

# Solving Stochastic Optimal Control Problems by a Wiener Chaos Approach

Tony Huschto · Sebastian Sager

Received: date / Accepted: date

**Abstract** We introduce a novel generic methodology to solve continuous finite-horizon stochastic optimal control problems (SOCPs). We treat controlled stochastic differential equations (SDEs) within the Wiener chaos framework by utilizing Malliavin calculus and developing innovative ideas to preserve the feedback character of optimal Markov decision rules.

This allows a direct reformulation of SOCPs into deterministic ones. Hence, it facilitates using Bock's direct multiple shooting method for solving SOCPs and pioneers the extension of sophisticated methods for deterministic control to the broad context of SDEs.

Numerical examples validate this new framework with huge computational advantages compared to standard ideas in SOC.

**Keywords** Stochastic Optimal Control · Markov Control · Wiener Chaos · Malliavin Calculus · Direct Multiple Shooting

**Mathematics Subject Classification (2000)** MSC 93E20 · MSC 60H35 · MSC 49J15

## 1 Introduction

Solving *continuous finite-horizon stochastic optimal control problems* has attracted increasing interest in the past decades. In economics and finance, and especially in portfolio management, starting with classical questions of, e. g., Merton [9], such problems driven by stochastic differential equations are usually used as the underlying modeling framework. Here the detection of optimal *decision rules* becomes a crucial task.

---

T. Huschto  
Interdisciplinary Center for Scientific Computing, Ruprecht-Karls University of Heidelberg  
Tel.: +49-6221-548809  
Fax: +49-6221-545444  
E-mail: tony.huschto@iwr.uni-heidelberg.de

S. Sager  
Institute for Mathematical Optimization, Otto-von-Guericke University Magdeburg

Basically, problems belonging to that class can be solved if we consider the corresponding *Hamilton-Jacobi-Bellman* (HJB) equation [17, 18, 33]. In most cases, however, this second-order partial differential equation is not analytically solvable or does not even admit a global solution at all, which is why appropriate numerical ideas have to be applied to obtain optimal decision rules.

Apart from discretizing the HJB equation, the common methods of choice are *direct* numerical ideas. Often they are based on the application of *Markov chains*, i. e., discretizing the continuous process in space and time. For example, the approach of Kushner [20, 21] uses finite-difference and finite-element ideas to obtain the transition probabilities of the Markov chain from the HJB equation. In contrast the work of Krawczyk [3, 19] builds on *weak approximation* schemes [17] for stochastic differential equations (SDE) to design a *Markov decision chain* from the original problem. The resulting Bellman equation can then be solved using value iteration. A third related idea introduced by Pagès [35] is based on the *quantization* of stochastic processes [27, 28, 36]. There the original process is projected onto a random vector taking values on a finite grid following a closest neighbor rule [35]. The resulting problem can, again, be solved by a dynamic programming procedure.

In this paper we propose an entirely different methodology to solve continuous finite-horizon stochastic optimal control problems. We utilize the *polynomial chaos* (PC) or *Wiener chaos* framework developed by Norbert Wiener [42] and generalized by Cameron and Martin later on [8] to reformulate the original stochastic optimal control problem as a (larger) deterministic optimal control problem.

By means of the chaos expansion we express the considered stochastic processes in a Fourier-like fashion in terms of deterministic coefficient functions and orthonormal basis polynomials spanning the underlying *Wiener chaos space*. Following this construction the driving force of the stochastic processes, i. e., the *Brownian motion*, can be represented in terms of Gaussian random variables. This is closely related to its *Karhunen-Loève expansion* [14, 24].

However, in order to reformulate the appearing SDE as a system of ordinary differential equations (ODEs), the stochastic integral characterizing the diffusion has to be treated cautiously, because one has to integrate with respect to a function that is nowhere differentiable. Thus it cannot be approximated straightforwardly. In the context of this paper we utilize *Malliavin calculus* [30, 31] to overcome this difficulty. The resulting *propagator*, appearing in a related fashion in [25, 26] for solving a class of partial differential equations with random forcing, is enhanced afterwards to controlled SDEs. It implicitly includes all randomness of the original SDE, which is due to the underlying Hilbert space theory. But in doing so it is crucial to implement feedback formulations of the control process to preserve the non-anticipativity of optimal decision rules.

The emphasis of this paper is on one side on the idea of transforming a stochastic optimal control problem into a deterministic one, which can be solved by sophisticated methods of optimal control afterwards, e. g., Bock's *direct multiple shooting* approach [6], which is based on the general *multiple shooting* ideas of Osborne and Bulirsch [34, 7]. Additionally, the proposed transformation opens the possibility to apply state-of-the-art methods for deterministic optimal control in the broad context of random processes. On the other side, we illustrate the methodology introduced here with numerical examples that yield very encouraging results, establishing the potential for huge computational savings compared to standard approaches in stochastic optimal control.

The paper is organized as follows. In Section 2 we introduce the general chaos expansion and give an overview on the important parts of Malliavin calculus we need for our approach. In Section 3 we apply these ideas to stochastic differential equations and in Section 4 to stochastic optimal control problems, resulting in a transformation to deterministic ones. In that context we establish a method to preserve the feedback character of the control policy. In Section 5 we shortly review the direct multiple shooting method for convenience. In Section 6 two numerical examples illustrate our novel approach to solve continuous finite-horizon stochastic optimal control problems.

## 2 Notations and Mathematical Preliminaries

### 2.1 Statement of the Problem

Let  $\{X_t\}_{t \in [0, T]}$  be a  $n_X$ -dimensional stochastic process within the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and  $\{B_t\}_{t \in [0, T]}$  a  $n_B$ -dimensional Brownian motion. Then in the further course of this work we consider the finite-horizon stochastic optimal control problem

$$\min_{u \in \mathcal{A}} \mathbb{E} \left[ \int_0^T L(t, X_t, u_t) dt + G(T, X_T) \right] \quad (1a)$$

$$\text{s.t.} \quad dX_t = b(t, X_t, u_t) dt + \sigma(t, X_t, u_t) dB_t, \quad (1b)$$

$$X_0 = x_0, \quad (1c)$$

with cost functional (1a) and  $b$  and  $\sigma$  describing the drift and diffusion parts of the random state process where  $X_t \in \mathcal{X} \subset \mathbb{R}^{n_X}$  for all  $t \in [0, T]$ . For notational ease we often write only  $X = \{X_t\}$ , whenever the time interval  $[0, T]$  is obvious. Assuming its existence in  $L^2(\Omega \times [0, T])$  for the moment, this process  $\{X_t\}$  is driven by the controlled Itô stochastic differential equation (1b) with initial condition (1c). We assume to stop the process at the final time  $T$  rather than letting it evolve until it leaves a predefined region  $\mathcal{G}$ . The control  $u_t \in \mathcal{U} \subset \mathbb{R}^{n_u}$  that is chosen over a set  $\mathcal{A}$  of *admissible* controls to minimize the cost functional (1a) then has to be a stochastic process  $\{u_t\} = \{u(t, \omega)\}$  as well. Moreover, it has to be at least  $\mathcal{F}_t$ -adapted since the decisions we take at time instant  $t \in [0, T]$  can only be depending on what already happened up to  $t$ . The most common choices of admissible control functions [33] are

- *deterministic* controls  $u(t, \omega) = u(t)$ ,
- *open-loop controls*  $u(t, \omega)$  which are non-anticipative with respect to the Brownian motion  $\{B_t\}$ , and
- *Markov controls*  $u(t, \omega) = u_M(t, X_t(\omega))$  with a non-random and Lebesgue-measurable function  $u_M$ . With such a control the process  $\{X_t\}$  becomes an Itô diffusion. Furthermore, we assume  $u_M$  to be sufficiently smooth.

In the following we restrict ourselves to Markov controls, writing simply  $u_t = u(t, X_t)$ .

The goal of this work is to transform the stochastic optimal control problem (1) into a deterministic optimal control problem that can be solved by existing sophisticated methods like, e.g., the direct multiple shooting approach [6]. One

important aspect is that we have to preserve the feedback character of the Markov control.

## 2.2 The Wiener Chaos Expansion

The basic concept of the presented idea is the *Wiener chaos expansion* going back to the work of Norbert Wiener [42] who introduced the *homogeneous chaos*, an orthogonal development for nonlinear functionals based on a Gaussian measure. Cameron and Martin [8] generalized this first idea as they used Fourier-Hermite functionals to construct an orthogonal basis for nonlinear functions. The first attempts of applying the idea to problems with random phenomena have again been done by Wiener [43]. Based on [8], Wiener's expansion of functionals can be generalized to hold for arbitrary random processes [32, 40]. Especially in the past few years, the concept of polynomial chaos has attracted much attention again, e. g., in technical applications like calculations of von Mises stress [1] or robustness in shape optimization [39]. Further on, there is a deep connection between the polynomial chaos and the general Hilbert space theory, [11] and [13] give a detailed explanation and discussion on that topic.

Based on [42], we have the following theorem, cf. [8, Theorem 1] as well as [31, Theorem 1.1.1]:

**Theorem 1 (Cameron and Martin)** *Assume that the process  $\{X_t\}$  satisfies the integrability condition  $\mathbb{E} \left[ \int_0^T |X_t|^2 dt \right] < \infty$ , i. e.,  $\{X_t\} \in L^2(\Omega \times [0, T])$ . Then  $X_t$  can be expanded in  $[0, T]$  as*

$$X_t = \sum_{\alpha} x_{\alpha}(t) \Psi^{\alpha}(\xi) \quad (2)$$

with (deterministic) coefficient functions  $x_{\alpha}(t)$  and  $\{\Psi^{\alpha}(\xi)\}_{\alpha}$  being an orthonormal basis of the Wiener chaos space  $L^2(\Omega \times [0, T])$ .

Within this theorem,  $\alpha$  denotes a multi-index from the set

$$\mathcal{I} = \left\{ \alpha = (\alpha_i)_{i \geq 1} \mid \alpha_i \geq 0, |\alpha| = \sum_{i=1}^{\infty} \alpha_i < \infty \right\}. \quad (3)$$

For convenience, in the rest of this paper we shortly write

$$\begin{aligned} \alpha = \mathbf{0} & \quad \text{if } \alpha_i = 0 \text{ for all } i, \\ \alpha = \mathbf{e}_j & \quad \text{if } \alpha_i = \delta_{ij}. \end{aligned}$$

To construct the basis polynomials  $\Psi^{\alpha}(\xi)$  we first need [26, Chapter 2.3]

**Lemma 1** *Let  $\{m_i(t)\}$  be an orthonormal basis of the Hilbert space  $L^2([0, T])$ . Then the Itô integrals*

$$\xi_i = \int_0^T m_i(t) dB_t \quad (4)$$

define independent standard Gaussian random variables.

**Definition 1** With the one-dimensional Hermite polynomials

$$H_n(x) = \frac{(-1)^n}{n!} e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2} \quad (5)$$

the basis polynomials of the Wiener chaos space are defined by

$$\Psi^\alpha(\xi) = \sqrt{\alpha!} \prod_i H_{\alpha_i}(\xi_i), \quad (6)$$

where  $\alpha!$  as the product of all component's factorials.

From that definition and the properties of the Hermite polynomials one directly shows the orthonormality of the basis polynomials, which are often referred to as *Wick polynomials*.

Hence, the coefficient functions  $x_\alpha(t)$  can be interpreted as projections of the process  $X_t$  onto the corresponding chaos basis as in a Fourier-related manner. It holds

$$x_\alpha(t) = \mathbb{E}[X_t \Psi^\alpha].$$

Particularly the coefficient function of order zero obtains a special meaning. It coincides with the expectation of the process  $X_t$ , as the basis polynomial of order zero is identically one.

$$x_{\mathbf{0}}(t) = \mathbb{E}[X_t \Psi^{\mathbf{0}}] = \mathbb{E}[X_t]. \quad (7)$$

In a similar fashion the variance of  $X_t$  and, consequently, all higher moments can be expressed in terms of the coefficient functions  $x_\alpha(t)$  only, compare [1].

Lemma 1 additionally reveals a connection to the *Karhunen-Loève expansion* (KLE) of the Brownian motion  $\{B_t\}$ . Due to the scaling property of this process we can restrict the considerations to the time interval  $[0, 1]$ . Then the KLE is determined from the eigenvalues and eigenvectors of the covariance between two time points  $s, t$  [14, 24]. In the special case of the Brownian motion with covariance function  $C(s, t) = \min\{s, t\}$  one calculates the eigenvalues  $\lambda_i$  and their corresponding eigenfunctions  $v_i(t)$

$$\lambda_i = \frac{1}{\left(i - \frac{1}{2}\right)^2 \pi^2}, \quad v_i(t) = \sqrt{2} \sin\left(\left(i - \frac{1}{2}\right) \pi t\right), \quad i = 1, \dots, \infty,$$

obtaining the KLE

$$B(t, \xi) = \sum_{i=1}^{\infty} \frac{\sqrt{2} \sin\left(\left(i - \frac{1}{2}\right) \pi t\right)}{\left(i - \frac{1}{2}\right) \pi} \xi_i. \quad (8)$$

For larger time horizons  $[0, T]$  one has to modify the time variable as it is done, e. g., in [4] or [26]. On the other hand, keeping (4) in mind, one can rewrite  $B_t$  by its Fourier expansion [26]

$$B(t) = \sum_{i=1}^{\infty} \xi_i \int_0^t m_i(s) ds. \quad (9)$$

This expansion converges in the mean square sense. Letting the orthonormal basis  $\{m_i(t)\}$  of  $L^2([0, 1])$  be given by

$$m_i(t) = \sqrt{2} \cos\left(\left(i - \frac{1}{2}\right) \pi t\right),$$

one obtains (8) again.

### 2.3 Malliavin Calculus

The second major concept that we need for the transformation of the stochastic problem (1) into a deterministic one is Malliavin calculus. Extensive introductions to the topic can be found in, e.g., [30,31]. Here we only want to summarize the essential ingredients needed for the presented method to solve problems of type (1). Again, we use the Hermite polynomials (5) and the chaos basis functions  $\Psi^\alpha$  (6).

Let  $W = \{W(h) \mid h \in H\}$  denote an isonormal Gaussian process defined in  $(\Omega, \mathcal{F}, \mathbb{P})$  and associated with the Hilbert space  $H$ . (Compare Lemma 1, i.e., the definition of the variables  $\xi_i$  depending on  $m_i \in L^2$ ,  $\xi_i = W(m_i)$ .) Further on, assume  $F$  to be a square-integrable, smooth random variable of the form

$$F = f(W(h_1), \dots, W(h_n)) \quad (10)$$

with  $f \in C_p^\infty(\mathbb{R}^n)$  and  $h_i \in H$ ,  $i = 1, \dots, n$ .

**Definition 2 (Malliavin derivative  $D$ , [31])** The *Malliavin derivative* of a smooth random variable of the form (10) with respect to the chance parameter  $\omega \in \Omega$  is the  $H$ -valued random variable

$$DF = \sum_{i=1}^n \partial_i f(W(h_1), \dots, W(h_n)) \cdot h_i. \quad (11)$$

By that definition, we interpret  $DF$  as a directional derivative. We denote the domain of  $D$  in  $L^2(\Omega)$  by  $\mathbb{D}^{1,2}$ . This space is again a Hilbert space with the scalar product

$$\langle F, G \rangle_{1,2} = \mathbb{E}[FG] + \mathbb{E}[\langle DF, DG \rangle_H].$$

Additionally, the derivative of a random variable  $F \in \mathbb{D}^{1,2}$  is a stochastic process denoted by  $\{D_t F\}_{t \in [0, T]}$ .

To give an illustrating example, let us calculate the Malliavin derivative of the basis polynomial  $\Psi^\alpha(\xi)$ , which is indeed a random variable in  $\mathbb{D}^{1,2}$ .

*Example 1* Consider the basis polynomial  $\Psi^\alpha(\xi)$  defined by (6) for fixed  $\alpha \in \mathcal{I}$ . We deduce (exploiting the rules for differentiating Hermite polynomials of the form (5) and the definition of the random variable  $\xi_i$ )

$$\begin{aligned} D_s \Psi^\alpha(\xi) &= D_s \left( \sqrt{\alpha!} \prod_i H_{\alpha_i}(W(m_i)) \right) \\ &= \sum_{j=1}^{\infty} \sqrt{\alpha!} \prod_{\substack{i \\ i \neq j}} H_{\alpha_i}(\xi_i) H_{\alpha_j-1}(\xi_j) m_j(s) \\ &= \sum_{j=1}^{\infty} \sqrt{\alpha_j} m_j(s) \Psi^{\alpha^-(j)}(\xi) \end{aligned} \quad (12)$$

with the diminished multi-index  $\alpha^-(j)$  defined as

$$\alpha_i^-(j) = \begin{cases} \alpha_i, & i \neq j \\ \alpha_i - 1, & i = j \end{cases}. \quad (13)$$

The next component we need is the *divergence* operator.

**Definition 3 (Divergence  $\delta$ , [31])** We denote by  $\delta$  the *adjoint* of the operator  $D$ . That is,  $\delta$  is an unbounded operator on  $L^2(\Omega; H)$  with values in  $L^2(\Omega)$  such that the domain  $\text{Dom } \delta$  of the divergence is the set of  $H$ -valued square integrable random variables  $u \in L^2(\Omega; H)$  that are bounded according to

$$|\mathbb{E}[\langle DF, u \rangle_H]| \leq c \|F\|_2$$

for all  $F \in \mathbb{D}^{1,2}$  and a constant  $c$  depending on  $u$ .

If  $u \in \text{Dom } \delta$  then  $\delta(u) \in L^2(\Omega)$  is characterized for any  $F \in \mathbb{D}^{1,2}$  by

$$\mathbb{E}[F\delta(u)] = \mathbb{E}[\langle DF, u \rangle_H]. \quad (14)$$

Equation (14) provides the important *integration by parts formula* we need later on to transform the SDE determining the process  $\{X_t\}$  into a system of ODEs. In the context of stochastic integrals we have:

**Proposition 1 ([31])** Let  $\{B_t\}_{t \in [0, T]}$  be a  $d$ -dimensional Brownian motion and consider by  $L_a^2$  the closed subspace of adapted processes in  $L^2([0, T]^d \times \Omega)$ . Then  $L_a^2 \subset \text{Dom } \delta$  and the divergence operator restricted to  $L_a^2$  coincides with the Itô stochastic integral, i. e.,

$$\delta(u) = \sum_{i=1}^d \int_0^T u_t^i dB_t^i.$$

Following from that proposition and (14) we obtain a customized integration by parts formula restricted to the time interval  $[0, t]$ .

**Lemma 2** Let  $X_t$  be a square integrable and  $\mathcal{F}_t$ -measurable random variable for all  $t \in [0, T]$ . Then for all  $F \in \mathbb{D}^{1,2}$  we obtain the formula

$$\mathbb{E}\left[F \cdot \int_0^t X_s dB_s\right] = \mathbb{E}\left[\int_0^t D_s F \cdot X_s ds\right]. \quad (15)$$

*Proof* We directly calculate

$$\begin{aligned} \mathbb{E}\left[F \cdot \int_0^t X_s dB_s\right] &= \mathbb{E}\left[F \int_0^T X_s \mathbb{1}_{\{s \leq t\}} dB_s\right] = \mathbb{E}\left[F\delta(\hat{X})\right] = \mathbb{E}\left[\langle DF, \hat{X} \rangle_H\right] \\ &= \mathbb{E}\left[\int_0^T D_s F \cdot X_s \mathbb{1}_{\{s \leq t\}} ds\right] = \mathbb{E}\left[\int_0^t D_s F \cdot X_s ds\right], \end{aligned}$$

where  $\hat{X}$  simply denotes the restriction of the process  $X$  to the interval  $[0, t]$ .

### 3 Stochastic Differential Equations and the Wiener Chaos

As only a small number of stochastic differential equations can be solved analytically, a very important issue is the numerical approximation of their solution. [17] describes many different algorithms for that task, using adaptations of deterministic numerical integration schemes to the stochastic necessities, e. g., stochastic Taylor approximations. However, all of these algorithms make use of random numbers. In this section we show how the aforementioned ideas of the Wiener chaos can be used to transform a given Itô SDE into a system of ODEs which, thereafter, can be solved by sophisticated methods for deterministic differential equations.

Let us start with an uncontrolled (one-dimensional) process  $\{X_t\}$  defined by the autonomous SDE on the time interval  $[0, T]$ , i. e.,

$$dX_t = f(X_t) dt + \phi(X_t) dB_t, \quad X_0 = x_0, \quad (16a)$$

or, conveniently, written in its integral form

$$X_t = x_0 + \int_0^t f(X_s) ds + \int_0^t \phi(X_s) dB_s. \quad (16b)$$

The generalization to multi-dimensional processes  $\{X_t\}$  and  $\{B_t\}$  follows straightforwardly. Here we focus on the basic idea. We can apply the Wiener chaos expansion (2) from Theorem 1 if the SDE (16) has a square integrable solution in  $[0, T]$ .

**Theorem 2** *Let  $\{X_t\}$  be given by (16) and assume that  $\{X_t\} \in L^2([0, T] \times \Omega)$ . Then  $X_t$  can be written in its Wiener chaos expansion*

$$X_t = \sum_{\alpha \in \mathcal{I}} x_\alpha(t) \Psi^\alpha(\xi)$$

with the coefficients  $x_\alpha(t)$  determined by the following propagator on  $[0, T]$

$$\dot{x}_\alpha(t) = f(X_t)_\alpha + \sum_{j=1}^{\infty} \sqrt{\alpha_j^-} m_j(t) \phi(X_t)_{\alpha^-(j)}, \quad (17a)$$

$$x_\alpha(0) = \mathbb{1}_{\{\alpha=\mathbf{0}\}} \cdot x_0. \quad (17b)$$

Within this system of ODEs,  $f_\alpha$  and  $\phi_\alpha$  denote again the  $\alpha$ -coefficients of the chaos expansions of the functions  $f$  and  $\phi$  (depending on  $X_t$ ) and  $\alpha^-(j)$  the diminished multi-index as defined in (13).

*Proof* Inserting the expansion (2) into (16b), multiplying with the basis polynomial  $\Psi^\beta(\xi)$ ,  $\beta \in \mathcal{I}$ , and calculating expectations yields for all  $\beta \in \mathcal{I}$  and  $t \in [0, T]$

$$x_\beta(t) = x_0 \cdot \mathbb{1}_{\{\beta=\mathbf{0}\}} + \int_0^t \mathbb{E} \left[ f(X_s) \Psi^\beta(\xi) \right] ds + \mathbb{E} \left[ \Psi^\beta(\xi) \cdot \int_0^t \phi(X_s) dB_s \right].$$

While the first appearing integral is deterministic and the expectation within can be represented by the corresponding coefficient function of the expansion of  $f(X_t)$ ,



the second integral has to be treated with the integration by parts formula (15). This yields

$$\begin{aligned}
x_\beta(t) &= x_0 \cdot \mathbb{1}_{\{\beta=\mathbf{0}\}} + \int_0^t f(X_s)_\beta \, ds + \mathbb{E} \left[ \int_0^t D_s \Psi^\beta(\xi) \phi(X_s) \, ds \right] \\
&= x_0 \cdot \mathbb{1}_{\{\beta=\mathbf{0}\}} + \int_0^t f(X_s)_\beta \, ds \\
&\quad + \sum_{j=1}^{\infty} \int_0^t \sqrt{\beta_j} m_j(s) \mathbb{E} \left[ \phi(X_s) \Psi^{\beta^-(j)}(\xi) \right] \, ds \\
&= x_0 \cdot \mathbb{1}_{\{\beta=\mathbf{0}\}} + \int_0^t f(X_s)_\beta \, ds + \sum_{j=1}^{\infty} \int_0^t \sqrt{\beta_j} m_j(s) \phi(X_s)_{\beta^-(j)} \, ds. \quad (18)
\end{aligned}$$

As for all  $\beta \in \mathcal{I}$  there are only a finite number of non-zero components  $\beta_i$ ,  $i \geq 1$ , compare (3), the formally infinite sum on the right-hand side of (18) is in fact finite. Hence, the assertion follows after differentiating with respect to  $t$ .

*Remark 1* Keeping in mind the definition of the random variables  $\xi_i$  (4) and their significance in constructing the Brownian motion (9), we see that all information about the behavior of the stochastic process  $\{X_t\}$  is implicitly captured within the *deterministic* ODE system (17).

In the context of a special class of stochastic partial differential equations driven by Gaussian white noise, a similar propagator is derived in a somewhat related fashion in [25], the references within, and, based on that, [26].

One major advantage of this approach of solving SDEs is that the expectation of the process,  $\mathbb{E}[X_t]$ , is directly given by the zero-order coefficient  $x_{\{\alpha=\mathbf{0}\}}(t)$ . Hence, it need not be calculated by, e.g., *Monte Carlo* methods, where a huge amount of sample paths (computed by some standard stochastic integration method) is necessary. Here, if one is interested in special sample paths, they can be determined by using different realizations of the random vector  $\xi$  *after* the system has been solved. Similar to the expectation, the variance of the process and all higher moments are completely specified by the deterministic coefficient functions of the chaos expansion.

## 4 Stochastic Optimal Control

### 4.1 The Propagator of the Control Problem

Let us return to the stochastic optimal control problem (1). In particular, the cost functional (given by (1a)) or value function

$$\Phi(t, x) = \min_{u \in \mathcal{A}} \mathbb{E} \left[ \int_t^T L(s, X_s, u_s) \, ds + G(T, X_T) \mid X_t = x \right] \quad (19)$$

satisfies for each initial value  $(t, x)$  the Hamilton-Jacobi-Bellman partial differential equation (PDE) [18, 33]

$$0 = \inf_u \left\{ L(t, x, u) + \frac{\partial \Phi}{\partial t}(t, x) + \sum_{i=1}^{n_x} b^i(t, x, u) \frac{\partial \Phi}{\partial x_i}(t, x) + \frac{1}{2} \sum_{i,j=1}^{n_x} (\sigma \sigma^\top)^{ij}(t, x, u) \frac{\partial^2 \Phi}{\partial x_i \partial x_j}(t, x) \right\} \quad (20)$$

with the final condition

$$\Phi(T, x) = G(T, x) \quad (21)$$

for each (initial) value  $(t, x)$ .

The approaches of solving such kind of problems include tackling this PDE directly [9, 18]. However, one cannot guarantee the assumptions in all cases, as the determination of appropriate boundaries is intricate in most applications, [33]. Therefore, we prefer the use of direct methods. Successful approaches based on Markov chains can be found in, e. g., [20, 21] or [3, 19].

As an alternative, our novel methodology applies the ideas of the previous section to the control problem (1). Therefore, we have to consider a controlled SDE (compare especially Equation (1b)), where we proceed in a similar way as above to obtain the propagator of the system. Besides the expansion of the state process  $\{X_t\}$  we have to include a second chaos expansion determining the control process  $\{u_t\}$ , i. e.,

$$u_t = \sum_{\alpha \in \mathcal{I}} u_\alpha(t) \Psi^\alpha(\xi). \quad (22)$$

*Remark 2* By incorporating expansion (22) directly in the propagator obtained for a controlled SDE, we cannot guarantee the assumed feedback character of the Markov control  $u_t = u(t, X_t)$  anymore.

From a computational point of view, there is another disadvantage of directly implementing the expansion (22) of the control process: The final *deterministic* optimal control problem we want to deduce would contain the same number of state and control functions  $u_\alpha(t)$  and  $x_\alpha(t)$ . Hence the resulting problem would be very hard to solve numerically.

The remedy to both problems lies in

**Theorem 3** *Assume that the Markov control  $u_t = u(t, X_t)$  can be Taylor-expanded in terms of  $X_t$ . Then by considering the  $q$ -th order polynomial*

$$u^q(t, X_t) = \sum_{i=0}^q \hat{u}_i(t) X_t^i, \quad (23)$$

*the original control coefficients  $u_\alpha(t)$  of (22) are characterized completely by the  $q+1$  new control functions  $\hat{u}_i(t)$ ,  $i = 0, \dots, q$ , and the state coefficients  $x_\alpha(t)$ . Furthermore, the resulting control  $u_t^q$  is automatically non-anticipative and tends to  $u_t$  for  $q \rightarrow \infty$ .*

*Proof* In contrast to expanding  $u(t, X_t)$  in  $t$ , for calculating the expansion in terms of  $X_t$  we do not need a stochastic Taylor expansion. Thus, an (infinite) Taylor expansion in  $X_t = a$  yields

$$u_t = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{\partial}{\partial x} u^{(n)}(t, X_t) \Big|_{X_t=a} (X_t - a)^n$$

which can always be rewritten in powers of  $X_t$ . Hence, with defining new control functions  $\hat{u}_i(t)$  as the coefficient terms of these powers, one arrives at the infinite version of (23). Similarly, a finite version up to order  $q$  can be defined with the  $q$ -th term corresponding to the remaining error. The convergence to  $u_t$  follows directly and so does the non-anticipativity as we express  $u_t$  through the state process which fulfills the property by definition.

Now if we compare (22) and (23)

$$\sum_{\alpha \in \mathcal{I}} u_{\alpha}(t) \Psi^{\alpha}(\xi) = \sum_{i=0}^q \hat{u}_i(t) X_t^i \quad (24)$$

by inserting the chaos expansion (2) of  $X_t$  and projecting the resulting expression onto the chaos bases, we obtain a system describing the original control coefficients  $u_{\alpha}(t)$  by the new control functions  $\hat{u}_i(t)$  and the state coefficients  $x_{\alpha}(t)$ , while having the feedback character of the Markov control included implicitly.

*Example 2* Assume  $q = 2$ . Then the quadratic and non-anticipative expansion of the control process  $\{u_t\}$  is given by (22), where the coefficients  $u_{\alpha}(t)$  are defined by the system

$$\begin{aligned} u_{\alpha}(t) &= \hat{u}_0(t) \cdot \mathbb{1}_{\{\alpha=0\}} + \hat{u}_1(t) \cdot x_{\alpha}(t) \\ &\quad + \hat{u}_2(t) \cdot \sum_{\beta \in \mathcal{I}} \sum_{0 \leq \gamma \leq \alpha} C(\alpha, \gamma, \beta) x_{\alpha-\gamma+\beta}(t) x_{\gamma+\beta}(t) \end{aligned} \quad (25)$$

for all  $\alpha \in \mathcal{I}$  and  $C(\alpha, \gamma, \beta)$  given by (compare [26])

$$C(\alpha, \gamma, \beta) = \sqrt{\binom{\alpha}{\gamma} \binom{\gamma+\beta}{\beta} \binom{\alpha-\gamma+\beta}{\beta}}. \quad (26)$$

All multi-index operations are defined component-wise, including the binomial coefficient that is calculated as the product of the component's binomial coefficients. Note as well that by  $0 \leq \gamma \leq \alpha$  it holds  $\alpha + \gamma - \beta \in \mathcal{I}$  and  $\gamma + \beta \in \mathcal{I}$ .

Combining Theorems 2 and 3 we obtain a deterministic reformulation of the controlled SDE (1b). Hence, the only missing part of our transformation method is the objective function (1a) of the original stochastic optimal control problem. But as this is already formulated as an expectation value, it can be rewritten directly in terms of the deterministic coefficients  $x_{\alpha}(t)$  of the state process  $\{X_t\}$  and the (new) control functions  $\hat{u}_i(t)$ . We give a detailed example in Section 6.

*Remark 3* To reflect that optimal controls can be discontinuous, our preference for solving the resulting deterministic control problem after applying expansion (23) is Bock's *direct multiple shooting approach*, cf. Section 5. Within this method controls are identified on a discrete multiple shooting grid, see (30), allowing discontinuous control profiles for each  $\hat{u}_i(t)$ ,  $i = 0, \dots, q$ .

## 4.2 Truncating the Propagator

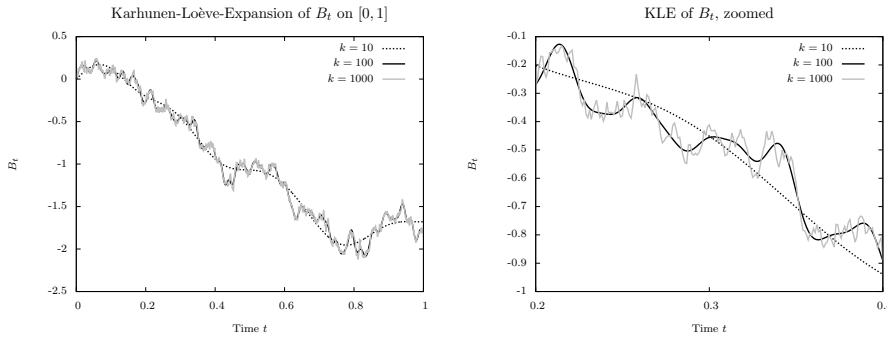
For numerical applications the propagator certainly has to be truncated. Basically, there are three major types of truncation:

- The order  $q$  of the corresponding expansion from the approximation (23) of the Markov control.
- The number  $k$  of random variables  $\xi_i$ ,  $i = 1, \dots, k$  that are used within the construction of the basis polynomials  $\Psi^\alpha(\xi)$ ,  $\xi = (\xi_i)_{i=1, \dots, k}$ . Note that due to their construction (4) via the basis polynomials  $m_i(t)$  of  $L^2([0, T])$ , the random variables  $\xi_i$  give less information for increasing index  $i$ .
- The maximum order  $p$  of the basis polynomials. Note that the coefficient functions with corresponding basis polynomial  $\Psi^\alpha$  with  $|\alpha| = p$  give less information for increasing  $p$ .

From the second and third type we obtain the (simply) truncated multi-index set

$$\mathcal{I}_{k,p} = \left\{ \alpha = (\alpha_1, \dots, \alpha_k) \mid \alpha_i \geq 0, \sum_{i=1}^k \alpha_i \leq p \right\}. \quad (27)$$

*Example 3* If we consider again the Brownian motion process  $\{B_t\}$  by its expansion (8) or (9), the importance of the truncation number  $k$  is shown in Figure 1.



**Fig. 1** Example paths for the truncated Karhunen-Loève expansion of a Brownian Motion  $B_t$ . The left figure shows sample paths for varying  $k$  on the time horizon  $t \in [0, 1]$ , the right one a zoomed part of that plot.

## 4.3 Sparse Truncation

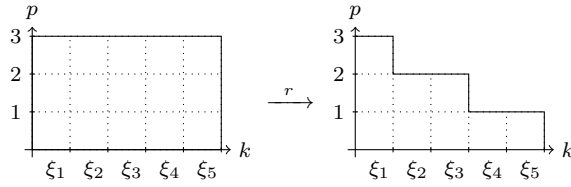
As already stated in the previous section, the importance of the coefficients  $x_\alpha(t)$  decays depending on the order  $p$  of the basis polynomials  $\Psi^\alpha$  and the decaying rate of the Gaussian expansion, i. e., the index of the random variables  $\xi_i$ ,  $i \in \{1, \dots, k\}$ , used for the construction of  $\Psi^\alpha(\xi)$ . Especially if we consider coefficients with index  $\bar{\alpha}$ , where  $|\bar{\alpha}|$  is large and  $\bar{\alpha}$  consists of a combination of random variables  $\xi_j$  with large indices  $j$ , the information gained is very low.

Hence, we define a *sparse index* for truncating the index set  $\mathcal{I}$  (compare [10, 26]).

**Definition 4** Let  $p$  be the maximum order of the index  $\alpha$ . Then the sparse index  $r = (r_1, \dots, r_k)$  fulfills  $p = r_1 \geq r_2 \geq \dots \geq r_k$  and we define the sparse index set

$$\mathcal{I}_{k,p}^r = \{(\alpha_1, \dots, \alpha_k) \mid |\alpha| \leq p, \alpha_i \leq r_i \forall i\}. \quad (28)$$

*Example 4* Let  $k = 5$  and  $p = 3$ . Then a possible choice of  $r$  is  $r = (3, 2, 2, 1, 1)$ . Figure 2 visualizes the sparse index set  $\mathcal{I}_{k,p}^r$ . For constructing the first order polynomials all five random variables (and the corresponding first-order Hermite polynomials) can be used. The second order polynomials are comprised by all possible combinations of first-order Hermite polynomials depending on  $\xi_1, \dots, \xi_5$  and the second order Hermite polynomials of  $\xi_1, \xi_2, \xi_3$ . Analogously, the third order polynomials are constructed.



**Fig. 2** Schematic example of a sparse index  $r = (3, 2, 2, 1, 1)$  in comparison with the full index set for  $k = 5$  random variables and maximum order  $p = 3$  of the chaos basis polynomials  $\Psi^\alpha(\xi)$ .

*Remark 4* By using this sparse index set  $\mathcal{I}_{k,p}^r$  the number of coefficient functions  $x_\alpha(t)$  appearing within the propagator can be reduced drastically without impairing the solution much. In the above Example 4 the full index set  $\mathcal{I}_{k,p}$  consists of  $\frac{(k+p)!}{k!p!} = 56$  terms [1,26], whereas the sparse truncated index set includes 42 terms (compare Tables 1/3).

An even better reduction can be achieved if we use an *adaptive* index  $r^p$  that depends on the actual order of the polynomials  $\Psi^\alpha$  with  $p = |\alpha|$ , e.g.,  $r^1 = (1, 1, 1, 1, 1)$ ,  $r^2 = (2, 2, 2, 1, 0)$ ,  $r^3 = (3, 2, 0, 0, 0)$ . That means, in constructing basis polynomials of order  $|\alpha| = 3$  we can use all combinations of Hermite polynomials depending on the first two random variables  $\xi_1$  and  $\xi_2$  up to orders 3 and 2, respectively. Thus, these are  $\sqrt{6}H_3(\xi_1)$ ,  $\sqrt{2}H_2(\xi_1)H_1(\xi_2)$ , and  $\sqrt{2}H_1(\xi_1)H_2(\xi_2)$  (see (6) and [26]).

In [26] there is an analysis of the errors made through truncating the propagator for several examples of solving stochastic partial differential equations via a Wiener chaos approach. In the context of stochastic optimal control this is a far more complex task. In Section 6 we give an impression of the error performance by numerical investigations.

While the usual sparse-grid ideas for polynomial chaos based on the Smolyak scheme [41,16] cannot be deployed directly in our intrusive-type method, alternative schemes can be developed from the *sparsity-of-effects* principle and the *least angle regression* or *compressed sensing* [5,29], which provides a more general truncation than the heuristic approach introduced above.

## 5 The Direct Multiple Shooting Approach

Before we take a closer look on numerical examples of solving stochastic optimal control problems by the introduced chaos approach, we give a short introduction to *Bock's direct multiple shooting approach* [6], which is a progression of Bulirsch's general multiple shooting idea [7]. It provides a state-of-the-art simultaneous method to solve optimization and simulation tasks at the same time. More information on this technique can be found, e. g., in [23].

The following deterministic optimal control problem represents the class of problems we want to solve in the following:

$$\min_{x,u} \int_{t_0}^{t_f} L(t, x(t), u(t)) dt \quad (29a)$$

$$\text{s.t.} \quad \dot{x}(t) = f(t, x(t), u(t)), \quad (29b)$$

$$x(t_0) = x_0, \quad (29c)$$

$$0 = r_e(x(t_0), x(t_f)), \quad (29d)$$

$$0 \leq r_i(x(t_0), x(t_f)), \quad (29e)$$

$$0 \leq g(t, x(t), u(t)), \quad (29f)$$

for  $t \in [t_0, t_f]$  almost everywhere, with differential states  $x: [t_0, t_f] \rightarrow \mathbb{R}^{n_x}$ , control functions  $u: [t_0, t_f] \rightarrow \mathbb{R}^{n_u}$ , and an objective function of Lagrange type. All functions considered are assumed to be sufficiently smooth.

Let  $x$  and  $u$  be the vectors of states  $x_i$  and controls  $u_i$ , then equation (29b) represents the ODE model with a right hand side  $f$  depending on time  $t \in [t_0, t_f]$ . Initial values  $x_0$  are given in (29c), (29d) and (29e) summarize (optional) equality respectively inequality boundary conditions, and (29f) contains optional state and path constraints.

The direct multiple shooting methods is based on a *first discretize, then optimize* approach to transform the control problem first to a nonlinear program (NLP), before this finite-dimensional optimization problem is solved to optimality.

The continuous controls are discretized by replacing them with base functions with local support, such as piecewise constant or piecewise linear functions. These functions can be described by finitely many parameters, i. e., we select a time grid

$$t_0 = \tau_0 < \tau_1 < \dots < \tau_m = t_f, \quad m \in \mathbb{N}$$

and with  $I_i := [\tau_i, \tau_{i+1}] \forall i \in \{0, \dots, m-1\}$  set

$$u(t) \Big|_{I_i} = \phi_i(t, w_i), \quad w_i \in \mathbb{R}^{\mu_i}, \quad (30)$$

where the  $\phi_i$  are the used base functions. Now we have transformed the infinite-dimensional control  $u$  into a finite vector  $w = (w_0, \dots, w_{m-1})$ .

The states  $x$  are discretized using *multiple shooting*. For efficiency and simplicity we choose the same time grid as for the controls. In theory, this is no limitation of generality, as we can refine the grids such that they match and add some constraints. We introduce  $m+1$  new variables  $s_0, \dots, s_m$  which represent the initial

values of the ODE on each interval  $I_i$ , respectively the final value  $s_m$ . Now we solve  $m$  independent initial value problems  $\forall i \in \{0, \dots, m-1\}$ ,

$$\dot{x}(t; \tau_i, s_i) = f(t, x_i(t), \phi_i(t, w_i)), \quad (31a)$$

$$x(\tau_i; \tau_i, s_i) = s_i, \quad (31b)$$

$$t \in [\tau_i, \tau_{i+1}]. \quad (31c)$$

For the numerical results presented in this paper, a Runge-Kutta-Fehlberg method has been used to solve the initial value problems.

To ensure equivalence to the original problem (29), we have to add *matching conditions*, which are the equality constraints

$$s_{i+1} = x(\tau_{i+1}; \tau_i, s_i) \quad \forall i \in \{0, \dots, m-1\}. \quad (32)$$

As the objective function is separable, it can be computed separately on each interval by

$$\int_{t_0}^{t_f} L(t, x(t), \phi(t, w)) \, dt = \sum_{i=0}^{m-1} L_i(\tau_{i+1}) \quad (33)$$

with

$$L_i(t) = \int_{\tau_i}^t L(t', x(t'; \tau_i, s_i), \phi_i(t', w_i)) \, dt'$$

and  $\phi(t, w) := \phi_i(t, w_i)$  for  $t \in I_i$ .

The optional continuous constraints  $g(t, x(t), u(t), p) \geq 0$  are evaluated pointwise on the grid, i. e.,

$$g(\tau_i, x(\tau_i; \tau_i, s_i), \phi_i(\tau_i, w_i)) \geq 0, \quad \forall i \in \{0, \dots, m\}.$$

Note that due to this the equivalence of the discretized problem to (29) is limited. However, [37] provides a possibility to check whether the path constraints (29f) are satisfied over the complete interval and reiterate, if necessary. The transformed boundary conditions and initial values read

$$\begin{aligned} r(s_0, s_m) &= 0, \\ s_0 &= x_0. \end{aligned}$$

The infinite-dimensional optimal control problem (29) has been transformed into a finite-dimensional NLP, which can be solved with a structure-exploiting sequential quadratic programming (SQP) method [38, 44].

## 6 Numerical Applications

### 6.1 A Linear Stochastic Regulator Problem

Our first example for solving optimal control problems driven by SDEs by the help of the novel chaos approach developed in Sections 3 and 4 is the standard linear-quadratic stochastic regulator problem [17, 33]. The advantage of this academic example is that we can solve the corresponding HJB equation, i. e., we have an exact solution to compare our results with.

We consider the one-dimensional stochastic regulator problem

$$\min_{u \in \mathcal{A}} \mathbb{E} \left[ \frac{1}{2} \int_0^1 (X_t^2 + u_t^2) dt + \frac{1}{2} X_1^2 \right] \quad (34a)$$

$$\text{s.t.} \quad dX_t = (X_t + u_t) dt + \sigma dB_t, \quad (34b)$$

$$X_0 = x_0, \quad (34c)$$

where the coefficient  $\sigma$  determining the diffusion term is merely a scalar. Then the optimal Markov feedback rule can be calculated as

$$u_t(\omega) = u^*(t, X_t(\omega)) = \left( \sqrt{2} \tanh \left( \sqrt{2}(t-1) \right) - 1 \right) \cdot X_t(\omega). \quad (35)$$

Thereby we have to keep in mind that the feedback rule at each instant of time  $t$  depends on the actual state of the system, as each such pair of time and state can be interpreted as the initial point of a separate problem. The Markov control  $u_t$  depends linearly on the state  $X_t$  of the system and explicitly on the time  $t$ . The optimal cost of the problem is

$$\Phi^*(t_0 = 0, x_0) = \frac{1}{2} \left( 1 + \sqrt{2} \tanh \sqrt{2} \right) x_0^2 + \frac{1}{2} \sigma^2 \left( 1 + \ln \cosh \sqrt{2} \right)$$

and the expectation and variance of the solution process can be calculated analytically as well.

Applying the propagator of Sections 3 and 4.1 to the SDE (34b) in its integral form, we deduce

$$\begin{aligned} x_\alpha(t) &= x_0 \cdot \mathbb{1}_{\{\alpha=0\}} + \int_0^t (x_\alpha(s) + u_\alpha(s)) ds \\ &\quad + \sum_{j=1}^{\infty} \int_0^t \sqrt{\alpha_j} m_j(s) \mathbb{E} \left[ \sigma \Psi^{\alpha^-(j)}(\xi) \right] ds, \end{aligned}$$

with the control coefficients  $u_\alpha(t)$  given by (23) up to some order  $q$ . The expectation value within the last summand is only  $\neq 0$  if  $\alpha^-(j) = 0$  or, equivalently,  $\alpha = \mathbf{e}_j$ ,  $j \geq 1$ . Therefore, the resulting system of ordinary differential equations reads

$$\begin{aligned} \dot{x}_\alpha(t) &= x_\alpha(t) + u_\alpha(t) + \sigma m_j(t) \cdot \mathbb{1}_{\{\alpha=\mathbf{e}_j\}}, \\ x_\alpha(0) &= x_0 \cdot \mathbb{1}_{\{\alpha=0\}}. \end{aligned}$$

As stated before, we can transform the objective function (34a) by directly inserting the chaos expansions (2) and (23) of  $X_t$  and the Markov control  $u_t$ . However, our investigations showed that it is numerically beneficial to convert Mayer-type objectives into their corresponding Lagrange form. Despite a slightly better convergence behavior, the computational costs are reduced notably.

Hence, applying Itô's formula [17] on  $z(X_t) = \frac{1}{2} X_t^2$  yields

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{2} X_1^2 \right] &= \frac{1}{2} x_0^2 + \mathbb{E} \left[ \int_0^1 \left( X_t (X_t + u_t) + \frac{1}{2} \sigma^2 \right) dt \right] + \underbrace{\mathbb{E} \left[ \int_0^1 \sigma X_t dB_t \right]}_{=0} \\ &= \frac{1}{2} x_0^2 + \frac{1}{2} \sigma^2 + \mathbb{E} \left[ \int_0^1 X_t (X_t + u_t) dt \right]. \end{aligned}$$



This changes the objective (34a) to

$$\begin{aligned}
& \mathbb{E} \left[ \frac{1}{2} \int_0^1 (X_t^2 + u_t^2) dt + \frac{1}{2} X_1^2 \right] \\
&= \frac{1}{2} (x_0^2 + \sigma^2) + \mathbb{E} \left[ \frac{1}{2} \int_0^1 (X_t^2 + u_t^2 + 2 X_t (X_t + u_t)) dt \right] \\
&= \frac{1}{2} (x_0^2 + \sigma^2) + \mathbb{E} \left[ \frac{1}{2} \int_0^1 ((X_t + u_t)^2 + 2 X_t^2) dt \right] \\
&= \frac{1}{2} (x_0^2 + \sigma^2) + \frac{1}{2} \int_0^1 \sum_{\alpha \in \mathcal{I}} ((x_\alpha(t) + u_\alpha(t))^2 + 2 x_\alpha^2(t)) dt.
\end{aligned}$$

Finally, for numerical investigations we have to truncate the index set  $\mathcal{I}$  appropriately. In the sequel we assume a quadratic approximation of the control rule, i. e.,  $q = 2$ . Remember that the exact control (35) is only linear in  $X_t$ . Additionally, we use different choices of (simply and adaptively) truncated index sets  $\mathcal{I}_{k,p}^r$ . We obtain the deterministic optimal control problem

$$\min_{\hat{u}_0(\cdot), \hat{u}_1(\cdot), \hat{u}_2(\cdot)} \left\{ \frac{1}{2} (x_0^2 + \sigma^2) + \frac{1}{2} \int_0^1 \sum_{\alpha \in \mathcal{I}_{k,p}^r} ((x_\alpha(t) + u_\alpha(t))^2 + 2 x_\alpha^2(t)) dt \right\} \quad (36a)$$

$$\text{s.t.} \quad \dot{x}_\alpha(t) = x_\alpha(t) + u_\alpha(t) + \sigma m_j(t) \cdot \mathbb{1}_{\{\alpha = \mathbf{e}_j\}} \quad (36b)$$

$$x_\alpha(0) = x_0 \cdot \mathbb{1}_{\{\alpha = \mathbf{0}\}} \quad (36c)$$

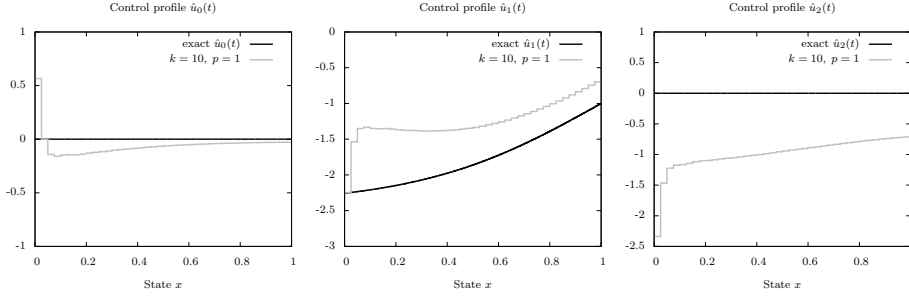
with

$$\begin{aligned}
u_\alpha(t) &= \hat{u}_0(t) \cdot \mathbb{1}_{\{\alpha = \mathbf{0}\}} + \hat{u}_1(t) x_\alpha(t) \\
&\quad + \hat{u}_2(t) \cdot \sum_{\beta \in \mathcal{I}_{k,p}^r} \sum_{0 \leq \gamma \leq \alpha} C(\alpha, \gamma, \beta) x_{\alpha - \gamma + \beta}(t) x_{\gamma + \beta}(t) \quad (36d)
\end{aligned}$$

as given in Example 2.

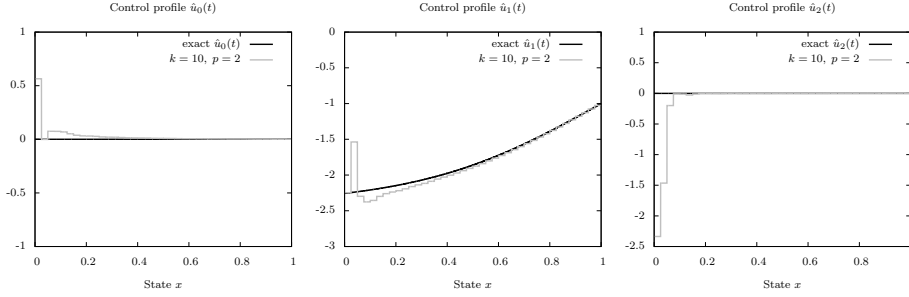
The resulting problem (36) can now be solved by sophisticated methods of deterministic optimal control as it does not explicitly involve random components anymore. All stochastic information is included within the system (36b). Our preferred choice is the software MUSCOD-II [22] based on Bock's direct multiple shooting approach [6], compare Section 5. The problem (36) includes  $|\mathcal{I}_{k,p}^r|$  state functions corresponding to the coefficients  $x_\alpha(t)$  of the chaos expansion and three control functions as we use a quadratic approximation of the feedback rule. The following numerical experiments have been performed using the initial values  $x_0 = \frac{1}{2}$  and  $x_0 = 1$ , the diffusion parameter  $\sigma = 0.15$ , and different truncation numbers  $k$  and approximation orders  $p$  of the chaos expansion.

Figures 3 and 4 illustrate the behavior of the new control functions  $\hat{u}_i(t)$ ,  $i = 0, \dots, p$ , that we introduce to preserve the feedback character of the Markov control  $u_t = u(t, X_t)$  of the original stochastic problem (34). Note that we use a quadratic expansion (23) although the exact feedback rule is only linear in  $X_t$  (compare (35)). The solutions shown in Figure 3 are computed by a Gaussian approximation of the chaos space, i. e., by truncating the index set  $\mathcal{I}$  with an



**Fig. 3** Optimal controls of the linear-quadratic stochastic regulator problem (34). The plots depict solutions of the deterministic optimal control problem (36) resulting from the chaos methodology, i. e., the new control functions  $\hat{u}_i(t)$ ,  $i = 0, \dots, q$  introduced in the expansion (23) with  $q = 2$  to preserve the non-anticipativity of the Markov control in its chaos expansion. In comparison the exact functions (compare (35)) are shown.

Here the new control functions are obtained from truncating the index set  $\mathcal{I}$  of the chaos expansion with  $k = 10$  random variables and approximation order  $p = 1$ , resulting in eleven basis polynomials that describe the stochastic system. Because of that simple Gaussian approximation the quadratic expansion of the Markov control collapses to a linear one, whereas these apparently wrong solutions yield good results.



**Fig. 4** Optimal controls of the linear-quadratic stochastic regulator problem (34) as in Figure 3. Now the controls are computed for  $k = 10$  random variables and approximation order  $p = 2$ . Thus, we come very close to the desired results, including  $\hat{u}_2(t) \approx 0$  as supposed.

approximation order  $p = 1$ . Hence, the system of control coefficients  $u_\alpha(t)$  induced by (23) as in Example 2 is not quadratic in the state coefficients  $x_\alpha(t)$  anymore. To justify this, we calculate

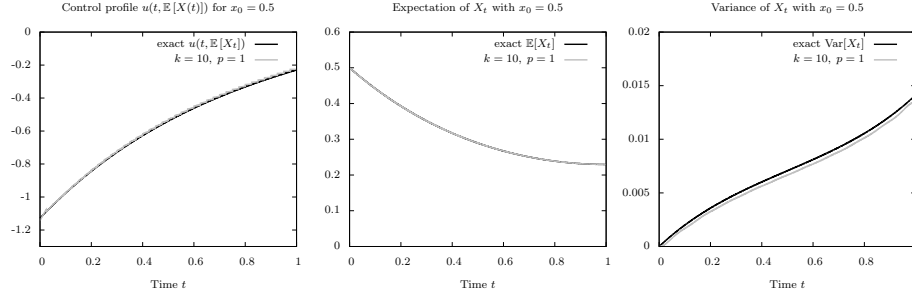
$$u_{\mathbf{0}}(t) = \hat{u}_0(t) + \hat{u}_1(t) x_{\mathbf{0}}(t) + \hat{u}_2(t) \cdot \sum_{\beta \in \mathcal{I}_{k,1}} C(\mathbf{0}, \mathbf{0}, \beta) x_\beta^2(t),$$

where the last term can merely be seen as a multiple of the process's variance plus the quadratic expectation, and for  $\alpha \neq \mathbf{0}$

$$\begin{aligned} u_\alpha(t) &= \hat{u}_1(t) x_\alpha(t) \\ &\quad + \hat{u}_2(t) \cdot \sum_{\beta \in \mathcal{I}_{k,1}} (C(\alpha, \mathbf{0}, \beta) x_{\alpha+\beta}(t) x_\beta(t) + C(\alpha, \alpha, \beta) x_\beta(t) x_{\alpha+\beta}(t)) \\ &= \hat{u}_1(t) x_\alpha(t) + 2\hat{u}_2(t) x_{\mathbf{0}}(t) x_\alpha(t), \end{aligned}$$

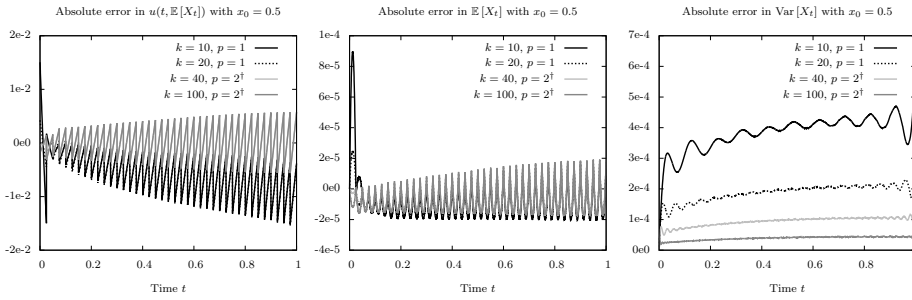
as all appearing coefficients  $\alpha$ ,  $\beta$ , and  $\alpha + \beta$  have to be within the index set  $\mathcal{I}_{k,1}$ . This explains the differences of the exact solutions and the ones shown in the figure.

Moreover, only  $k = 10$  random variables  $\xi_i$  are included in the construction of the basis polynomials  $\Psi^\alpha$  used to obtain the plotted solutions. By increasing the order  $p$ , the system of control coefficients becomes quadratic in the state coefficients, which is why the solutions shown in Figure 4 (computed with  $k = 10$  and  $p = 2$ ) come closer to the exact ones, including  $\hat{u}_2(t) \approx 0$  as anticipated.



**Fig. 5** Solution paths of the linear-quadratic stochastic regulator problem (34) with initial value  $x_0 = 0.5$ . All plots show again a comparison of exact solutions and the corresponding path obtained by the introduced chaos approach, i. e., a solution to the transformed deterministic optimal control problem (36) with  $k = 10$  included random variables and an approximation order  $p = 1$ .

The first plot depicts the control profile  $u(t, \mathbb{E}[X_t])$ , the second one the expectation of the state process  $\mathbb{E}[X_t]$ , and the third figure its variance  $\text{Var}[X_t]$ .



**Fig. 6** Absolute errors of the solutions of the linear-quadratic stochastic regulator problem (34) computed by the novel chaos approach for different numbers  $k$  of involved random variables and orders  $p$  to construct the basis polynomials  $\Psi^\alpha(\xi)$ . The sequence of plots is as in Figure 5, i. e., the first showing the errors within the control profile  $u(t, \mathbb{E}[X_t])$ , the second errors in the expectation  $\mathbb{E}[X_t]$ , and the third errors in the variance  $\text{Var}[X_t]$ .

Figure 5 shows different solution paths of the transformed deterministic optimal control problem (36) in comparison with the appropriate exact solutions of the original stochastic problem (34) for given initial values  $x_0$ . Again the results of the chaos approach are obtained with the simple truncation (27) using  $k = 10$  random variables and an approximation order  $p = 1$  for constructing the basis polynomials  $\Psi^\alpha(\xi)$  with  $\alpha \in \mathcal{I}_{k,p}$ .

Within the figure the first plot depicts the optimal control profile depending on the time  $t \in [0, 1]$  and the expectation of the process at that time, i. e., the control

**Table 1** Comparison of optimal values and numerical expenses for solving the deterministic optimal control problem (36) for initial values  $x_0 = 0.5$  (columns 5–7) and  $x_0 = 1$  (columns 8–10) depending on the type and accuracy of truncating the index set  $\mathcal{I}$ . We notice that the accuracy of the objective function value mainly depends on the number  $k$  of incorporated random variables  $\xi_i$ . Runtime increases with the dimension of the resulting deterministic problems and the associated coupling of the state variables within the system.

The dagger symbol  $\dagger$  denotes the adaptive index sets (depending on  $k$  and  $p$ ) that are used in Figure 6. The symbol “–” in the  $r$ -column indicates that the simple truncation (27) was used, “sp” marks the use of a sparse and “ad” of an adaptive index (compare Remark 4 and Table 4 for a detailed description of the appropriate index denoted by the reference number).

$k$	$p$	$r$	# coeff. $x_\alpha$	objective value	time in s	# SQP	objective value	time in s	# SQP
				$x_0 = 0.5$			$x_0 = 1.0$		
10	1	–	11	0.301731	2.0	44	1.147880	2.7	52
10	2	–	66	0.301731	186.8	135	1.147879	360.5	300
10	2	sp <sup>2</sup>	61	0.301731	117.7	101	1.147880	103.1	111
20	1	–	21	0.301898	8.8	46	1.148046	14.4	70
20	2	–	231	0.301898	3341.1	103	1.148046	4948.7	150
20	2	ad <sup>2</sup>	71	0.301898	168.5	90	1.148046	240.1	124
20	3	ad <sup>3</sup>	125	0.301898	1053.3	119	1.148046	1020.7	103
40	1	–	41	0.301979	57.8	54	1.148127	85.3	76
40	2 <sup>†</sup>	ad <sup>4</sup>	91	0.301979	694.6	100	1.148127	1211.0	233
100	1	–	101	0.302027	585.6	48	1.148174	888.7	71
100	2 <sup>†</sup>	ad <sup>6</sup>	151	0.302026	3669.9	132	1.148175	3233.5	119
exact				0.302054			1.148191		

$u(t, \mathbb{E}[X_t | X_0 = x_0])$ . By viewing this uncommon profile we get an impression of the accuracy of the numerically obtained control at states where the process will be most likely at time  $t$ . The remaining two plots of Figure 5 show the corresponding expectation and variance of the process.

From purely visual comparison we see how well the introduced chaos method works, even for very low approximations of the Wiener Chaos space and even as the new control functions  $\hat{u}_i(t)$ ,  $i = 0, \dots, q$ , deviate from their exact counterparts as we saw in Figure 3. This holds especially if we are interested in calculating the objective, expectations and possibly higher moments of the solution process for a given initial value to the original problem because they are a direct byproduct of the new methodology.

Figure 6 illustrates the absolute errors of  $u(t, \mathbb{E}[X_t])$ ,  $\mathbb{E}[X_t]$ , and  $\text{Var}[X_t]$  over time for  $x_0 = 0.5$  and different choices of truncation. We notice that the error decreases if the number of random variables  $k$  and the approximation order  $p$  are increased. E. g., the absolute error of the expectation process  $\mathbb{E}[X_t]$  in the time interval  $[0, 1]$  is at most  $1 \cdot 10^{-4}$  for the low approximation  $(k, p) = (10, 1)$  and decays to  $2 \cdot 10^{-5}$  for  $(k, p) = (40, 2)$ , which is very astonishing. Particularly the enhancement of the approximation order has a great influence on the error performance, which is seen most clearly in the error plots of the control profile  $u(t, \mathbb{E}[X_t])$  and the variance  $\text{Var}[X_t]$ . The jagged behavior of the graphs is due to our choice of constant control base functions (30) for  $\hat{u}_i(t)$ ,  $i = 0, \dots, q$ , on each multiple shooting interval, which is carried over to all solution processes.

Furthermore, Table 1 presents additional information about the performance of the chaos approach for solving stochastic optimal control problems, depending on the type and accuracy of the truncation of the index set. Therein we see that (at least in this first example) the order  $p$  of the used basis polynomials  $\Psi^\alpha$  and their corresponding state coefficient functions  $x_\alpha$  is less important than the number of incorporated random variables  $k$  if we desire a good result of the objective function. Moreover, Table 1 gives the dimensions of the resulting deterministic optimal control problems (36) and the computational effort to solve them numerically. Note that by using sparse or adaptive index sets  $\mathcal{I}_{k,p}^r$  the number of coefficient functions within the deterministic system (and, therefore, computation time) can partly be reduced drastically without impairing the solution. The most astonishing result is that if we are interested in the objective value, the expectation of the resulting state process, and its variance for a given initial value  $x_0$ , we can obtain these items with very little effort, the appearing relevant systems can be solved in a few seconds.

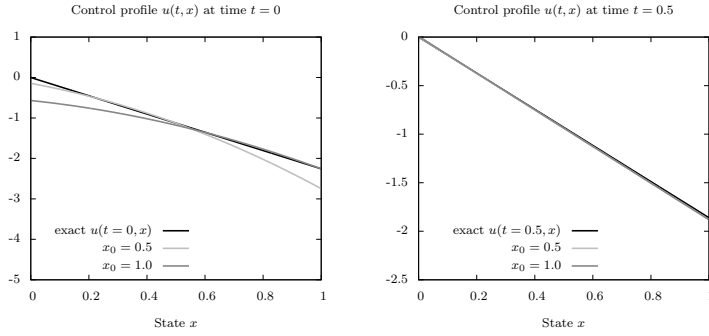
However, if one is not only interested in the solution to the stochastic optimal control problem for one certain initial value  $x_0$ , but possibly for an environment of  $x_0$ , the low approximation of Figure 5 is too inaccurate, as Figure 8 illustrates.

From the left plot we see that using a low chaos approximation, e.g., Gaussian ( $p = 1$ ) with  $k = 10$  random variables, the control obtained via solving the resulting deterministic optimal control problem (36) is only accurate for the initial value  $x_0$  employed. When we move further in time the control is very precise for states that the process will attain most likely (see again the left plots in Figure 5), but additionally there is a certain robustness against deviations from that states. This is natural due to the randomness that is implicitly captured within the deterministic system of differential equations (36b). Nevertheless, if we are interested in applying one (optimal) control—that is obtained through one specified initial value  $x_0$ —for several control problems depending on a whole environment of initial values around  $x_0$ , a low chaos approximation is useless. In that case more information of the stochastic behavior of the system is needed within the deterministic transformation. In particular, the crucial factor of a better approximation here is the order  $p$  rather than the number of incorporated random variables  $\xi_i$ ,  $i = 1, \dots, k$ . In this first example it is sufficient to apply the following truncation with  $k = 40$ ,  $p = 2$ , and an adaptive index  $r$  (compare Remark 4) to obtain the desired robustness property of the optimal control.

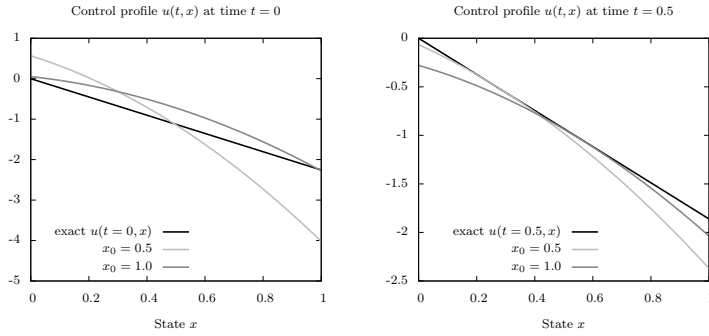
$$\begin{aligned} \mathcal{I}_{40,2}^r = \{ & \alpha = (\alpha_1, \dots, \alpha_{40}) \mid 0 \leq \alpha_i \leq r_i^l \ \forall i \in \{1, \dots, k\} \ \forall l \in \{0, 1, 2\}, |\alpha| \leq 2; \\ & r^0 = \mathbf{0}, r^1 = (1, \dots, 1), r^2 = (\underbrace{2, \dots, 2}_5, \underbrace{1, \dots, 1}_5, \underbrace{0, \dots, 0}_{30}) \} \end{aligned} \quad (37)$$

Figure 7 validates this. In general, we notice a connection of the behavior of the control profiles  $u(t, x)$ ,  $t$  fixed, shown in Figures 8 and 7 and the new control functions  $\hat{u}_i(t)$ ,  $i = 0, \dots, q$ , in Figures 3 and 4. The better those new control functions coincide with their exact counterparts, the better the state dependent profiles at fixed time instants fit and the more robust the solutions become.

Altogether, this simple example shows that up to this point the results of our novel chaos reformulation of a continuous finite-horizon stochastic optimal control problem as a deterministic optimal control problem are very promising. In the next section we consider a problem that cannot be solved analytically anymore.



**Fig. 7** Control profiles as in Figure 8 but for an advanced chaos approximation (37).



**Fig. 8** Control profiles of the linear-quadratic stochastic regulator problem (34) computed by the novel chaos approach in comparison with the exact solutions. Both plots show controls  $u(t, x)$  for fixed time instants  $t$  depending on the state  $x$ . They are calculated with a quadratic approximation of the Markov control and a low truncation of the index set  $\mathcal{I}$ , i. e.,  $k = 10$  and  $p = 1$ .

One notices that using this truncation the controls at time  $t = 0$  are only accurate for the initial value  $x_0$  of the solved deterministic problem. In the course of time this behavior gets better, which is due to the implicit capture of  $X_t$ 's variance within the deterministic system.

In fact, the state process  $\{X_t\}$  is determined by a stochastic differential equation with state dependent diffusion term and a drift that is a nonlinear combination of  $X_t$  and the control.

## 6.2 A Nonlinear Example

Let us consider a stochastic optimal control problem with the same objective function as for the stochastic regulator of the previous section, but with a nonlinear diffusion driving the state process  $\{X_t\}_{t \in [0,1]}$ .

$$\min_{u \in \mathcal{A}} \mathbb{E} \left[ \frac{1}{2} \int_0^1 (X_t^2 + u_t^2) dt + \frac{1}{2} X_1^2 \right] \quad (38a)$$

$$\text{s.t. } dX_t = X_t u_t dt + \sigma X_t dB_t, \quad (38b)$$

$$X_0 = x_0. \quad (38c)$$

Because of this enhancement, problem (38) cannot be solved analytically. Nevertheless, in [12] it is shown that a solution to (38) exists, even in a more general

formulation. Therefore we can apply our chaos methodology again to transform this stochastic control problem into a deterministic one. Then the propagator of the SDE (38b) reads as follows (with  $C(\alpha, \gamma, \beta)$  and  $\alpha^-(j)$  defined as before and  $\{m_i(t)\}$  denoting the basis functions of  $L^2([0, T])$ ):

$$\begin{aligned} x_\alpha(t) = & x_0 \cdot \mathbb{1}_{\{\alpha \equiv \mathbf{0}\}} + \int_0^t \sum_{\beta \in \mathcal{I}} \sum_{0 \leq \gamma \leq \alpha} C(\alpha, \gamma, \beta) x_{\alpha-\gamma+\beta}(s) u_{\gamma+\beta}(s) ds \\ & + \sigma \int_0^t \sum_{j=1}^{\infty} \sqrt{\alpha_j} m_j(s) x_{\alpha^-(j)}(s) ds. \end{aligned} \quad (39)$$

Therein, the control coefficients  $u_\alpha(\cdot)$  are again defined via their feedback formulation (23) depending on the new control functions  $\hat{u}_i(\cdot)$ ,  $i = 0, \dots, q$ . The first integral in (39) follows from the chaos expansion of  $X_t u_t$ .

To reformulate the objective function (38a) in terms of the deterministic coefficient functions, we start again by converting the Mayer-type part using Itô's formula. Then inserting the chaos expansions of  $X_t$  and  $u_t$  yields the desired form. With  $\mathcal{I}_{k,p}^r$  denoting the truncated index set as before and approximating the Markov control by a quadratic expansion (compare (36d)), we obtain the deterministic optimal control problem corresponding to (38),

$$\min_{\hat{u}_0(\cdot), \hat{u}_1(\cdot), \hat{u}_2(\cdot)} \left\{ \frac{1}{2} x_0^2 + \frac{1}{2} \int_0^1 \sum_{\alpha \in \mathcal{I}_{k,p}^r} \left[ (1 + \sigma^2) x_\alpha^2(t) + u_\alpha^2(t) \right. \right. \quad (40a)$$

$$\left. \left. + 2 \sum_{\beta \in \mathcal{I}_{k,p}^r} \sum_{0 \leq \gamma \leq \alpha} C(\alpha, \gamma, \beta) x_{\alpha-\gamma+\beta}(t) x_{\gamma+\beta}(t) u_\alpha(t) \right] dt \right\}$$

$$\begin{aligned} \text{s.t. } \dot{x}_\alpha(t) = & \sum_{\beta \in \mathcal{I}_{k,p}^r} \sum_{0 \leq \gamma \leq \alpha} C(\alpha, \gamma, \beta) x_{\alpha-\gamma+\beta}(t) u_{\gamma+\beta}(t) \\ & + \sigma \sum_{j=1}^{\infty} \sqrt{\alpha_j} m_j(t) x_{\alpha^-(j)}(t) \end{aligned} \quad (40b)$$

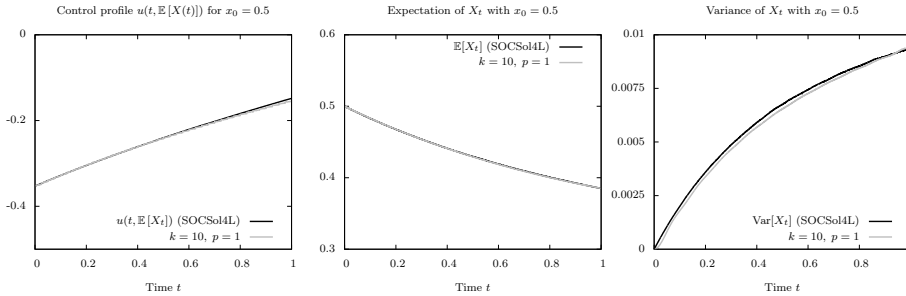
$$x_\alpha(0) = x_0 \cdot \mathbb{1}_{\{\alpha \equiv \mathbf{0}\}}. \quad (40c)$$

As the solution of problem (38) cannot be deduced analytically, we have to compare the results of our chaos approach with other numerical methods. It is intricate to solve the partial differential HJB equation induced by (38) numerically because we have no information about appropriate boundary conditions for the region of interest. In financial problems this can often be overcome by economic argumentation, however, in this case it is not possible. Therefore, we use the software package *SOCsol4L* [3, 19] for obtaining reference solutions. It transforms the original continuous stochastic control problem into a Markov decision chain by utilizing the *weak Euler-Maruyama approximation* scheme of the SDE on a predefined time and space grid for the region  $\mathcal{G}$  of interest. Afterwards, this Markov decision chain problem can be solved by a dynamic programming technique. Table 2 gives an overview of optimal values obtained with that software and the computational effort needed therefor. Note that while our new chaos methodology provides the expectation of the process  $\{X_t\}$  and the optimal cost automatically,

within SOCSol4L these quantities have to be approximated by using a Monte Carlo simulation.

**Table 2** Optimal values of problem (38) calculated with the software package *SOCSol4L*. The problem was solved in the state space  $\mathcal{G} = [-0.7, 1.2]$  for different space and time discretizations  $\Delta_x$  and  $\Delta_t$ . After calculating optimal policies, the optimal values have been approximated by a Monte Carlo simulation with 100 000 and 1 million sample paths and different simulation step sizes  $\Delta_{\text{Sim}}$  for the weak approximation scheme of the SDE. Each simulation therefore gives a different result. The runtimes (in *min*) include both solving the Markov decision process by a dynamic programming technique and performing the Monte Carlo simulation to eventually obtain the desired result.

discretization	# simulations	$\Delta_{\text{Sim}}$	optimal value	runtime in <i>min</i>
$\Delta_x = 0.005$ $\Delta_t = 0.001$	100 000	0.01	0.2113440	100
			0.2117336	
			0.2115376	
	1 000 000		0.2115948	925
$\Delta_x = 0.002$ $\Delta_t = 0.001$	100 000	0.001	0.2112635	1 000
			1 000 000	0.2114620



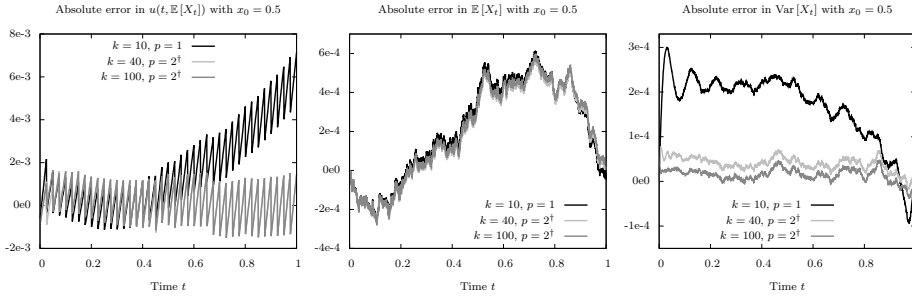
**Fig. 9** Solution paths of the stochastic optimal control problem (38) with initial value  $x_0 = 0.5$ . All plots show a comparison of paths obtained by the introduced chaos approach, i. e., a solution to the transformed deterministic optimal control problem (40) with  $k = 10$  included random variables and approximation order  $p = 1$ , and a reference solution obtained with SOCSol4L ( $\mathcal{G} = [-0.7, 1.2]$ ,  $\Delta_x = 0.002$ ,  $\Delta_t = 0.001$  and a Monte Carlo simulation with 300 000 samples).

The first plot depicts the control profile  $u(t, \mathbb{E}[X_t])$ , the second one the expectation of the state process  $\mathbb{E}[X_t]$ , and the third figure its variance  $\text{Var}[X_t]$ .

The reference solution that is used within the following figures is obtained with a state discretization step size  $\Delta_x = 0.002$  and a time discretization step  $\Delta_t = 0.001$ . All expectations and variances are simulated with 300 000 sample paths.

Figure 9 shows solution paths of the deterministic optimal control problem (40) in comparison with this reference solution for the initial value  $x_0 = 0.5$  and a diffusion parameter  $\sigma = 0.3$ . These solutions are again obtained with a simple truncation and  $k = 10$  random variables and order  $p = 1$  for constructing the chaos basis polynomials  $\Psi^\alpha(\xi)$ . The order of the plots within is as in Figures 5–6, i. e., the left plot depicting the control  $u(t, \mathbb{E}[X_t])$ , the middle one the expectation  $\mathbb{E}[X_t]$





**Fig. 10** Absolute errors of the solutions of the stochastic control problem (38) computed by the novel chaos approach for different numbers  $k$  of involved random variables and orders  $p$  to construct the basis polynomials  $\Psi^\alpha(\xi)$  (in comparison to the reference solutions calculated with SOCSol4L). The sequence of plots is as in Figure 5, i. e., the first showing the errors within the control profile  $u(t, \mathbb{E}[X_t])$ , the second errors in the expectation  $\mathbb{E}[X_t]$ , and the third errors in the variance  $\text{Var}[X_t]$ .

**Table 3** Optimal values and numerical expenses for solving the stochastic optimal control problem (38) with the chaos methodology, i. e., solving the deterministic problem (40) with MUSCOD-II. We use the initial values  $x_0 = 0.5$  (columns 5–7) and  $x_0 = 1$  (columns 8–10), the diffusion parameter  $\sigma = 0.3$ , and different types and accuracies of truncating the index set  $\mathcal{I}$ . The accuracy of the objective function values does not only depend on the number  $k$  of incorporated random variables, but as well on the approximation order  $p$ .

The dagger symbol  $\dagger$  denotes again the adaptive index sets that are used in Figure 10. The symbol “–” in the  $r$ -column indicates that the simple truncation (27) was used, “sp” marks the use of a sparse index (4), and “ad” of an adaptive index (compare Tables 1 and 4).

$k$	$p$	$r$	# coeff. $x_\alpha$	$x_0 = 0.5$			$x_0 = 1.0$		
				objective value	time in s	# SQP	objective value	time in s	# SQP
5	2	–	21	0.211509	42.2	234	0.620330	46.7	224
5	3	–	56	0.211502	476.0	215	0.620325	330.6	143
5	3	sp <sup>1</sup>	42	0.211503	170.2	156	0.620326	190.8	163
10	1	–	11	0.211733	6.9	90	0.620489	3.6	46
10	2	–	66	0.211458	759.6	212	0.619768	740.1	192
10	2	sp <sup>2</sup>	61	0.211458	631.2	202	0.619769	544.8	161
10	3	–	286	0.211451	27311.5	144	0.619761	25650.7	123
10	3	ad <sup>1</sup>	42	0.211462	233.3	165	0.619799	189.2	120
20	1	–	21	0.211714	14.0	42	0.620214	19.0	53
20	2	–	231	0.211432	18293.5	179	0.619469	11484.4	79
20	2	sp <sup>3</sup>	216	0.211432	13557.4	121	0.619470	22281.2	179
20	2	ad <sup>2</sup>	71	0.211439	1557.2	241	0.619498	1337.0	187
20	3	ad <sup>3</sup>	125	0.211433	7041.0	254	0.619493	5933.2	196
40	2 <sup>†</sup>	ad <sup>4</sup>	91	0.211430	2256.8	139	0.619359	2813.2	156
40	3	ad <sup>5</sup>	145	0.211423	9300.2	118	0.619353	13341.8	153
100	1	–	101	0.211698	1962.9	79	0.619986	3389.7	125
100	2 <sup>†</sup>	ad <sup>6</sup>	151	0.211424	12985.6	145	0.619273	14423.5	149
approx. (300 000 sim.)				0.211707	169159.0		0.619434	168315.7	

and the right one the variance  $\text{Var}[X_t]$ . And again these paths show that by this low approximation we obtain very good results if we are interested in the optimal value,

expectation, and related quantities, even as the state equation (38b) is much more complex than in the regulator problem and, hence, the deterministic system (40b) much more coupled. This quality of the solution is confirmed by the corresponding error plots in Figure 10 and the optimal values stated in Table 3. The very noisy shape of the absolute errors is caused by the Monte Carlo approximation—even with the large amount of 300 000 sample simulations the values deviate notably (compare Table 2).

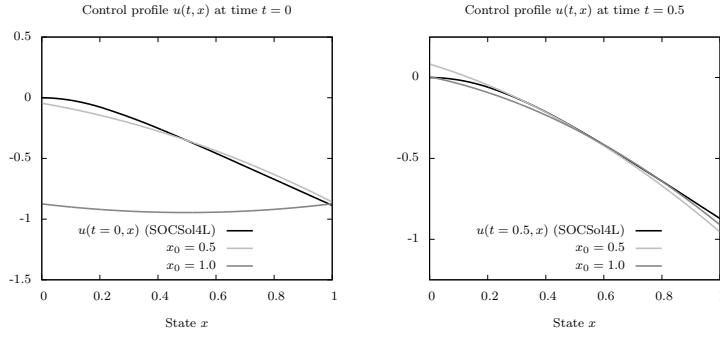
Like for the regulator problem we notice from Figure 10 that especially increasing the approximation order  $p$  leads to a decrease of the absolute errors in the control profile  $u(t, \mathbb{E}[X_t])$  and the variance  $\text{Var}[X_t]$ , while the expectation is already very accurately approximated with order  $p = 1$ . If we compare the optimal objective values in Table 3 we see that for this nonlinear example it is not sufficient to increase the number of incorporated random variables  $\xi_i$ ,  $i = 1, \dots, k$ , to obtain better results; here increasing the order  $p$  is important as well.

From a computational point of view Table 3 shows that the deterministic optimal control problem (40) is much more challenging than (36). While the number of SQP iterations needed to solve the problems remains at a comparable level, the runtime for (40) is notably higher. This results from a distinctive coupling of the deterministic state functions  $x_\alpha$  within the ODE system (40b) and, therefore, a higher expense for calculating their derivatives.

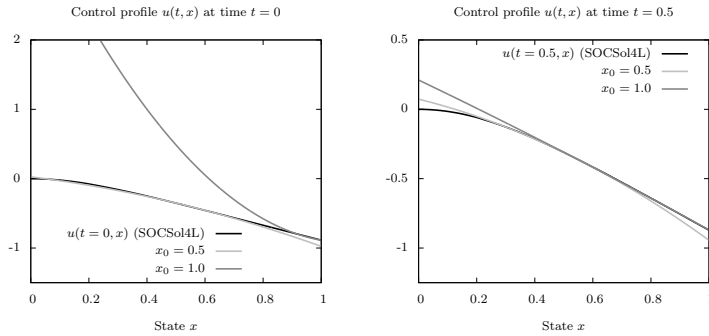
Finally, let us take a look on the control profiles for fixed times  $t$  depending on the state. As in the previous section, the control obtained via the low truncation ( $k = 10$ ,  $p = 1$ ) does not allow its application to initial values deviating from  $x_0$ , see Figure 11. If we want to guarantee a certain robustness of the validity of a control  $u$  (calculated with initial value  $x_0$ ) for applying it to initial values in an environment of  $x_0$ , we have to enhance the accuracy of the chaos approximation. Figure 12 shows the control profiles  $u(t, x)$  calculated for the adaptive truncation (37), i. e.,  $k = 40$ ,  $p = 2$ ,  $r$  given by  $\text{ad}^4$ . However, in comparison to the quadratic regulator problem the impact of this enhancement turns out lower. This might originate from the fact that the variance of the process  $\{X_t\}$  in the actual example is generally smaller than the variance of the regulator process, hence the process will not deviate that heavily from its expectation. Nevertheless, qualitatively better approximations can be obtained by increasing  $k$  and  $p$  or the order  $q$  of the Markov control expansion further—at the cost of higher computation times.

## 7 Results

In this paper we developed a novel generic methodology to solve finite horizon stochastic optimal control problems. Using Wiener’s chaos expansion and Malliavin calculus we are able to transform the underlying stochastic differential equations driving the state process into a system of ordinary differential equations. Additionally, we ensure the feedback character of the Markov control process by expanding it in a Taylor-like fashion. Hence, after reformulating the original objective function in terms of the chaos expansion coefficients, we obtain a deterministic optimal control problem that implicitly contains all the random information of the original stochastic problem. This resulting problem then can be solved efficiently by sophisticated methods of deterministic control.



**Fig. 11** Control profiles of the control problem (38) computed by the novel chaos approach in comparison with the reference solutions. Both plots show controls  $u(t, x)$  for fixed time instants  $t$  depending on the state  $x$ . They are calculated with a quadratic approximation of the Markov control and a low truncation of the index set  $\mathcal{I}$ , i. e.,  $k = 10$  and  $p = 1$ . One notices again that using this truncation the controls at time  $t = 0$  are only accurate for the initial value  $x_0$  of the solved deterministic problem.



**Fig. 12** Control profiles as in Figure 11 but for the advanced chaos approximation (37) with  $k = 40$  included random variables and approximation order  $p = 2$ , using an adaptive index  $r$ .

Two numerical examples show that our approach yields very promising results. If one is mainly interested in the optimal objective value, the expectation of the optimally controlled state process, and corresponding higher moments for only one or a small number of initial values, one can obtain fast and reliable results with quite low approximations of the chaos space. At that, making use of sparse and adaptive truncation schemes is beneficial to reduce the overall computational effort without impairing the obtained solution. If the optimizer instead emphasizes the control profile calculated for a fixed initial value to be robust against deviations in the initial values to some extent, he has to apply more accurate approximations of the chaos with more incorporated random variables and a higher order.

However, for more general problems including higher non-linearities or long-time integration intervals  $[0, T]$ , higher numerical effort may become necessary [2, 15]. Approaches to overcome this issue are, e. g., partitioning of the random space or combinations of the Wiener chaos idea with targeted Monte Carlo corrections [26].

A very important feature of the introduced chaos approach is that we obtain the objective value, expectations, variances, and all higher moments as a byproduct of solving the resulting deterministic control problem. Hence we do not need any simulation to deduce these quantities.

Therefore, we get an efficient alternative in solving this challenging class of problems, apart from the Hamilton-Jacobi-Bellman theory and dynamic programming approaches based on Markov chain approximations or quantization of the stochastic processes. Furthermore, we open the field of applying state-of-the-art methods of deterministic optimization and control in the broad context of random processes and stochastic differential equations.

### Acknowledgements

The authors like to express their gratitude to *Prof. M. Podolskij* from the University of Heidelberg for his helpful comments and suggestions.

This research was supported by the Heidelberg Graduate School *Mathematical and Computational Methods for the Sciences* and by the *European Union Seventh Framework Programme FP7/2007-2013* under grant agreement *n° FP7-ICT-2009-4 248940*.

### References

1. F. Augustin, A. Gilg, M. Paffrath, P. Rentrop, and U. Wever. Polynomial chaos for the approximation of uncertainties: Chances and limits. *European Journal of Applied Mathematics*, 19:149–190, 2008.
2. F. Augustin and P. Rentrop. Stochastic Galerkin Techniques for Random Ordinary Differential Equations. *Numerische Mathematik*, 122:399–419, 2012.
3. J.D. Azzato and J.B. Krawczyk. SOCSol4L: An Improved MATLAB Package For Approximating The Solution To A Continuous-Time Stochastic Optimal Control Problem. Working Paper of the Victoria University of Wellington, 2006.
4. M. Barczy and E. Iglói. Karhunen-Loève expansions of  $\alpha$ -Wiener bridges. *Central European Journal of Mathematics*, 9(1):65–84, 2011.
5. G. Blatman and B. Sudret. Sparse Polynomial Chaos Expansions Based on an Adaptive Least Angle Regression Algorithm. In *19ème Congrès Français de Mécanique*, Marseille, 2009.
6. H.G. Bock and K.J. Plitt. A Multiple Shooting algorithm for direct solution of optimal control problems. In *Proceedings of the 9th IFAC World Congress*, pages 242–247, Budapest, 1984. Pergamon Press. Available at <http://www.iwr.uni-heidelberg.de/groups/agbock/FILES/Bock1984.pdf>.
7. R. Bulirsch. Die Mehrzielmethode zur numerischen Lösung von nichtlinearen Randwertproblemen und Aufgaben der optimalen Steuerung. Technical report, Carl-Cranz-Gesellschaft, Oberpfaffenhofen, 1971.
8. R.H. Cameron and W.T. Martin. The Orthogonal Development of Non-Linear Functionals in Series of Fourier-Hermite Functionals. *Annals of Mathematics*, 48(2):385–392, April 1947.
9. W.H. Fleming and R.W. Rishel. *Deterministic and Stochastic Optimal Control*. Springer New York, 1982.
10. Philipp Frauenfelder, Christoph Schwab, and Radu Alexandru Todor. Finite elements for elliptic problems with stochastic coefficients. *Comput. Methods Appl. Mech. Engrg.*, 194:205–228, 2005.
11. R.G. Ghanem and P.D. Spanos. *Stochastic Finite Elements*. Springer, Heidelberg, 1991.
12. Floyd B. Hanson. *Stochastic Digital Control System Techniques*, volume 76 of *Control and Dynamic Systems: Advances in Theory and Applications*, chapter Techniques in Computational Stochastic Dynamic Programming, pages 103–162. Academic Press, New York, 1996.
13. S. Janson. *Gaussian-Hilbert Spaces*. Cambridge University Press, 1997.
14. K. Karhunen. *Über lineare Methoden in der Wahrscheinlichkeitsrechnung*, volume 37. *Annales Academiæ Scientiarum Fennicæ, Series A.I*, 1947.

15. G. E. Karniadikis, C.-H. Su, D. Xiu, D. Lucor, C. Schwab, and R. A. Todor. Generalized Polynomial Chaos Solution for Differential Equations with Random Inputs. Report Nr. 2005-01, ETH Zürich, 2005.
16. A. Keese and H. G. Matthies. Numerical methods and Smolyak quadrature for nonlinear stochastic partial differential equations. Informatikbericht Nr.: 2003-5, TU Braunschweig, 2003.
17. P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin, 3rd edition, 1999.
18. R. Korn and E. Korn. *Option Pricing and Portfolio Optimization*. Graduate Studies in Mathematics Volume 31. Oxford University Press, 2001.
19. J.B. Krawczyk. A Markovian Approximated Solution To A Portfolio Management Problem. *ITEM*, 1(1), 2001.
20. H.J. Kushner. Numerical Methods for Stochastic Control Problems in Continuous Time. *SIAM Journal on Control & Optimization*, 28(5):999–1048, 1990.
21. H.J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, 2001.
22. D.B. Leineweber. Analyse und Restrukturierung eines Verfahrens zur direkten Lösung von Optimal-Steuerungsproblemen. Diploma thesis, Ruprecht-Karls-Universität Heidelberg, 1995.
23. D.B. Leineweber, I. Bauer, H.G. Bock, and J.P. Schlöder. An Efficient Multiple Shooting Based Reduced SQP Strategy for Large-Scale Dynamic Process Optimization. Part I: Theoretical Aspects. *Computers & Chemical Engineering*, 27:157–166, 2003.
24. M. Loève. *Probability Theory II*. Springer, 4th edition, 1978.
25. Sergey Lototsky and Boris Rozovskii. *From Stochastic Calculus to Mathematical Finance*, chapter Stochastic Differential Equations: A Wiener Chaos Approach, pages 433–506. Springer, Berlin, Heidelberg, 2006.
26. W. Luo. *Wiener Chaos Expansion and Numerical Solutions of Stochastic Partial Differential Equations*. Caltech, 2006.
27. Harald Luschgy and Gilles Pagès. Functional quantization of Gaussian processes. *Journal of Functional Analysis*, 196:486–531, 2002.
28. Harald Luschgy and Gilles Pagès. Functional quantization of a class of Brownian diffusions: A constructive approach. *Stochastic Processes and their Applications*, 116:310–336, 2006.
29. L. Mathelin and K. A. Gallivan. A Compressed Sensing Approach for Partial Differential Equations with Random Data Input. *Communications in Computational Physics*, 12:919–954, 2012.
30. Ivan Nourdin and Giovanni Peccati. *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality*. Cambridge University Press, 2012.
31. David Nualart. *The Malliavin Calculus and Related Topics*. Springer, Berlin, Heidelberg, New York, second edition, 2006.
32. H. Ogura. Orthogonal functionals of the Poisson process. *IEEE Transactions on Information Theory*, 18(4):473–481, 1972.
33. B. Øksendal. *Stochastic Differential Equations*. Springer, Berlin, Heidelberg, 6th edition, 2007.
34. M.R. Osborne. On shooting methods for boundary value problems. *Journal of Mathematical Analysis and Applications*, 27:417–433, 1969.
35. Gilles Pagès and Huyèn Pham. A Quantization Algorithm for Multidimensional Stochastic Control Problems. Pre-print 697, Laboratoire de Probabilités & Modèles Aléatoires, Universités de Paris 6/7, 2001.
36. Gilles Pagès and Jacques Printems. Optimal quadratic quantization for numerics: the Gaussian case. *Monte Carlo Methods and Appl.*, 9(2):135–165, 2003.
37. A. Potschka. Handling Path Constraints in a Direct Multiple Shooting Method for Optimal Control Problems. Diploma thesis, Universität Heidelberg, 2006.
38. M.J.D. Powell. A fast algorithm for nonlinearly constrained optimization calculations. In G.A. Watson, editor, *Numerical Analysis, Dundee 1977*, volume 630 of *Lecture Notes in Mathematics*, Berlin, 1978. Springer.
39. Claudia Schillings and Volker Schulz. On the influence of robustness measures on shape optimization with stochastic uncertainties. Technical report, Universität Trier, 2012.
40. A. Segall and T. Kailath. Orthogonal functionals of independent-increment processes. *IEEE Transactions on Information Theory*, 22(3):287–298, 1976.
41. SA. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Soviet Math. Dokl.*, 4:240–243, 1963.

42. N. Wiener. The Homogeneous Chaos. *American Journal of Mathematics*, 60:897–936, 1938.
43. N. Wiener. *Nonlinear Problems in Random Theory*. Technology Press of the Massachusetts Institute of Technology and John Wiley & Sons Inc., New York, 1958.
44. R.B. Wilson. *A simplicial algorithm for concave programming*. PhD thesis, Harvard University, 1963.

## A Sparse and Adaptive Indices

**Table 4** List of (sp)arse and (ad)aptive indices used for the numerical examples in Section 6. The reference numbers coincide with those in Tables 1 and 3, the dagger symbol  $\dagger$  denotes the combinations used in the figures.

symbol	$k$	$p$	index $r$
sp <sup>1</sup>	5	3	$r = (3, 3, 2, 1, 1)$
sp <sup>2</sup>	10	2	$r = (2, 2, 2, 2, 2, 1, 1, 1, 1, 1)$
sp <sup>3</sup>	20	2	$r = (2, 2, 2, 2, 2, 1, \dots, 1)$
ad <sup>1</sup>	10	3	$r^1 = (1, \dots, 1)$ $r^2 = (2, 2, 2, 2, 2, 2, 0, 0, 0, 0)$ $r^3 = (3, 3, 3, 0, \dots, 0)$
ad <sup>2</sup>	20	2	$r^1 = (1, \dots, 1)$ $r^2 = (2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 0, \dots, 0)$
ad <sup>3</sup>	20	3	$r^1 = (1, \dots, 1)$ $r^2 = (2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 0, \dots, 0)$ $r^3 = (3, 3, 3, 2, 2, 2, 0, \dots, 0)$
ad <sup>4</sup>	40	2 <sup>†</sup>	$r^1 = (1, \dots, 1)$ $r^2 = (2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 0, \dots, 0)$
ad <sup>5</sup>	40	3	$r^1 = (1, \dots, 1)$ $r^2 = (2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 0, \dots, 0)$ $r^3 = (3, 3, 3, 2, 2, 2, 0, \dots, 0)$
ad <sup>6</sup>	100	2 <sup>†</sup>	$r^1 = (1, \dots, 1)$ $r^2 = (2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 0, \dots, 0)$