delta$. Hence, for all $N \geq N_1$, we have $\nu(\varepsilon, p, N, M) \geq \nu(\varepsilon, p) - \delta$. Since δ can be chosen arbitrarily small, this implies (7.1).

Let $N \geq N_0$ be given. Then we can choose a sequence $(u_S^M)_M$ such that for all $M \in \{1, 2, 3, \ldots\}$, we have $u_S^M \in X_M(T)$ and $\lim_{M\to\infty} \|u_S^M - u_S\|_{2(0,T)} = 0$;

[11, Prop. 2] implies that h_N is sequentially weakly upper semicontinuous. Hence, we have $\limsup_{M\to\infty} h_N(u_S^M) \le h_N(u_S) < 0$. This implies that for M sufficiently large, we have $h_N(u_S^M) < 0$. Hence, for M sufficiently large, u_S^M is feasible for the finite-dimensional convex optimization problem $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$. This implies that $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$ has a solution. Due to the strict convexity of the objective function, this solution is uniquely determined.

Let $\tilde{u}$ denote a solution of $(N_{\mathsf{prob}})(\varepsilon, p)$; such a solution exists according to Theorem 6.3. The definition of E_N^ω as a truncation of E^ω implies that we have $h_N(\tilde{u}) \le 0$. We can choose a sequence $(u^M)_M$ such that for all $M \in \{1, 2, 3, \ldots\}$, we have $u^M \in X_M(T)$ and $\lim_{M\to\infty} \|u^M - \hat{u}\|_{2(0,T)} = 0$. Similarly as above, [11, Prop. 2] implies that we have $\limsup_{M\to\infty} h_N(u^M) \le h_N(\hat{u}) \le 0$. We define λ_M as follows: If $h_N(u_M) \le 0$, let $\lambda_M = 0$. If $h_N(u_M) > 0$, let

$$\lambda_M = \frac{h_N(u_M)}{|h_N(u_S^M)| + h_N(u_M)} \in (0, 1).$$

Then $\lim_{\to\infty} \lambda_M = 0$. Define the modified sequence $(u_M^{(F)})_M$:

$$u_M^{(F)} = (1 - \lambda_M)u_M + \lambda_M u_S^M.$$

Then we have $u_M^{(F)} \in X_M(T)$, $\lim_{M\to\infty} \|u_M^{(F)} - \hat{u}\|_{L^2(0,T)} = 0$, and the definition of λ_M and the convexity of h_N imply the inequality $h_N(u_M^{(F)}) \le 0$. Hence, $u_M^{(F)}$ is feasible for $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$. Thus, we have $\nu(\varepsilon, p, N, M) \le \|u_M^{(F)}\|^2_{L^2(0,T)}$. This implies that

$$\limsup_{M\to\infty} \nu(\varepsilon, p, N, M) \le \limsup_{M\to\infty} \|u_M^{(F)}\|^2_{L^2(0,T)} = \|\hat{u}\|^2_{L^2(0,T)} = \nu(\varepsilon, p),$$

and (7.2) follows. Now (7.1) and (7.2) imply (7.3).

The remaining parts of the assertion follow as in the proof of Theorem 6.3. ∎

Remark 7.1. Lemma 7.1 indicates that in order to approximate a solution of $(N_{\mathsf{prob}})(\varepsilon, p)$, both discretization parameters N and M in $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$ should be increased simultaneously. This is also illustrated in the numerical examples; see Figure 3 and the corresponding comments below.

7.1. First example. Here, we assume that the nominal (unperturbed, expected) initial state is given by $y_0(x) = x$. With the concrete problem data as specified above, it follows from Theorem 3.1 that the optimal deterministic control in problem (N_{det}) is the bang-bang control

$$u^*(t) = 1/4 \quad \forall t \in (0, 2); \quad u^*(t) = -1/4 \quad \forall t \in (2, 4). \tag{7.4}$$

This solution takes a nominal deterministic initial state $y_0(x) = x$ and $y_1(x) = 0$ to a position of rest, i.e., a terminal state with zero energy, within the time $T = 4$. In contrast to the deterministic case, for uncertain initial values, this same optimal control will no longer take the string to rest (apart from the unlikely event that the uncertain initial value coincides with the nominal or expected initial value). Instead, we search for a control which takes the initial state with a certain sufficiently high probability to a terminal state with an energy level below the chosen value of $\varepsilon = 0.1$. In order to approximate the theoretical solution of problem $(N_{\mathsf{prob}})(\varepsilon, p)$, we solved problem $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$ for $N = 100$ and $M = 256$ with the probability level p

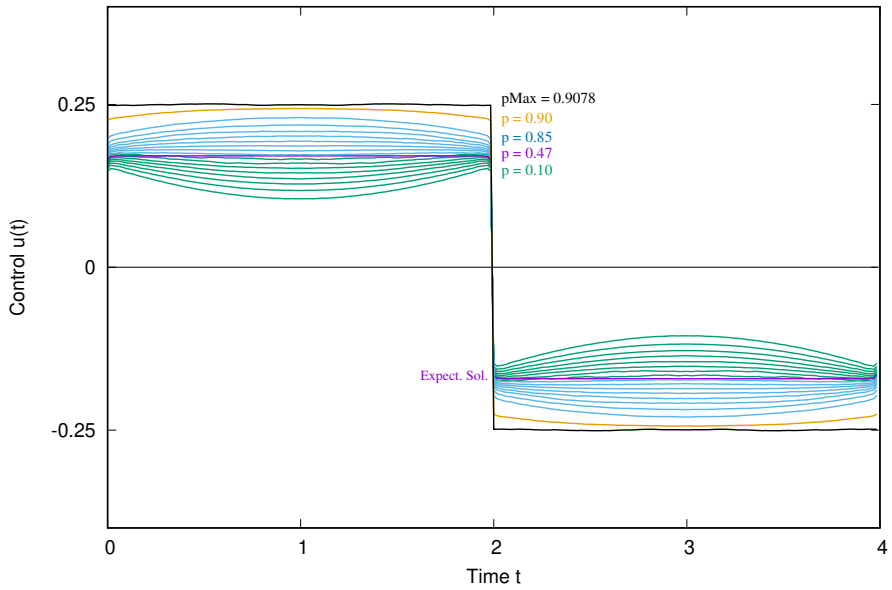

FIG. 1. *The figure shows the optimal controls under probabilistic terminal energy constraint as solutions of problem* $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$ *for a tolerance of* $\varepsilon = 0.1$ *and for different probability levels* p. *The two bang-bang controls in the figure refer to the optimal solution of the deterministic problem* (N_{det}) *(zero terminal energy) (larger amplitude) and to the solution of the expected value counterpart of problem* $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$ *(lower amplitude; see text).*

varying from $p = 0.1$ to the maximum possible value $p = 0.9078$ in steps of 0.05. The corresponding optimal controls are illustrated in Figure 1.

Not surprisingly, unlike the piecewise constant deterministic solution (7.4) of problem (N_{det}) (largest control in Figure 1), these controls are quite nonlinear due to the presence of the probabilistic constraint, and they are increasing in amplitude for increasing probability level p. Moreover, they exhibit the same symmetry patterns as the deterministic controls. More surprisingly, the profiles change their shape from convex-like to concave-like when passing a certain medium probability level $p \approx 0.47$. At that probability, the optimal control is bang-bang again but with lower amplitude than the solution of (N_{det}). It turns out that this is the solution of problem $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$ when replacing the probabilistic constraint $\phi_N(u) \geq p$ by the constraint

$$\int_0^L y_x(T,x)^2 + \frac{1}{c^2}\, y_t(T,x)^2 dx \leq \varepsilon, \tag{7.5}$$

which bounds the terminal energy of the nominal (expected) initial state by ε (therefore, the control can afford a lower amplitude than that of (N_{det}), which would correspond to the stricter bound $\varepsilon = 0$).

Even less evident is the fact that the largest possible probability $p = 0.9078$ is achieved again by a bang-bang solution which is exactly the one of the deterministic problem (N_{det}) imposing zero terminal energy while starting with the nominal initial state. Beyond that maximum probability, the feasible set of $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$ becomes empty, and costs jump to infinity.

In order to illustrate the effect of the calculated solutions, we simulate a sample of 10 random scenarios for the initial state around the nominal (expected) initial

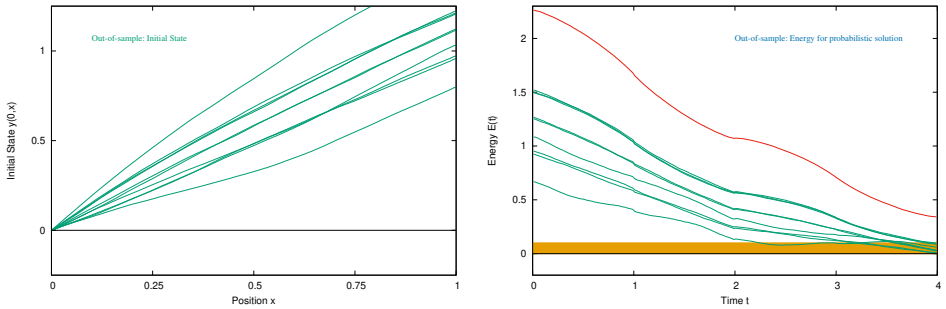

FIG. 2. *Illustration of* 10 *simulated scenarios for the initial state (left diagram) and corresponding evolution of energy over time.*

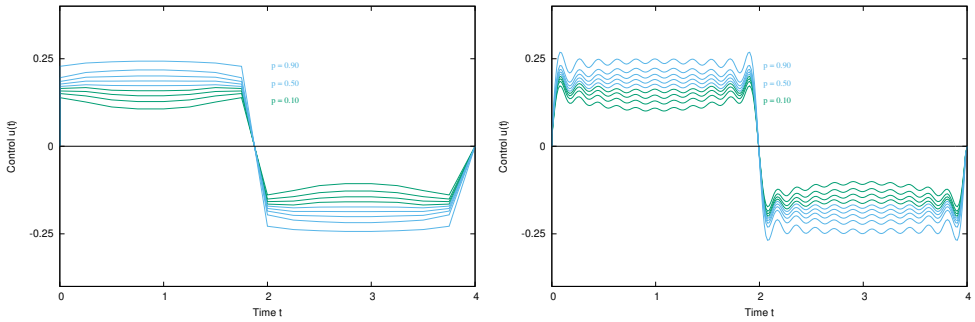

FIG. 3. *The figure show the optimal control for the discretized problem with the energy* E_N *with* $N = 10$. *The control grid is equidistant with* $M = 16$ *(left) and* $M = 256$ *points, respectively.*

state $y_0(x) = x$ according to the chosen Gaussian distribution of the multiplicative perturbations a_n of the nominal Fourier coefficients (see problem data above). They are illustrated in the left diagram of Figure 2. Taking the optimal control for the (still feasible) probability level $p = 0.9$ (see Figure 2) and applying it to these 10 scenarios for the initial state yields a time-dependent development of the corresponding energy as illustrated in the right diagram of Figure 2. At terminal time $T = 4$, 9 out of these 10 scenarios reach a terminal energy within the ε band around zero. This is in expected correspondence with the chosen probability 0.9 (of course, slight deviations could occur when repeating the simulation). Note, however, that this is just an out-of-sample test (posterior test) and that the computation of optimal controls has not been based on simulated scenarios but on the parameters of the underlying continuous multivariate distribution.

Figure 3 illustrates that it is important both to increase the stochastic dimension and to refine the discretization grid for the control simultaneously to approximate the solution of the infinite-dimensional problem numerically. In the discretized problems, the controls are approximated by piecewise constant functions. Refinement of the control grid while keeping $N = 10$ constant leads to convergence to the optimal control $u^*(\varepsilon, p, N)$, which shows the typical features of the convergence of Fourier series, namely, the overshooting at the jump discontinuities (Gibbs phenomenon). Figure 1 shows the corresponding solution for $N = 100$ and $M = 256$. With this choice of N, the spillover phenomenon is no longer visible.

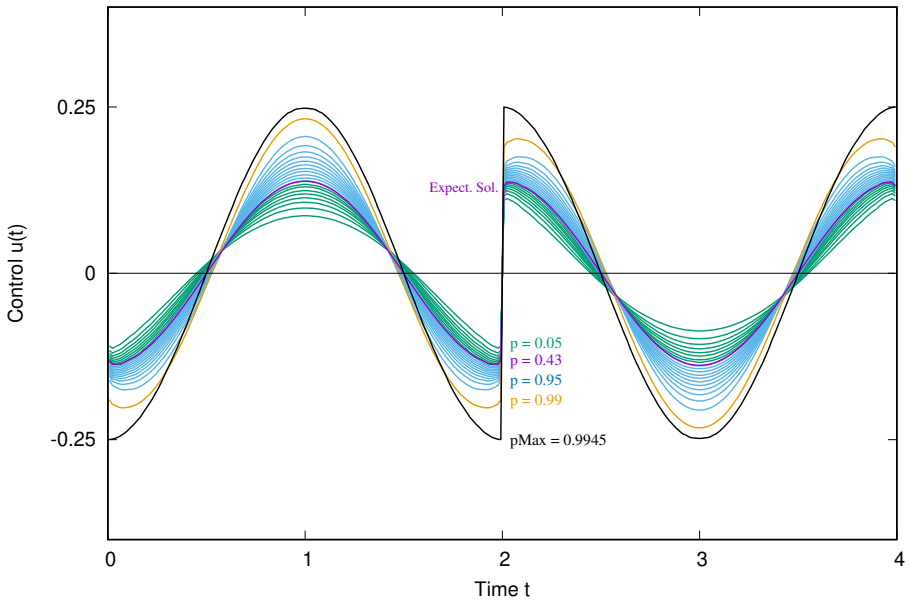

FIG. 4. *The meaning of the figure is analogous to that in the previous example but now for a problem with different initial state.*

7.2. Second example. We repeat the numerical experiment with the more nonlinear nominal initial state $y_0(x) = \pi^{-1}\sin(\pi x)$. The associated optimal controls are illustrated in Figure 4. Except monotonicity of profiles with respect to the probability level, we detect similar effects as in the previous example: Again, the family of profiles passes, when increasing the probability level, through the solution of problem $(N_{\mathsf{prob}})(\varepsilon, p, N, M)$ when replacing the probabilistic constraint $\phi_N(u) \geq p$ by the constraint (7.5) (terminal energy of expected initial state smaller than ε) at $p = 0.43$ and reaches its maximum probability at $p = 0.9945$, where it becomes identical to the deterministic solution of problem (N_{det}). Of course, these two special profiles associated with deterministic problems are nonlinear now due to the nonlinearity of the nominal initial state. As this relation between the probabilistic and the two deterministic solutions is repeatedly observed in examples, we strongly believe, without having a proof yet, that it is generally true.

The left diagram of Figure 5 illustrates the dependence of the maximum achievable probability on the chosen tolerance ε for the terminal energy. Of course, the larger this tolerance, the higher the maximum probability that can be achieved. It turns out that a slight increase of the tolerance from zero to 0.05 already ensures a sufficiently high maximum probability of around 0.9. The right diagram of Figure 5 illustrates the dependence of the cost for the control on the chosen probability level. It can be seen that the level can be quickly increased up to around 0.8–0.9 at a very moderate increase of costs. However, when approaching the maximum possible probability level, the additional costs are considerable.

8. Conclusions. We have studied optimal control problems with systems governed by the wave equation where the initial state is uncertain. In order to take into account the uncertainty, we have prescribed a probabilistic terminal constraint for the energy of the system. In the probabilistic constraint, an upper bound ε for the

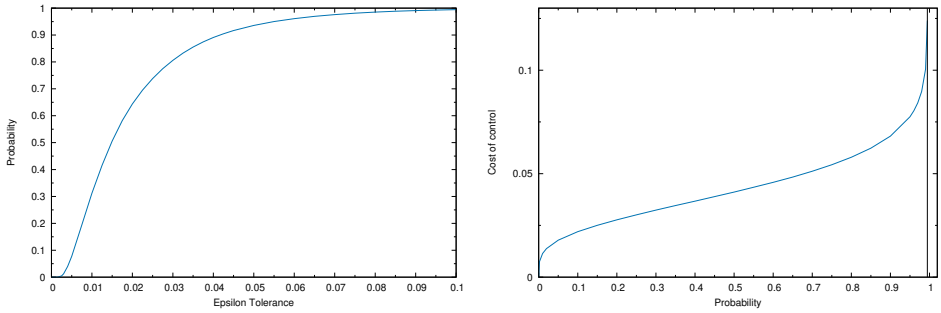

FIG. 5. *Plot of the maximum achievable probability as a function of the energy tolerance ε (left diagram) and of the cost for the control as a function of the chosen probability level (right diagram).*

energy and a desired probability level p appear as parameters. We have shown that for reasonable choices of these parameters, optimal controls exist that solve the optimal control problems with a probabilistic terminal constraint. Examples illustrate that the optimal controls can also be approximated numerically. In this presentation, we did not consider further deterministic control constraints in addition to the probabilistic constraint. However, it would not cause additional problems to include, for example, box constraints for the control.

Since the uncertainty of the initial states occurs for many optimal control problems, it is also interesting to study this type of problem for more complex nonlinear dynamics (see, for example, [5]). This will be a topic of future research.

REFERENCES

[1] W. VAN ACKOOIJ AND R. HENRION, *Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions*, SIAM J. Optim., 24 (2014), pp. 1864–1889.

[2] M. K. BANDA, M. HERTY, AND A. KLAR, *Coupling conditions for gas networks governed by the isothermal Euler equations*, Netw. Heterog. Media 1, (2006), pp. 295–314.

[3] A. BRESSAN, S. CANIC, M. GARAVELLO, M. HERTY, AND B. PICCOLI, *Flows on networks: Recent results and perspectives*, EMS Rev. Math. Sci., 1 (2014), pp. 47–111.

[4] N. BURQ AND N. TZVETKOV, *Random data Cauchy theory for supercritical wave equations* I: *Local theory*, Invent. Math., 173 (2008), pp. 449–475.

[5] J.-B. CAILLAU, M. CERF, A. SASSI, E. TRÉLAT, AND H. ZIDANI, *Solving chance constrained optimal control problems in aerospace via kernel density estimation*, Optim. Control Appl. Methods, 39 (2018), pp. 1833–1858.

[6] R. M. COLOMBO AND H. HOLDEN, *Isentropic fluid dynamics in a curved pipe*, Z. Angew. Math. Phys., 67 (2013), 131.

[7] S. CONTI, M. RUMPF, R. SCHULTZ, AND S. TÖLKES, *Stochastic dominance constraints in elastic shape optimization*, SIAM J. Control Optim., 56 (2018), pp. 3021–3034.

[8] J. CUZICK AND T. L. LAI, *On random Fourier series*, Trans. Amer. Math. Soc., 261 (1980), pp. 53–80.

[9] I. DEÁK, *Subroutines for computing normal probabilities of sets—Computer experiences*, Ann. Oper. Res., 100 (2000), pp. 103–122.

[10] Y. ERMOLIEV AND V. NORKIN, *Stochastic optimization of risk functions via parametric smoothing*, in Dynamic Stochastic Optimization, Lecture Notes in Economics and Mathematical Systems, K. Marti, et al., eds., Springer, Berlin, 2004.

[11] M. H. FARSHBAF-SHAKER, R. HENRION, AND D. HÖMBERG, *Properties of chance constraints in infinite dimensions with an application to PDE constrained optimization*, Set-Valued Var. Anal. 26 (2018), pp. 821–841.

[12] A. GELETU, A. HOFFMANN, M. KLOPPEL, AND P. LI, *An inner-outer approximation approach to chance constrained optimization*, SIAM J. Optim., 27 (2017), pp. 1834–1857.

[13] A. GENZ AND F. BRETZ, *Computation of Multivariate Normal and t Probabilities*, Lect. Notes Stat. 195, Springer, Dordrecht, the Netherlands, 2009.

[14] T. Gonzalez Grandon, H. Heitsch, and R. Henrion, *A joint model of probabilistic/robust constraints for gas transport management in stationary networks*, Comput. Manag. Sci., 14 (2017), pp. 443–460.
[15] M. Gugat, *Norm-minimal Neumann boundary control of the wave equation*, Arab. J. Math., 4 (2015), pp. 41–58.
[16] M. Gugat, *Optimal Boundary Control and Boundary Stabilization of Hyperbolic Systems*, Springer Briefs in Control, Automation and Robotics, Springer, New York, New York, 2015.
[17] M. Gugat and G. Leugering, *Regularization of L^∞-optimal control problems for distributed parameter systems*, Comput. Optim. Appl., 22 (2002), pp. 151–192.
[18] M. Gugat, G. Leugering, and G. Sklyar, *L^p-optimal boundary control for the wave equation*, SIAM J. Control Optim., 44 (2005), pp. 49–74.
[19] A. Hantoute, R. Henrion, and P. Pérez-Aros, *Subdifferential characterization of probability functions under Gaussian distribution*, Math. Program., 174 (2019), pp. 167–194.
[20] M. Heinkenschloss, B. Kramer, T. Takhtaganov, and K. Willcox, *Conditional-value-at-risk estimation via reduced-order models*, SIAM/ASA J. Uncertain. Quantif., 6 (2018), pp. 1395–1423.
[21] M. Hill, *Convergence of Random Fourier Series*, REU participant papers, University of Chicago, 2012.
[22] J.-P. Kahane, *Some Random Series of Functions*, Cambridge University Press, Cambridge, 1985.
[23] D. P. Kouri and T. M. Surowiec, *Existence and optimality conditions for risk-averse PDE-constrained optimization*, SIAM/ASA J. Uncertain. Quantif., 6 (2018), pp. 787–815.
[24] I. Lasiecka and J. Sokolowski, *Sensitivity analysis of optimal control problems for wave equations*, SIAM J. Control Optim., 29 (1991), pp. 1128–1149.
[25] R. Lepp, *Stochastic optimization type algorithm for maximization of probability function*, Communications of Academy of Sciences in Estonia, Physics and Math., 32 (1983), pp. 150–156.
[26] J. L. Lions, *Exact controllability, stabilization and perturbations of distributed systems*, SIAM Rev., 30 (1988), pp. 1–68.
[27] J. Martínez-Frutos and F. Periago Esparza, *Optimal Control of PDEs under Uncertainty*, Springer, Cham, Switzerland, 2018.
[28] A. Prékopa, *Stochastic Programming*, Kluwer, Dordrecht, the Netherlands, 1995.
[29] P. Schmidt, A. Geletu, and P. Li, *Stochastische Optimierung parabolischer PDE-Systeme unter Wahrscheinlichkeitsrestriktionen am Beispiel der Temperaturregelung eines Stabes*, Automatisierungstechnik, 66 (2018), pp. 975–985.
[30] F. Shan, L. Zhang, and X. Xiao, *A smoothing function approach to joint chance-constrained programs*, J. Optim. Theory Appl., 163 (2014), pp. 181–199.
[31] E. Zuazua, *Propagation, observation, and control of waves approximated by finite difference methods*, SIAM Rev., 47 (2005), pp. 197–243.