UC Berkeley
UC Berkeley Previously Published Works

Title

Stochastic Nonlinear Model Predictive Control with Efficient Sample Approximation of Chance Constraints

Permalink

https://escholarship.org/uc/item/0c5904nk

Authors

Streif, Stefan
Karl, Matthias
Mesbah, Ali

Publication Date

2014-10-16

Peer reviewed

Stochastic Nonlinear Model Predictive Control with Efficient Sample Approximation of Chance Constraints

Stefan Streif[a,b,d], Matthias Karl[b], Ali Mesbah[c]

[a]*Institute for Automation and Systems Engineering, Ilmenau University of Technology, 98684 Ilmenau, Germany*
[b]*Institute for Automation Engineering, Otto-von-Guericke Universität Magdeburg, 39106 Magdeburg, Germany*
[c]*Department of Chemical and Biomolecular Engineering, University of California, Berkeley, California 94720, USA*
[d]*Corresponding author (Stefan.Streif@TU-Ilmenau.de).*

Abstract

This paper presents a stochastic model predictive control approach for nonlinear systems subject to time-invariant probabilistic uncertainties in model parameters and initial conditions. The stochastic optimal control problem entails a cost function in terms of expected values and higher moments of the states, and chance constraints that ensure probabilistic constraint satisfaction. The generalized polynomial chaos framework is used to propagate the time-invariant stochastic uncertainties through the nonlinear system dynamics, and to efficiently sample from the probability densities of the states to approximate the satisfaction probability of the chance constraints. To increase computational efficiency by avoiding excessive sampling, a statistical analysis is proposed to systematically determine a-priori the least conservative constraint tightening required at a given sample size to guarantee a desired feasibility probability of the sample-approximated chance constraint optimization problem. In addition, a method is presented for sample-based approximation of the analytic gradients of the chance constraints, which increases the optimization efficiency significantly. The proposed stochastic nonlinear model predictive control approach is applicable to a broad class of nonlinear systems with the sufficient condition that each term is analytic with respect to the states, and separable with respect to the inputs, states and parameters. The closed-loop performance of the proposed approach is evaluated using the Williams-Otto reactor with seven states, and ten uncertain parameters and initial conditions. The results demonstrate the efficiency of the approach for real-time stochastic model predictive control and its capability to systematically account for probabilistic uncertainties in contrast to a nonlinear model predictive control approaches.

1. Introduction

Model predictive control (MPC) is widely used in the process industry owing to its ability to deal with multi-variable complex dynamics and to incorporate system constraints into the optimal control problem [1, 2]. However, parametric uncertainties and exogenous disturbances are ubiquitous in real-world systems, and the classical MPC framework is inherently limited to systematically account for uncertainties [3]. This consideration has led to the development of numerous robust MPC formulations that deal with uncertainties. The robust MPC approaches can be broadly categorized as deterministic and stochastic approaches based on the representation of uncertainties and the handling of constraints.

In deterministic robust MPC approaches (for a review see, e. g., [4]), uncertainties are often assumed to be bounded. The control law is determined such that the control objective is minimized with respect to worst-case uncertainty realizations, and/or such that the constraints are satisfied for all admissible values of uncertainties. Hence, robust MPC approaches discard statistical properties of uncertainties and are conservative [5, 6] if the worst-case uncertainty realizations have a small probability of occurrence.

In stochastic MPC (SMPC) approaches (e. g., see early work [7–9]) uncertainties are described by probability distributions (instead of bounded sets), which can often be readily obtained from state or parameter estimations. Such a stochastic approach to MPC not only alleviates the conservatism of worst-case control, but also enables tuning robustness against performance by allowing prespecified levels of risk during operation. The trade-off between control performance and robustness is achieved using chance (or probabilistic) constraints, which ensure the satisfaction of constraints with a desired probability level.

Stochastic MPC has recently become an active research area [8, 10–21]. These stochastic optimal control approaches are often limited to linear systems or restricted to certain types of uncertainty distributions (e. g. Gaussian uncertainties). The reason is twofold: first, the complexity of predicting the propagation of general uncertainty distributions through nonlinear system dynamics; second, chance constrained optimization problems are in general computationally intractable due to computation of multi-dimensional probability integrals. These integrals can only be evaluated exactly if special probability distributions are assumed (see, e. g., [22, 23]). However, such assumptions may not hold for the probability distributions of the states, especially in the presence of nonlinear system dynamics.

The restriction to special distributions is alleviated in sampling or scenario-based methods, or randomized algorithms (see, e. g. [5, 6, 8, 15, 24–30]). In scenario-based approaches, for instance, a suitable number of randomly extracted scenarios of uncertainty and disturbances are used to formulate an optimization problem that replaces the chance constrained control problem [5, 15, 26–30]. However, these approaches usually assume convexity of the optimization problem, which often implies linearity of the system dynamics and constraints. Furthermore, high-dimensional uncertainties often require a large number of samples for accurate prediction of the system dynamics, which may be prohibitively expensive for real-time control. Bounds on the required sample size in the scenario approach are given e. g. in [31, 32]. Thus, efficient sampling and evaluation of the randomized constraints is crucial for the efficiency of these approaches. [1]

In addition to the need for accurate and computationally efficient approximation of chance constraints, the evaluation of gradients of chance constraints is also critical for real-time control applications. As shown e. g. in [36], providing analytic forms of the gradients of the objective function and constraints can significantly improve the speed and accuracy of online optimization. However, the computation of gradients for nonlinear chance constraints is particularly challenging for general probability distributions, as analytic expressions of the gradients cannot be readily derived [23]. For sample-based approaches, it is known that a finite-difference approximation of the gradients can be very inaccurate even for a large number of samples [24, 37].

The contribution of this work is a stochastic nonlinear MPC (SNMPC) framework based on sample approximation of the chance constraints (Sec. 4) and their gradients (Sec. 5). In addition, a statistical analysis is presented to determine a-priori (i.e., before the real-time optimizations) the required constraint tightening and the number of samples that guarantee a desired feasibility probability for a prespecified robustness (or risk) level (Sec. 4). The presented SNMPC approach is applicable to a broad class of nonlinear systems subject to time-invariant uncertainties in model parameters and initial conditions. The system dynamics and constraints are required to be analytic with respect to the states and separable with respect to the inputs, states and parameters. The generalized polynomial chaos (PC) framework is used to obtain a computationally efficient surrogate for uncertainty propagation through the nonlinear system dynamics in order to generate a large number of Monte-Carlo (MC) samples. [2] In the PC framework, spectral expansions in terms of orthogonal polynomials are used to present stochastic quantities [42–45], which allows sampling in a computationally efficient manner (Sec. 3). The performance of the proposed SNMPC approach is demonstrated for the Williams-Otto reactor using extensive simulation studies (Sec. 6).

Notation

Bold symbols (e. g. x, ξ) denote vectors. Subscripts at vectors are used for indexing vector elements. n_x, n_u, ... denote dimension of the vectors indicated by the subscripts. N (e. g. N_x, N_S) denote the number of constraints or samples. Sets are denoted by calligraphic letters (e. g. $\mathcal{X}$, $\mathcal{U}$). Tilde (e. g. $\widetilde{x}$, $\widetilde{p}$) denotes coefficients in the polynomial chaos expansion of the corresponding variable, and $\widetilde{P}$ denotes the number of terms/coefficients in the polynomial chaos expansion of order P. Probability distributions are denoted by μ. Expectation of a random variable is denoted by $\mathbb{E}[\cdot]$, variance by $\mathbb{V}[\cdot]$, higher order moments by $\mathbb{E}^m[\cdot]$, and probability by $\mathbb{P}[\cdot]$. Superscripts $[i]$ (e. g. $\xi^{[i]}$) indicate independent and identically distributed samples. Superscript $*$ denotes the optimal solution of an optimization problem.

---

[1] There exists a large literature on alternatives to sampling-based methods [20, 21, 33–35]. See also [23] for a recent review.

[2] The use of the PC framework for stochastic MPC and optimal control has also been investigated in [35, 38–41].

2. Stochastic Model Predictive Control Problem

Consider the continuous-time, nonlinear system

$$\dot{x}(t) = f(x(t), u(t), p), \quad x(0) = x_0, \tag{1}$$

where t, $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$ and $p \in \mathbb{R}^{n_p}$ denote time, the states, the inputs and the time-invariant parameters, respectively. x_0 denotes the initial states. The function $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \rightarrow \mathbb{R}^{n_x}$ represents the nonlinear system dynamics. To be able to efficiently employ the framework presented in the subsequent sections, it is assumed that f can be transformed into a polynomial-in-the-states representation [46]. A sufficient condition for this assumption to hold is that f is analytic with respect to the states, and separable with respect to the inputs, states and parameters.

2.1. Uncertainties

The system dynamics are subject to the following uncertainties. The time-invariant parameters p_i, $i = 1, \dots, n_p$ are assumed to be distributed according to the continuous probability distributions μ_{p_i} (denoted by $p_i \sim \mu_{p_i}$). Additionally, uncertain estimates of the states $x_i(t_k)$, $i = 1, \dots, n_x$ described by continuous probability distributions $\mu_{x_i(t_k)}$, can be used to recursively initialize (1). Such uncertainty descriptions for the initial conditions are often available from state estimation, for example from Kalman filters. Note that this formulation also allows considering exact state estimates by choosing $\mu_{x_i(t_k)}$ to be a Dirac distribution. For technical reasons, it is assumed that the parameters and initial conditions are uncorrelated and have finite variances (e. g., $\mathbb{V}[p_i] < \infty$ and $\mathbb{V}[x_i(t_k)] < \infty$).

2.2. Cost Function and Constraints

This work considers the stochastic optimal control of the system (1) on the finite-time horizon $[t_k, t_f]$ [3], while constraints on the inputs and states should be satisfied in the presence of uncertainties. The cost function of the stochastic nonlinear model predictive control approach is assumed to be deterministic and defined by

$$J(x(\cdot), u(\cdot)) := \int_{t_k}^{t_f} F\left(u(t), \mathbb{E}[x(t)], \mathbb{E}^2[x(t)], \dots\right) dt + E\left(\mathbb{E}[x(t_f)], \mathbb{E}^2[x(t_f)], \dots\right). \tag{2}$$

Here F and E denote the running and terminal cost functions, respectively, both of which can be functions of the moments of the states $\mathbb{E}^m[\cdot]$. Such a cost function enables shaping state distributions or, in a simpler case, minimizing the variance of state distributions (e.g., see [35, 47]).

In the following, we assume that (1) and (2) are time-discretized to integrate the nonlinear system dynamics and to impose constraints on the states and inputs at discrete time-points as described next.

State constraints[4] are imposed at N_x different time points $t_{x,i} \in [t_k, t_f]$, $i = 1, \dots, N_x$

$$X \in \mathcal{X} := \{g_i(X) \leq 0, \quad i = 1, \dots, N_g\} \subset \mathbb{R}^{n_x N_x}, \tag{3}$$

where $X := [x(t_{x,1})^\top, \dots, x(t_{x,N_x})^\top]^\top$, and N_g is the number of constraints. To efficiently employ the proposed control approach, functions $g_i(X)$ are assumed to satisfy the same conditions as f (i.e., being analytic with respect to the states and separable with respect to the inputs, states and parameters).

In addition to state constraints, inputs are constrained by a compact set $\mathcal{U}$

$$u(t) \in \mathcal{U} \subset \mathbb{R}^{n_u}, \quad \forall t \in [t_k, t_f]. \tag{4}$$

For notational simplicity, output constraints are not explicitly considered here, as output constraints can often be represented in terms of (3). Note that algebraic equations can be straightforwardly incorporated into the considered stochastic optimal control framework [38].

---

[3]Here $t_f = t_k + T$ (with T being the prediction horizon) for receding horizon control, and fixed t_f for shrinking horizon control.
[4]Output constraints can be considered similarly.

Under the uncertainties in the system parameters and initial conditions, the solution trajectories of system (1) may violate the constraints (3). In this work, inputs $u(t)$ are designed to satisfy (4) such that constraints (3) are fulfilled in a probabilistic manner in the presence of uncertainties. This is formalized by chance constraints [10]

$$\mathbb{P}\left[g(X) \leq 0\right] \geq \beta \tag{5}$$

where $\beta \in (0, 1]$ is a user-specified probability chosen according to process requirements: $\beta = 1$ corresponds to hard constraints that should hold at all times for all uncertainty realizations (i.e., the risk-free case); $\beta < 1$ allows for constraint violation with probability $1 - \beta$ in order to trade-off control performance with robustness.

When $N_g > 1$, (5) entails *joint chance constraints*, as all constraints $g_1(X) \leq 0, \ldots, g_{N_g}(X) \leq 0$ should be jointly satisfied with the probability level β. If $N_g = 1$ or constraints are defined independently using different β_i for each constraint $g_i(X)$, (5) will be referred to as *individual chance constraints* (e. g., see [10]).

2.3. Problem Formulation

This paper considers the nonlinear system (1) with parametric uncertainties, and in which the initial states at sampling time t_k are uncertain, e. g. due to uncertain state estimates. Let t_k denote the sampling time instances at which states $x(t_k)$ become available. Furthermore, denote by the vector $\pi \in \mathbb{R}^{n_\pi}$ and function $\overline{u}$ a parametrization of the continuous-time input such that $u(t) = \overline{u}(t, \pi)$, $t \in [t_k, t_f]$.[5]

This work considers the following main problem.

Finite-horizon stochastic nonlinear MPC with joint chance constraints: *At each sampling time t_k the following stochastic optimal control problem is solved*

$$\min_{\pi} \; J(\overline{x}(\cdot), \overline{u}(\cdot)) \tag{6a}$$

$$\textit{subject to: } \dot{\overline{x}}(t) = f(\overline{x}(t), \overline{u}(t, \pi), p), \;\; \forall t \in [t_k, t_f] \tag{6b}$$

$$\mathbb{P}\left[g(X) \leq 0\right] \geq \beta, \tag{6c}$$

$$\overline{u}(t, \pi) \in \mathcal{U}, \qquad\qquad \forall t \in [t_k, t_f] \tag{6d}$$

$$\overline{x}_i(t_k) \sim \mu_{x_i(t_k)}, \qquad\quad i = 1, \ldots, n_x \tag{6e}$$

$$p_i \sim \mu_{p_i}, \qquad\qquad\quad i = 1, \ldots, n_p \tag{6f}$$

where π denote the decision variables, $\overline{x}(t)$ denotes the states predicted by the nonlinear system model, and $\mu_{x_i(t_k)}$ denotes the probability distributions of the states at time t_k.

To facilitate closed-loop control, the stochastic optimal control problem (6) is often implemented in a receding-horizon mode or shrinking-horizon mode (cf. footnote 3). The closed-loop control that is applied to the system (1) is defined by the optimal solution π^* of (6) at the sampling instants: $u(t) = \overline{u}(t, \pi^*)$, $t \in [t_k, t_{k+1}]$.

This work presents a framework to efficiently solve (6). In particular, the following problems are addressed.

Problem 1: Propagation of the time-invariant probabilistic uncertainties (6e) **and** (6f) **through the nonlinear system dynamics** (6b)**.**

The problem is addressed using the polynomial chaos approach as presented in Sec. 3. This allows to efficiently sample from the probability distributions of the states and cost functions (Sec. 4) to address the next problem.

Problem 2: Efficient evaluation of the chance constraints (6c)**.**

Accurate prediction of uncertain system dynamics typically requires a large number of samples, which can still be prohibitive even in the polynomial chaos approach. Moreover, sample-based approximations of (6) may yield solutions that are infeasible for the original problem (6) [48]. To reduce the risk of infeasibility due to the statistical error made due to the finite sample size while avoiding excessive sampling, the constraints can be tightened to make the entire problem more robust. This is addressed in the following problem.

Problem 3: Tightening the chance constraints (6c) **by $\beta_{\mathrm{cor}} > \beta$ depending on the sample-size to guarantee a desired feasibility probability.**

---

[5]For a piecewise-constant control input parameterization $u(t) = \pi_i$, $t \in [t_{u,i}, t_{u,i+1}]$, $i = 1, \ldots, N_u$, partition the time horizon $[t_k, t_f]$ into N_u subintervals $[t_{u,i}, t_{u,i+1}]$ with $t_k = t_{u,1} < \ldots < t_{u,N_u} < t_{u,N_u+1} = t_f$.

A solution to this problem is presented in Sec. 4. The last problem addresses the efficiency of solving the stochastic optimal control problem (6) using gradient-based optimization algorithms.

Problem 4: Determining the gradients of the chance constraints (6c).

A solution to the latter problem is presented in Sec. 5 and the overall framework is demonstrated in Sec. 6.

3. Uncertainty Propagation for Nonlinear Systems Using Polynomial Chaos

This work uses the polynomial chaos framework to solve Problem 1. In the PC framework, spectral expansions in terms of orthogonal polynomials are used to represent stochastic variables and parameters [42–45]. This allows deriving surrogate models, based on which the moments of the system states of the original system can be readily obtained. As shown in Sec. 4, PC also allows for the sampling-based approximation and evaluation of chance constraints in a computationally efficient manner.

3.1. Polynomial Chaos Expansion

In the following, we assume that all uncertain parameters and uncertain initial conditions are functions of the standard random variables $\boldsymbol{\xi} \in \mathbb{R}^{n_\xi}$, which is denoted by $\boldsymbol{x}(t, \boldsymbol{\xi})$, $\boldsymbol{p}(\boldsymbol{\xi})$, etc. The random variables ξ_j, $j = 1, \ldots, n_\xi$ are assumed to be independent with known probability distribution functions (PDFs) μ_{ξ_j}, such that $\xi_j \in L^2(\Omega, \mathcal{F}, \mu)$. Here, $L^2(\Omega, \mathcal{F}, \mu)$ is the Hilbert space of all random variables ξ_j with finite variance $\mathbb{V}\left[\xi_j\right] < \infty$. The triple $(\Omega, \mathcal{F}, \mu)$ denotes the probability space on the basis of the sample space Ω, σ-algebra $\mathcal{F}$ of subsets of Ω, and probability measure μ on Ω.

Le $v(t, \boldsymbol{\xi})$ represent a state variable $x_i(t, \boldsymbol{\xi})$ or any other (possibly) nonlinear function such as $g_i(\boldsymbol{x}(t, \boldsymbol{\xi}))$ in the chance constraints (5). To explicitly derive the dependence of v on the random variables $\boldsymbol{\xi}$, the following polynomial chaos expansion can be used [42, 44, 49]

$$v(t, \boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha}_i \in \mathcal{I}_\infty} \widetilde{v}_{\boldsymbol{\alpha}_i}(t) \Psi_{\boldsymbol{\alpha}_i}(\boldsymbol{\xi}). \tag{7}$$

The variables $\widetilde{v}_{\boldsymbol{\alpha}_i}$ denote the (deterministic) coefficients of the PC expansion (PCE), and $\Psi_{\boldsymbol{\alpha}_i}(\boldsymbol{\xi})$ denote multivariate polynomials in the random variables $\boldsymbol{\xi}$ of total degree $\sum_{j=1}^{n_\xi}(\boldsymbol{\alpha}_i)_j$. The multivariate polynomials can be written as products of univariate polynomials: $\Psi_{\boldsymbol{\alpha}_i}(\boldsymbol{\xi}) := \prod_{j=1}^{n_\xi} \Phi_{\xi_j}^{(\alpha_i)_j}$. The polynomial $\Phi_{\xi_j}^m$ of the random variable ξ_j is of degree m, where $\{\Phi_{\xi_j}^m\}_{m=0}^P$, $j = 1, \ldots, n_\xi$ is an orthogonal basis[6] with respect to the corresponding probability measures $\mu(\xi_j)$.

In (7), $\boldsymbol{\alpha}_i$ is the ith multi-index vector from the set

$$\mathcal{I}_P := \left\{ \boldsymbol{\alpha} : \boldsymbol{\alpha} \in \mathbb{N}_{\geq 0}^{n_\xi}, \sum_{j=1}^{n_\xi} \alpha_j \leq P \right\}$$

with $P = \infty$. For practical reasons, the infinite (weighted) sum of polynomials in Eq. (7) is truncated after $\widetilde{P} := \frac{(n_\xi + P)!}{n_\xi! P!}$ terms, where P is called the order of the PC expansion. The truncation can be written in a compact form as

$$v(t, \boldsymbol{\xi}) \approx \widetilde{\boldsymbol{v}}(t)^\top \boldsymbol{\Psi}(\boldsymbol{\xi}), \tag{8}$$

with

$$\widetilde{\boldsymbol{v}}(t) := \left[\widetilde{v}_{\boldsymbol{\alpha}_1}(t), \widetilde{v}_{\boldsymbol{\alpha}_2}(t), \ldots, \widetilde{v}_{\boldsymbol{\alpha}_{\widetilde{P}}}(t) \right]^\top \in \mathbb{R}^{\widetilde{P}} \tag{9}$$

being the vector of coefficients for the PC expansion of variable v, and

$$\boldsymbol{\Psi}(\boldsymbol{\xi}) := \left[\Psi_{\boldsymbol{\alpha}_1}(\boldsymbol{\xi}), \Psi_{\boldsymbol{\alpha}_2}(\boldsymbol{\xi}), \ldots, \Psi_{\boldsymbol{\alpha}_{\widetilde{P}}}(\boldsymbol{\xi}) \right]^\top \in \mathbb{R}^{\widetilde{P}} \tag{10}$$

being the vector of the multivariate polynomials.

The next step is to determine the values of the PC expansion coefficients (9), which will be needed to approximate the probability distributions for the evaluation of the chance constraints (see Sec. 4). Two different approaches to determine the PC expansion coefficients are explained in the next two subsections.

---

[6]Such orthogonal bases are readily available for well-known standard distributions such as Normal, Uniform or Beta distributions [44], or, in general, can be constructed for any distribution using moments [50] or Gram-Schmidt orthogonalization [51].

3.2. Determination of the PC Expansion Coefficients Using Collocation

The PC expansion coefficients can be determined using the so-called probabilistic collocation methods (e.g., see [39, 52–54] and references within). In the collocation methods, N_{MC} samples are drawn from the known distributions of uncertainties and, subsequently, are used to solve the nonlinear process model (1). The PC expansion coefficients can then be obtained in a least squares sense through minimizing the residuals between the PC expansion and the nonlinear model predictions $v(t, \xi^{[j]})$. Note that an explicit analytic solution to the resulting unconstrained optimization problem $\min_{\widetilde{v}(t)} \sum_{j=1}^{N_{MC}} \left(\widetilde{v}(t)^\top \Psi(\xi^{[j]}) - v(t, \xi^{[j]}) \right)^2$ is readily available.

The complexity of process dynamics in terms of nonlinearities may require a large number of samples N_{MC} to obtain adequate estimations of the PC expansion coefficients. This may render real-time control applications computationally prohibitive. Next, an alternative approach is presented for determining the PC expansion coefficients that relies on the solution of an extended set of ordinary differential equations.

3.3. Determination of the PC Expansion Coefficients Using Galerkin Projection

In the following it is assumed that the considered system is polynomial in the states, i. e. the dynamics for the ith state

$$\dot{x}_i(t, \xi) = f_i(x(t, \xi), u(t), p(\xi)), \quad \forall i = 1, \ldots, n_x \tag{11}$$

is of the form $c \prod_{l=1}^{n_x} x_l^{\gamma_l}(t, \xi) \sigma_u(u(t)) \sigma_p(p(\xi))$, where c is a constant and γ_l is the degree of variable x_l. σ_p and σ_u are nonlinear functions of the parameters and inputs, respectively. Note that such a representation can always be obtained exactly by state-lifting or immersion [46] under the conditions given in Sec. 2. Such a polynomial model structure enables explicit derivation of differential equations for the PC coefficients $\widetilde{x}_i$ using Galerkin projection [43]. The Galerkin projection requires evaluation of multi-dimensional integrals, which can be solved exactly and efficiently offline for polynomial systems in the form (11). For generality, the (uncertain) parameters need not appear polynomially in (11). If σ_p is not polynomial, a PC approximation of σ_p can be determined using the collocation approach described in the previous section 3.2.

3.3.1. Galerkin Projection of the System Dynamics

To obtain the PC coefficients for the states $\widetilde{x}_i(t)$, an extended system of ordinary differential equations is derived using Galerkin projection [42, 44]. This system is obtained by formal derivative of the PC expansion (8) for the state variables (yielding $\dot{\widetilde{x}}_i(t)^\top \Psi(\xi)$) and by inserting the PC expansion of the state variables (8) into the system dynamics (11). Subsequently, the projection is performed by computing the inner products $\int f_i(\widetilde{x}, u(t), \widetilde{p}) \Psi_{\alpha_i}(\xi) \mu_1(d\xi_1) \cdots \mu_{n_\xi}(d\xi_{n_\xi})$ of the resulting equation and the different polynomials in (10). By employing orthogonality, this results in the set of ordinary differential equations (ODEs) describing the dynamics of the coefficients

$$\dot{\widetilde{x}}_i(t) = \widetilde{f}_i(\widetilde{x}_1(t), \ldots, \widetilde{x}_{n_x}(t), u(t), \widetilde{p}_1, \ldots, \widetilde{p}_{n_p}), \forall i = 1, \ldots, n_x, \tag{12}$$

where $\widetilde{x}_i$ and $\widetilde{p}_i$ are the vectors of coefficients of the PC expansions (cf. (9)) of the states and parameters, respectively. The system (12) has extended state space dimension $\widetilde{P} n_x$ and describes the dynamics of the PC expansion coefficients. Thus, by projection onto the orthogonal polynomials, the explicit dependencies on the random variables are eliminated. The solution to this set of ODEs can then be used for efficient sample evaluations (cf. Sec. 3.4).

To compute the inner products, multi-dimensional integration is required. Even though this is in general cumbersome, it is particularly easy for polynomial systems (11). In this case, the integral can be efficiently and exactly computed using Gauss Quadrature [55]. In addition, due to (power) orthogonality [55] most projection integrals ($\sim 97\%$) are zero such that the computational burden can be reduced significantly. For more details and the structure of (12), see [56].

3.3.2. Galerkin Projection of the Initial Conditions

Once (12) has been determined as described in Secs. 3.2 and 3.3, the initial conditions $\widetilde{x}_i(t_k)$ are needed for numerical solution of the set of ODEs. Since the initial conditions $x(t_k)$ are assumed to be functions of the uncertainties

$\underline{\xi}$, they can be obtained by projection of the corresponding PC expansion (7) onto the different orthogonal polynomials $\Psi_{\alpha_i}(\xi)$

$$\widetilde{x}_{i,j}(t_k) = \frac{\int x(t_k, \xi)\Psi_{\alpha_i}(\xi)\mu_1(d\xi_1)\cdots\mu_{n_\xi}(d\xi_{n_\xi})}{\int (\Psi_{\alpha_i}(\xi))^2 \mu_1(d\xi_1)\cdots\mu_{n_\xi}(d\xi_{n_\xi})}, \forall i = 1,\ldots, n_x, \forall j = 1,\ldots,\widetilde{P}. \tag{13}$$

3.4. Computation of the Moments and Efficient Sampling to Approximate Probability Distributions

Once the PC expansion coefficients (9) are determined, the probability density of $v(t,\xi)$ can be approximated using sampling. With that, the approximation of the probability distributions of stochastic variables v or the evaluation of chance constraints can be done efficiently as shown in Sec. 4.

Let $\xi^{[j]}$ be samples drawn from the multivariate distributions of ξ. Then $v(t,\xi_i^{[j]})$ is obtained from (8) by evaluating the orthogonal polynomials (10) for the different samples $\xi^{[j]}$, and by vector multiplications with the vector of the PC expansion coefficients obtained from the solution of (12).

Besides such a sampling-based approach, the probability distributions can be approximated based on moments (see discussion and references in [56]). The moments can be determined directly from polynomial chaos expansions without further approximations (see [47, 56]), which is computationally cheap for low-order moments. The moments are also required to compute the cost function (2) in (6).

4. Sample Approximation of Chance Constraints with Guaranteed Feasibility Probability

In this section, Problems 2 and 3 are addressed. In particular, the satisfaction probability of the joint chance constraint (5) is approximated using samples generated with the PC framework. Even though a large number of samples can be generated in a computationally efficient manner using the PC expansion (see Sec. 3), the evaluation of the nonlinear functions $g_i(\widetilde{X}, \xi)$, $i = 1,\ldots, N_g$ and their gradients (see Sec. 5) may still be prohibitive for a large sample size or large N_g. To increase the computational efficiency, it is therefore desired to evaluate as few samples as possible. However, fewer samples increase the width of the confidence interval (i. e. reduce the quality) of the estimate of the satisfaction probability. A low confidence bears the risk that a feasible solution to the sample-approximated chance constrained optimization problem is infeasible for the original problem (e. g. for a different or larger set of samples) [48]. To increase the *confidence level of feasibility*, which we call *feasibility probability*, one can make the controller more robust by tightening the chance constraints (i. e., using a $\beta_{\text{cor}} > \beta$ in Eq. (5)). This section proposes a statistical analysis to determine a-priori the the constraint tightening β_{cor} for which a desired feasibility probability $1 - \alpha$ ($\alpha \in (0,1)$) and a desired robustness level (satisfaction probability β) can be guaranteed.

Note that this section considers the PC expanded system (12) rather than the original nonlinear system (6b).

4.1. Satisfaction Probability

Suppose that the inequalities in the chance constraints (5) are expressed using the PC expansion

$$\mathbb{P}\left[g\left(\widetilde{X}, \xi\right) \leq 0\right] \geq \beta, \tag{14}$$

with β as in (5), and

$$\widetilde{X} := \left[\widetilde{x}_1(t_{x,1})^\top, \ldots, \widetilde{x}_{n_x}(t_{x,1})^\top, \ldots, \widetilde{x}_1(t_{x,N_x})^\top, \ldots, \widetilde{x}_{n_x}(t_{x,N_x})^\top\right]^\top$$

are given from the simulation of the PC expanded system (12). The probability of satisfaction in (14) is given by

$$\mathbb{P}\left[g\left(\widetilde{X}, \xi\right) \leq 0\right] := \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} I_{\mathcal{G}}(\xi)\,\mu_1(d\xi_1)\mu_2(d\xi_2)\cdots\mu_{n_\xi}(d\xi_{n_\xi}). \tag{15}$$

$I_{\mathcal{G}}$ is the indicator function

$$I_{\mathcal{G}}(\xi) := \begin{cases} 1 & \text{if } \xi \in \mathcal{G} \\ 0 & \text{otherwise,} \end{cases}$$

where $\mathcal{G} \in \mathbb{R}^{n_\xi}$ denotes the set where all constraints are satisfied

$$\mathcal{G} := \left\{ \boldsymbol{\xi} \in \mathbb{R}^{n_\xi} \mid g_i\left(\widetilde{X}, \boldsymbol{\xi}\right) \leq 0, \forall i = 1, \ldots, N_g \right\}.$$

Evaluation of the multidimensional integral in (15) is in general very difficult due to the non-convexity of the set $\mathcal{G}$ and the weighting with respect to the probability measures $\mu(\xi_i)$, $i = 1, \ldots, n_\xi$.

In this work, sampling is used to approximate the probability of satisfaction (15). For this purpose, n_ξ-dimensional independent and identically distributed samples $\boldsymbol{\xi}^{[1]}, \ldots, \boldsymbol{\xi}^{[N_S]}$ are drawn from the distributions $\mu_i(\xi_i)$, $i = 1, \ldots, n_\xi$. The sample-average approximation of $\mathbb{P}\left[\boldsymbol{g}\left(\widetilde{X}, \boldsymbol{\xi}\right) \leq \boldsymbol{0}\right]$ is given by

$$\mathbb{P}\left[\boldsymbol{g}\left(\widetilde{X}, \boldsymbol{\xi}\right) \leq \boldsymbol{0}\right] \approx \frac{1}{N_S} \sum_{j=1}^{N_S} I_\mathcal{G}\left(\boldsymbol{\xi}^{[j]}\right). \tag{16}$$

4.2. Required Sample Size and Constraint Tightening for a Guaranteed Feasibility Probability

The main results of this section are the following theorems that allow determining systematically the constraint tightening and sample size for which desired confidence level in the approximation (16) of (15) is guaranteed.

Theorem 1 (Constraint Tightening and Confidence in the Sample Approximation of the Chance Constraints): *If $\beta_{cor} > \beta$ is chosen such that*

$$1 - betainv\left(1 - \frac{\alpha}{2}, N_S + 1 - \lfloor \beta_{cor} N_S \rfloor, \lfloor \beta_{cor} N_S \rfloor\right) \geq \beta, \tag{17}$$

then the sample approximation (16) *of the tightened chance constraints* $\mathbb{P}\left[\boldsymbol{g}\left(\widetilde{X}, \boldsymbol{\xi}\right) \leq \boldsymbol{0}\right] \geq \beta_{cor}$ *using N_S samples guarantees a confidence level of $1 - \alpha$ of the original chance constraint* $\mathbb{P}\left[\boldsymbol{g}\left(\widetilde{X}, \boldsymbol{\xi}\right) \leq \boldsymbol{0}\right] \geq \beta$ *((14) resp. (6c)). Furthermore, β_{cor} is the least conservative constraint tightening that can be chosen.*
The operator $\lfloor \cdot \rfloor$ denotes rounding towards $-\infty$, and betainv denotes the inverse of the cumulative Beta-distribution.
Proof: *The proof uses standard results in statistics and is presented in Appendix A.* □

The theorem tightens the chance constraints β_{cor} to compensate for the statistical error made due to the finite sample size N_S. To this end, it employs the lower bounds of the confidence interval obtained from a statistical analysis (details see the proof). Note that the analysis is independent of the specifics of the optimization problems, in particular of the chance constraints and the system dynamics. Note also that the analysis neither depends on the number of decision variables nor requires convexity assumptions as in [5, 26–28, 57]. Furthermore, bounds on the sample complexity have been presented in [31, 32] in a different context.

Due to the independence of the analysis on the specifics of the optimization problems and constraints, Theorem 1 can be applied offline (and needs to be done only once) to systematically satisfy prespecified probabilistic requirements. The implicit relationships of (17) are plotted in Fig. 1 for selected values of α, β, and N_S as a reference, and can be derived for other values similarly. It is noteworthy that more samples does not necessarily mean a tighter confidence interval [58], which can be also seen at the lack of monotonicity of the curves in the inset of Fig. 1.

Theorem 1 can now be used to guarantee a certain confidence level of the feasibility of a solution to (6).

Theorem 2 (Feasibility Probability): *Consider the chance constrained stochastic optimal control problem* (6) *and let π' be a feasible solution to the sample-approximated form of problem* (6) *for a given β_{cor} according to Theorem 1. Then π' is a feasible point of* (6) *with probability not less than $1 - \alpha$.*
Proof: *See Appendix B.* □

Theorem 2 guarantees that a solution found by sample approximation is also a solution to the original problem (6) (i. e. for an infinite number of samples) with the specified confidence level $1 - \alpha$.

5. Sample Approximation of the Gradients for Efficient Optimization

This section proposes a solution to Problem 4, that is a sample-based approximation of the analytic expressions of the gradients of the chance constraints. The proposed approach avoids approximations of the gradients by finite
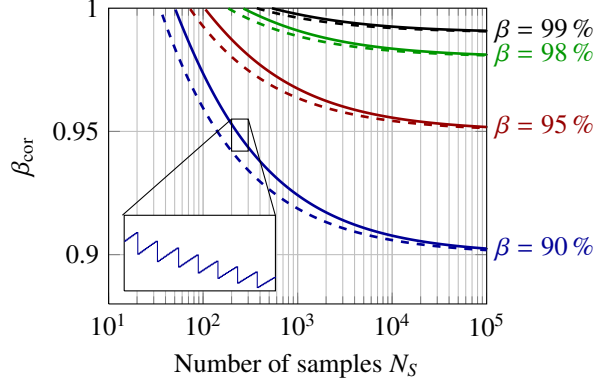
Figure 1: Sample size N_S and corrected satisfaction probability β_{cor} (i.e. chance constraint tightening) for different required satisfaction probabilities β for a feasibility probability of $1 - \alpha = 0.99$ (solid lines) and $1 - \alpha = 0.95$ (dashed lines). The inset shows the non-monotonicity of the curves.

difference methods, which typically slows down gradient-based optimization algorithms (see e. g. [36]). Furthermore, finite-differencing of the chance constraints can leads to poor estimates due to the discrete nature of sample-approximations as explained in Fig. 2a and [24, 37]. [7]

The gradients of the joint chance constraints with respect to the input parametrization $\boldsymbol{\pi}$ are formally given by

$$\frac{d\mathbb{P}\left[\boldsymbol{g}\left(\widetilde{\boldsymbol{X}},\boldsymbol{\xi}\right) \leq \boldsymbol{0}\right]}{d\boldsymbol{\pi}} = \frac{\partial\mathbb{P}\left[\boldsymbol{g}\left(\widetilde{\boldsymbol{X}},\boldsymbol{\xi}\right) \leq \boldsymbol{0}\right]}{\partial\widetilde{\boldsymbol{X}}} \frac{\partial\widetilde{\boldsymbol{X}}}{\partial\boldsymbol{\pi}}. \tag{18}$$

Note that the first-order sensitivities $\frac{\partial\widetilde{\boldsymbol{X}}}{\partial\boldsymbol{\pi}}$ are obtained from the solution of the sensitivity equations, which are almost always needed to speed up gradient-based optimization algorithms.

The main result of this section is the following proposition, namely sample-based approximation of the analytic gradients of (15) with respect to the input parametrization variables $\boldsymbol{\pi}$. Note that the dependence of the states $\widetilde{\boldsymbol{X}}$ on $\boldsymbol{\pi}$ is not explicitly written to shorten the notation. The derivation is illustrated in Fig. 2b and the constructive proof is given with all technical details in the Appendix C.

Proposition 1 (Gradients of the Sample-Approximated Chance Constraints): *Suppose the following to be given:*

1. *first-order sensitivities $\frac{\partial\widetilde{\boldsymbol{X}}}{\partial\boldsymbol{\pi}}$;*
2. *$\overline{\boldsymbol{\xi}}^{[j]}$, $j = 1, \ldots, N_S$ independent and identically distributed samples from the vector $\overline{\boldsymbol{\xi}} := [\xi_2, \ldots, \xi_{n_\xi}]$;*
3. *for each sample $\overline{\boldsymbol{\xi}}^{[j]}$ the vector $\boldsymbol{r}^{[j]} := [r_0^{[j]}, r_1^{[j]}, \ldots, r_{n_r^{[j]}+1}^{[j]}]$ with unique elements and sorted in ascending order; the vector is defined by the solutions of $\boldsymbol{g}(\widetilde{\boldsymbol{X}}, \xi_1, \overline{\boldsymbol{\xi}}^{[j]}) = 0$ with respect to ξ_1; $r_0^{[j]}$ ($r_{n_r^{[j]}+1}^{[j]}$) is the minimum (resp. maximum) of the support of ξ_1.*

Then a sample-based approximation of $\frac{\partial\mathbb{P}\left[\boldsymbol{g}\left(\widetilde{\boldsymbol{X}},\boldsymbol{\xi}\right)\leq\boldsymbol{0}\right]}{\partial\widetilde{\boldsymbol{X}}}$ in (18) is given by

$$\frac{\partial\mathbb{P}\left[\boldsymbol{g}\left(\widetilde{\boldsymbol{X}},\boldsymbol{\xi}\right) \leq \boldsymbol{0}\right]}{\partial\widetilde{\boldsymbol{X}}} \approx \frac{1}{N_S}\sum_{j=1}^{N_S}\left[\sum_{i=0}^{n_r^{[j]}} I_{\mathcal{G}}\left(\frac{r_i^{[j]} + r_{i+1}^{[j]}}{2}, \overline{\boldsymbol{\xi}}^{[j]}\right)\left(\mu_1\left(r_{i+1}^{[j]}\right)\frac{\partial r_{i+1}^{[j]}}{\partial\widetilde{\boldsymbol{X}}} - \mu_1\left(r_i^{[j]}\right)\frac{\partial r_i^{[j]}}{\partial\widetilde{\boldsymbol{X}}}\right)\right], \tag{19}$$

in which

$$\frac{\partial r_i^{[j]}}{\partial\widetilde{\boldsymbol{X}}} = \begin{cases} 0, & i \in \{0, n_r^{[j]} + 1\} \\ -\frac{\partial g\left(\widetilde{\boldsymbol{X}}, r_i^{[j]}, \overline{\boldsymbol{\xi}}^{[j]}\right)}{\partial\widetilde{\boldsymbol{X}}}\left(\frac{\partial g\left(\widetilde{\boldsymbol{X}}, r_i^{[j]}, \overline{\boldsymbol{\xi}}^{[j]}\right)}{\partial\xi_1}\right)^{-1}, & i \in \{1, \ldots, n_r^{[j]}\}. \end{cases} \tag{20}$$

---

[7]Note that more sophisticated sampling methods such as importance sampling could be used at the cost of much higher computational demands.
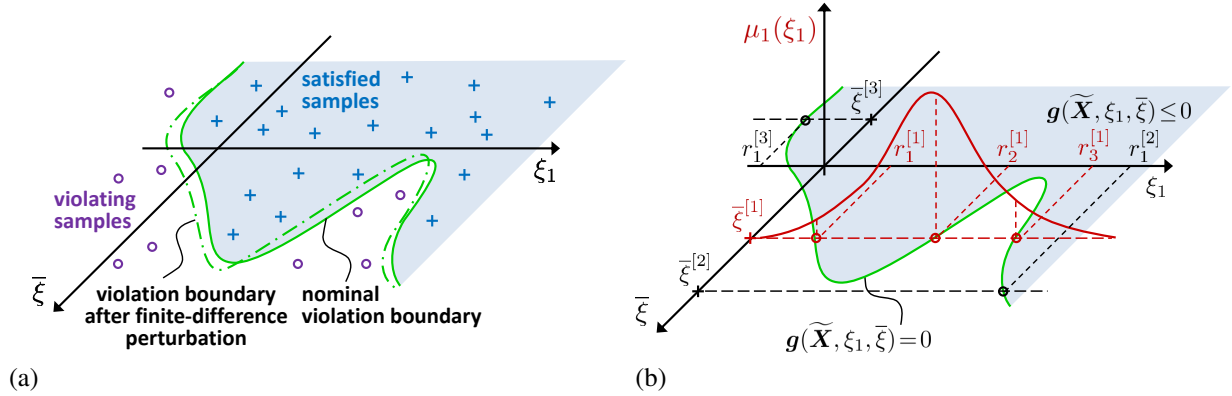
Figure 2: (a) Samples satisfying $g(\cdot) \leq 0$ are shown by blue **+**, and violating samples by purple ∘. The satisfaction probability is given by ratio "number of satisfying samples"/"total number of samples". This ratio changes discontinuously due to finite-difference perturbations $\widetilde{X} + \Delta\widetilde{X}$ because a discrete number of samples change their validity. Note that the ratio may not even change at all if the validity of none of the samples is affected. (b) The samples $\bar{\xi}^{[1]}, \bar{\xi}^{[2]}, \bar{\xi}^{[3]}$ (symbol **+**) are drawn at random from the distributions $\mu_2(\xi_2)$, where $\bar{\xi} = \xi_2$. The equation $g(\widetilde{X}, \xi_1, \bar{\xi}^{[j]}) = 0$ defines the violation boundary (shown in green) and it is solved for ξ_1. For $\bar{\xi}^{[1]}$, this yields three solutions $r^{[1]} = [r_1^{[1]}, r_2^{[1]}, r_3^{[1]}]$ (shown in red). A different number of (real-valued) solutions is obtained for the other samples as shown for the red and black dashed lines. Piecewise integration of the probability measure $\mu_1(\xi_1)$ is performed with integration limits given by the elements of $r^{[1]}$. Note that the integral between the limits $r_1^{[1]}$ and $r_2^{[1]}$ contributes to the approximation of the gradient only if the constraint $g(\widetilde{X}, \hat{r}_1^{[1]}, \bar{\xi}) \leq 0$ for an arbitrary point $\hat{r}_1^{[1]}$ (not shown) between the integration limits.

Proof: *See Appendix C.* □

The vector $r^{[j]}$ defines the limits of the integration of the PDF of the random variable ξ_1 (for further explanations see Fig. 2 and the proof). Note that it is required to solve polynomial equations g in ξ_1 to obtain the vector of integration limits. However, the polynomials can be solved either analytically for low order polynomials, or numerically using efficient root finding algorithms.

It is important to note that Proposition 1 makes the implicit assumption that only samples with unique roots $r^{[j]}$ are used. It is expected that samples with nonunique roots are relatively rare, such that the error will be negligible when such samples are discarded (in the case study in Sec. 6 no samples were discarded).

With Proposition 1, the gradients can be computed efficiently since time-consuming finite differencing is avoided. Furthermore, the approach does not suffer from the discretization effects shown in Fig. 2a. Note that the presented approach is inspired by [24]. However, the extension made in Proposition 1 allows for a much broader applicability such as polynomial and joint chance constraints, as well as non-gaussian probability distributions.

6. Stochastic NMPC of a Williams-Otto Reactor

In this section, the solution to (6) will be illustrated based on shrinking horizon SNMPC of a William-Otto semi-batch reactor. To this end, uncertainties are propagated through the nonlinear system dynamics using the polynomial chaos approach presented in Sec. 3. The sample-average approximation of the chance constraints (16) and of their gradients (19) proposed in Sec. 4 resp. Sec. 5 are used. The theorems presented in Sec. 4.2 are used to guarantee a desired feasibility probability by choosing the required constraint tightening and sample size according to equation (17).

6.1. Williams-Otto Reactor

The Williams-Otto semi-batch reactor is considered, in which the reactions $A + B \xrightarrow{\rho_1} C$, $C + B \xrightarrow{\rho_2} P + E$, and $P + C \xrightarrow{\rho_3} G$ take place with the associated reaction rates ρ_1, ρ_2, and ρ_3 [59, 60]. The reactant A is introduced into the reactor at the beginning of the batch process, whereas reactant B is fed into the reactor with feed rate $u_1(t)$. During

10

the exothermic reactions, the products P and E as well as the side-product G are formed. The reactor temperature is directly manipulated using input $u_2(t)$.

The dynamic model consists of seven differential equations

$$
\begin{aligned}
\dot{x}_1(t) &= -\rho_1(t) && && -\frac{x_1(t)u_1(t)}{x_7(t)} \\
\dot{x}_2(t) &= -\rho_1(t) \,-\rho_2(t) && +\frac{c_{B,in}u_1(t)}{x_7(t)} -\frac{x_2(t)u_1(t)}{x_7(t)} \\
\dot{x}_3(t) &= \;\;\;\rho_1(t) \,-\rho_2(t) -\rho_3(t) && -\frac{x_3(t)u_1(t)}{x_7(t)} \\
\dot{x}_4(t) &= \qquad\quad\;\; \rho_2(t) -\rho_3(t) && -\frac{x_4(t)u_1(t)}{x_7(t)} \\
\dot{x}_5(t) &= \qquad\quad\;\; \rho_2(t) && -\frac{x_5(t)u_1(t)}{x_7(t)} \\
\dot{x}_6(t) &= \qquad\qquad\qquad\;\; \rho_3(t) && -\frac{x_6(t)u_1(t)}{x_7(t)} \\
\dot{x}_7(t) &= \;\; u_1(t),
\end{aligned}
\tag{21}
$$

where $x(t) \coloneqq [x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t)]^\top$ denotes the molar concentrations (in units of mol/m^3) of A, B, C, P, E, and G with respect to the volume $x_7(t)$ (in units of m^3) contained in the reactor at time t (in seconds). $c_{B,in} = 5$ is the molar concentration of B in the inlet reactor feed u_1 (in units of m^3/s) and u_2 (in units of K). The last terms in the first six equations are the dilution effect due to inflow of reactant B. The nonlinear state and input dependent reaction rates are

$$
\begin{aligned}
\rho_1(t) &= k_1 x_1(t)x_2(t) \exp(-6666.7/u_2(t)) \\
\rho_2(t) &= k_2 x_2(t)x_3(t) \exp(-8333.3/u_2(t)) \\
\rho_3(t) &= k_3 x_3(t)x_4(t) \exp(-11111/u_2(t)).
\end{aligned}
$$

The parameter values of the reaction kinetics are taken from [59, 60] and have been converted to SI units. The initial conditions at the beginning of the batch are:

$$
x(0)^\top = [10\; 0\; 0\; 0\; 0\; 0\; 2].
$$

6.1.1. Stochastic Optimal Control Problem

The reaction rate constants k_1, k_2, and k_3 (in units of m^3/(mol s)) are uncertain and probabilistically distributed according to Normal distributions.

$$
\begin{aligned}
k_1 &\sim \mathrm{Norm}(1.6599 \cdot 10^6, 1.6599 \cdot 10^5) \\
k_2 &\sim \mathrm{Norm}(7.2117 \cdot 10^8, 7.2117 \cdot 10^7) \\
k_3 &\sim \mathrm{Norm}(2.6745 \cdot 10^{12}, 2.6745 \cdot 10^{11}),
\end{aligned}
$$

where the first argument in $\mathrm{Norm}(\cdot, \cdot)$ specifies the mean and the second the variance. The variances are chosen to be 10 % of the mean values taken from [60].

The objective of the process is to maximize the profit at the end of the batch, which is the difference between the sales of the products E and P (1.0 monetary units per mole given by $x_4(t_f)x_7(t_f)$ resp. $x_5(t_f)x_7(t_f)$ with $x_7(t)$ being the volume at time t) and the costs of raw material B (0.5 monetary units per mole). The objective is defined by

$$
\begin{aligned}
J(x(\cdot), u(\cdot)) = -0.5 c_{B,in} &\left(\mathbb{E}\left[x_7(t_f)\right] - \mathbb{E}\left[x_7(t_0)\right] \right) \\
&+ \mathbb{E}\left[x_5(t_f)x_7(t_f)\right] + 2\mathbb{E}\left[x_4(t_f)x_7(t_f)\right] \\
&\qquad\qquad - 10\left(\mathbb{V}\left[x_5(t_f)x_7(t_f)\right] + \mathbb{V}\left[x_4(t_f)x_7(t_f)\right] \right),
\end{aligned}
$$

where the expectations and variances are computed as shown in Sec. 3.4. The objective takes into account the mean values of the amounts of the desired end-products P and E, as well as the variances of the end products weighted by a positive constant (in units of 1/mol) to reduce the variance of the end-products.

During the batch, constraints on the inputs must be satisfied at the time-points $t_{u,k}$

$$0 \leq u_1(t) \leq 0.002 \tag{22a}$$

$$313 \leq u_2(t) \leq 363 \tag{22b}$$

$$|u_2(t) - u_2(t)| \leq 1. \tag{22c}$$

In addition, a constraint on the side-product G (x_6) at the final time t_f is defined to keep the amount of the undesired side-product at the end of the batch below a certain threshold. This is because the batch products would be worthless (due to expensive filtering or purifications) if the threshold is exceeded. Due to physical limitations, the volume x_7 at the final time t_f is also kept below a certain threshold. Since $x_6(t_f)$ and $x_7(t_f)$ depend on the random reaction rates and uncertain initial conditions, joint chance constraints are considered with required minimum constraint satisfaction probability $\beta = 0.98$ and a confidence level of $1 - \alpha = 0.99$. It can be seen in Fig. 1 that the desired satisfaction probability is guaranteed for $\beta_{\mathrm{cor}} = 0.985$ and a sample size of $N_S = 5\,000$ samples. This implies

$$\mathbb{P} \begin{bmatrix} x_6(t_F) \leq 0.6 \\ x_7(t_F) \leq 7.0 \end{bmatrix} \geq \beta_{\mathrm{cor}} = 0.985. \tag{23}$$

6.1.2. Polynomial Chaos Expansion

The system (21) is not polynomial, which makes computation of the integrals in the Galerkin projections difficult (cf. Sec. 3.3). A polynomial-in-the-states representation was obtained by defining a new state $x_8(t) := x_7(t)^{-1}$ and its derivative by $\dot{x}_8(t) = -u_1(t)/x_7(t)^2 = -x_8(t)^2 u_1(t)$. With this reformulation, (21) becomes polynomial-in-the-states where each term is separable in the states, parameters and the inputs. However, the employed transformations of the initial conditions $x_8(t_k) = x_7(t_k)^{-1}$ are not polynomials. Hence, the collocation approach in Sec. 3.2 was exploited to obtain a PC approximation (of order $P = 3$) of $x_8(t_k)$ by sampling $x_7(t_k)$. From the PC approximation of $x_8(t_k)$, the coefficients $\widetilde{x_8}(t_k)$ were used as the initial conditions for the extended set of ordinary differential equations obtained by Galerkin projections (12).

Note that similar transformations are always possible if the system has analytic nonlinearities [46] such as exponential terms, rational functions, etc. Therefore, the Galerkin-based PC expansion as presented in Sec. 3 is applicable to a broad class of systems.

The PC expansion in conjunction with the Galerkin projection was applied to the polynomial-in-the-states system using a PC order $P = 3$. The random variables were the reaction rates k_1, k_2, k_3 and the initial conditions $x_1(t_k), \ldots, x_7(t_k)$. This lead to $\widetilde{P} = 286$ terms in the expansion of each state (cf. Eq. (8)), resulting in $2\,002$ differential equations altogether. The first-order sensitivities of the states with respect to the inputs were used to compute the sensitivities and gradients, which resulted in a sensitivity system of $27\,456$ differential equations.

6.2. Closed-Loop Simulations

The shrinking horizon SNMPC was implemented with batch end-time $t_f = 4\,000\,\mathrm{s}$ and sampling times $t_k \in \{0, 250, 500 \ldots, 3\,500, 3\,750\}$. A piecewise-constant input parametrization was chosen with the initial values $u_1(t) = 0.002$ and $u_2(t) = 318$.

In the closed-loop simulations, the nonlinear model (21) was used as the true plant, whose random parameters were drawn from the uncertainty distributions given above. To update the the controller's state information at the beginning of each sampling time interval, normally distributed measurement/observer noise with a standard deviation of 1 % of the mean value was considered.

6.3. Numerical Results

The SNMPC was implemented in Matlab R2012a using `fmincon` with SQP method. Time-critical code (including the ODEs, the Jacobians, the sensitivity equations) was written in C using Matlab CMEX-functions. In particular, the integration of the differential equations was done using the *SUNDIALS* integrator suite [61] using the nonlinear integrator `Functional` in conjunction with the `Adams` method. The optimization was run on a Linux PC (Intel[TM]Core[TM]2 Q6600, 2.4 GHz, 4 cores, 4 GB Ram).

The results of the closed-loop SNMPC are shown in Fig. 3 and, for comparison, the results from the closed-loop NMPC are depicted in Fig. 4. The results show that the presented SNMPC approach guarantees satisfaction of the
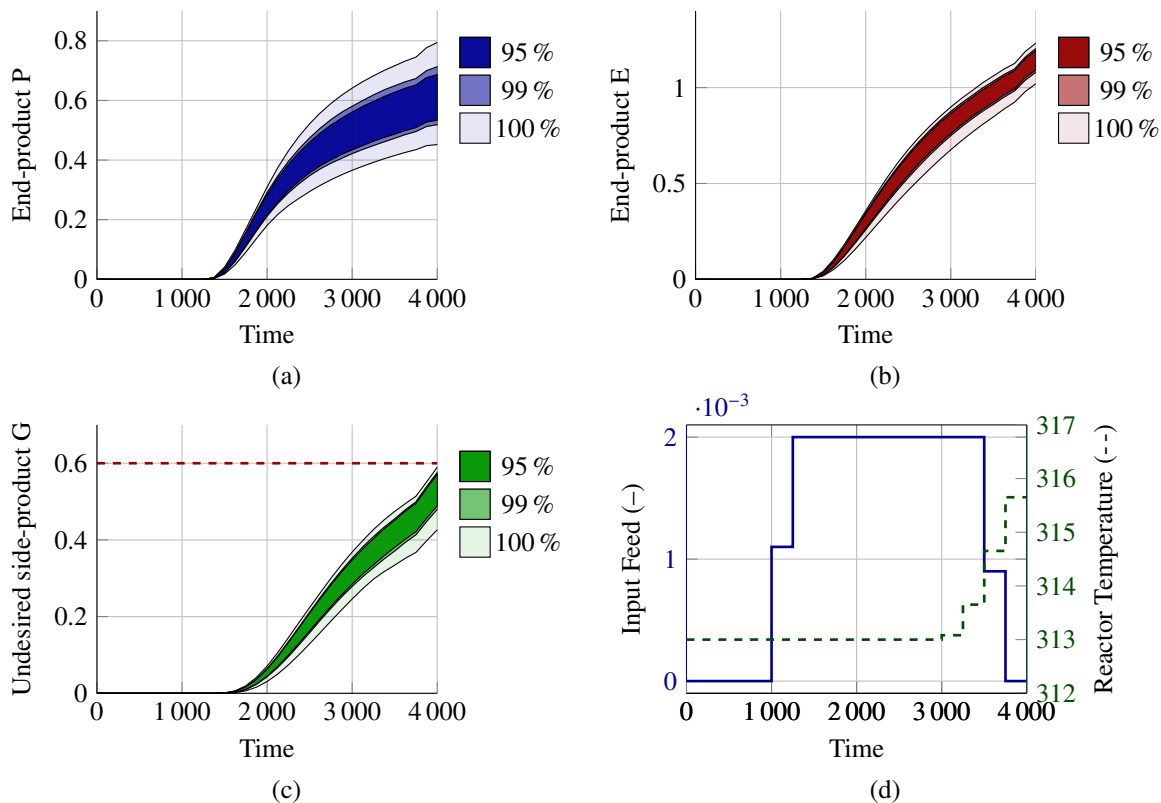
12

Figure 3: Shrinking horizon SNMPC. (a), (b) and (c) Concentrations of the end-products P (x_4), E (x_5), and side-product G (x_6). Depicted are regions containing 90 %, 95 % and 100 % of 350 closed-loop simulations. As can be seen in (c), the uncertainties lead to constraint violations in no cases. (d) Optimal input profile for one representative sample.

joint chance constraint. In contrast, the nominal NMPC caused violation of the constraints (i. e. infeasible solutions) in about 25% of the cases.

Table 1 summarizes the optimization time of the chance constrained optimal control problem at the first sampling time-point. It can be seen that the evaluation of the Jacobians and sensitivity equations, which are required for the computation of the gradients, took the longest. In case of finite-differencing approximation of the gradients, the optimization of the optimal control problem of the first horizon was prematurely stopped after several hours without finding a solution. This demonstrates that the presented approximation of the gradients speeds up the optimization significantly.

The simulation and evaluation of 5 000 samples using the PC approach was on average about 100 times faster (about 2 s) than simulation of the same number of Monte-Carlo samples (about 200 s). This also emphasizes the advantage of the PC approach over a pure sampling or scenario-based approaches in the investigated context. The accuracy of the PC approximation is compared with Monte-Carlo samples in Fig. 5 demonstrating adequate approximation quality.

7. Conclusions and Outlook

This work demonstrates the feasibility of a SNMPC approach for real-time control of a chemical process with uncertain parameters and initial conditions. Polynomial chaos is used for uncertainty propagation and Monte Carlo sampling, which is significantly more efficient than Monte Carlo simulations based on the nonlinear system model. The approximation of the probability densities reduces to solving an expanded set of differential equations (to get the values of the PC coefficients) and matrix multiplications to evaluate the chance constraints. The proposed sample approximation of the joint chance constraints is very flexible and the gradient approximation improves efficiency of
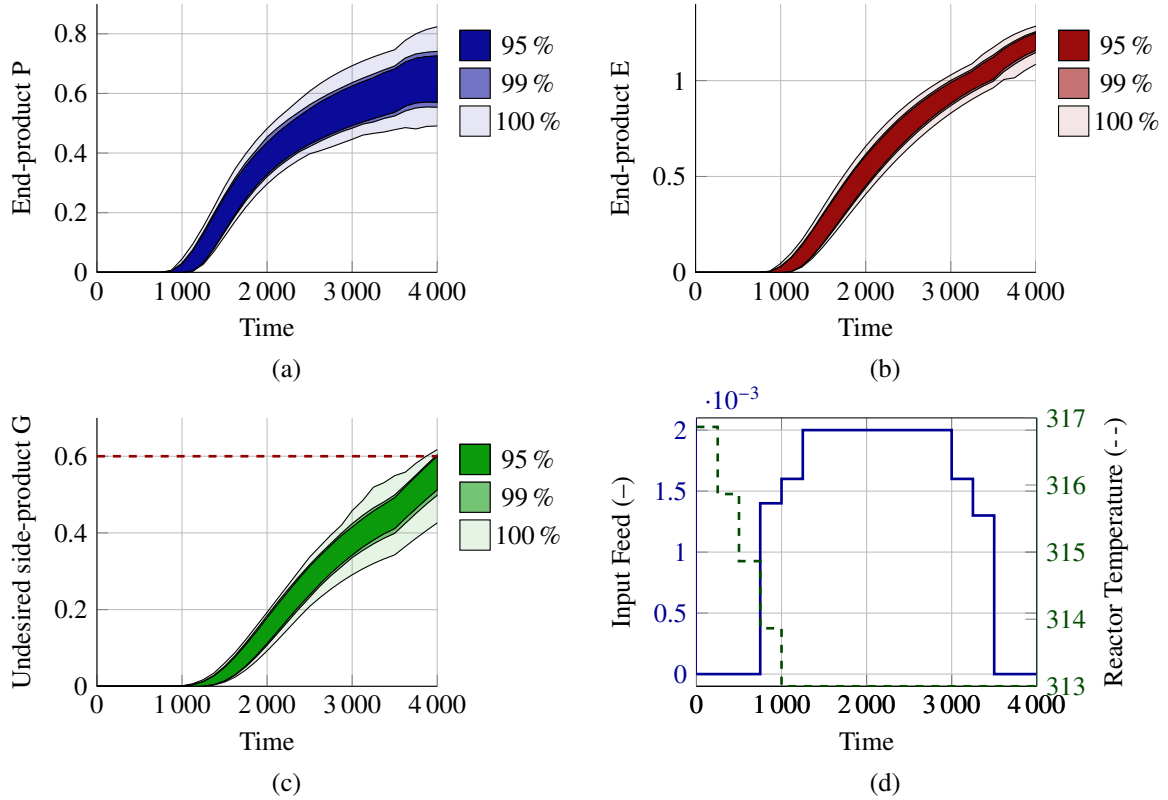
Figure 4: Shrinking horizon NMPC obtained by considering the mean values for each uncertain variable. (a), (b) and (c) Concentrations of the end-products P (x_4), E (x_5), and side-product G (x_6). Depicted are regions containing 90 %, 95 % and 100 % of 1 000 closed-loop simulations. Note that in 225 cases (22.5 %) fmincon failed to find a feasible solution. Out of the remaining 775 cases, 37 (4.8 %) do not satisfy the nominal MPC's inequality constraint $x_6(t_f) \leq 0.6$. (d) Optimal input profile for one representative sample.
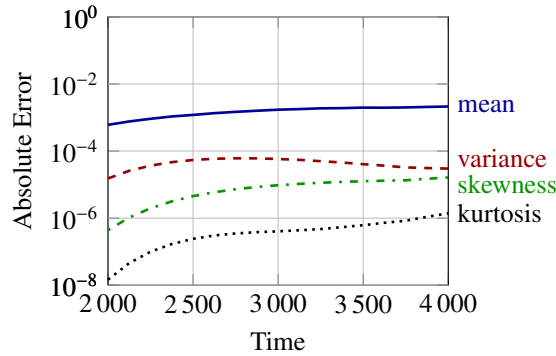

Figure 5: Accuracy of the PCE approximation illustrated through a comparison of 5 000 Monte Carlo samples and samples generated using the PCE. Depicted is the absolute error of the mean, the variance, the skewness and kurtosis for the concentration of G ($x_6(t)$). Note that the curves are not plotted for $t < 2\,000$ since the values are very small (i. e. approach $-\infty$ in the semi-logarithmic plot).

the overall optimization significantly. The confidence analysis provides a systematic way to determine the sample size and the chance constraint tightening required to guarantee a desired feasibility probability and constraint satisfaction probability. This allows the systematic trade-off between large sample sizes (i. e. accurate results, time-consuming) and small sample sizes (i. e. less accurate results, less time-consuming). This could also be useful when one wants to adapt the SNMPC approach to hardware with less computational power, or where timing-constraints have to be

Computing Times for Stochastic NMPC	
Joint chance constraints (5 000 samples, 12 calls)	
Probability of satisfaction and gradients	2.3 s
Integration (24 calls)	
System dynamics	65 s
Evaluation of the Jacobian	162 s
Sensitivity differential equations	288 s
Total time (including `fmincon` and	
further auxiliary functions)	404 s

Table 1: Averaged computing times in seconds for the chance constrained optimal control problem for the initial time-horizon [0, 4 000]. The optimization times for the subsequent sampling time-points were much shorter and the total time ranged from 60 to 150 seconds.

satisfied. Note that the sample approximations and confidence analysis is quite general and independent of the used method to generate samples, and independent of the properties (e. g. convexity) of the optimization problem.

It is important to remark that the probabilities of satisfaction and feasibility hold only for the PC approximation, but not necessarily for the original uncertain nonlinear system. However, as illustrated in the example in this work and mentioned in other work (see introduction), PC allows for accurate predictions of the propagation of stochastic uncertainties through (nonlinear) system dynamics. However, the accuracy is clearly system dependent and to the best of our knowledge no systematic method exist to determine a-priori or depending on the system dynamics the PC order required to achieve a certain accuracy of the estimates. However, there has been recent progress in the error analysis of PC expansions [62] and a line of future research could be to include the error analysis in the synthesis of a robust controller.

From Eq. (7) it becomes clear that the number of monomials required in the PCE grows rapidly with the chosen PC order and number of uncertain variables. Using sparse PCE, i. e. different polynomial orders for different variables [63], is one solution to reduce this complexity. It has been noted [51] that PCE provides accurate predictions for small time only and that the predictions may become inaccurate for increasing time. We believe, however, that the error is negligible in a controlled system with constant update by measurements. In any case, the prediction could be improved by adapting the orthogonal basis as proposed in [51]. Another interesting future extension could be the consideration of time-dependent disturbances (e. g. [40]).

Acknowledgements

Appendix A. Derivation and Proof of Theorem 1

To determine the least conservative constraint tightening, a statistical analysis for the sample approximation (16) (with a given number of samples N_S and confidence level α) is used.

Let $p := \mathbb{P}\left[g\left(\widetilde{X}, \xi\right) \le 0\right]$ (cf. (15)), and the sample approximation $\widehat{p} := \frac{1}{N_S} \sum_{j=1}^{N_S} I_{\mathcal{G}}\left(\xi^{[j]}\right)$ (cf. (16)). The estimation of p is a well-known problem in statistics and corresponds to estimating the success probability of a sequence of Bernuoulli trials [58, 64, 65]. This is due to the fact that testing whether a sample $\xi^{[j]}$ lies in $\mathcal{G}$ (i. e., evaluation of $I_{\mathcal{G}}(\xi^{[j]})$) is a Bernoulli trial with a "yes" or "no" outcome (i. e., "satisfied" or "violated"). Therefore, $\widehat{p}$ is a realization of the random variable $\widehat{P}$ that is distributed according to a binomial distribution, $\widehat{P} \sim \frac{1}{N_S} \text{Bin}(N_S, p)$. The confidence interval $[p_{\text{lb}}(\alpha, N_S, \widehat{P}), p_{\text{ub}}(\alpha, N_S, \widehat{P})]$ consists of a range of values that, with a certain confidence level of $1-\alpha$, contains the true value p. Furthermore, the probability of $p \ge \beta$ given the observation that $p_{\text{lb}}(\alpha, N_S, \widehat{P}) \ge \beta$ is not less than $1 - \alpha$:

$$\mathbb{P}\left[p \ge \beta \mid p_{\text{lb}}(\alpha, N_S, \widehat{P}) \ge \beta\right] \ge 1 - \alpha. \tag{A.1}$$

Since $p_{\text{lb}}(\alpha, N_S, \widehat{P})$ is monotonically increasing in $\widehat{P}$, it holds that

$$\exists \beta_{\text{cor}} \in [\beta, 1] \text{ s.t. } \widehat{P} \geq \beta_{\text{cor}} \text{ iff } p_{\text{lb}}(\alpha, N_S, \widehat{P}) \geq \beta. \tag{A.2}$$

It follows from (A.1) and (A.2) that

$$\mathbb{P}\left[p \geq \beta \mid \widehat{P} \geq \beta_{\text{cor}}\right] \geq 1 - \alpha.$$

Thus, the lower confidence bound is used to determine β_{cor}, since we demand that $p_{\text{lb}}(\alpha, N_S, \widehat{P}) \geq \beta$, meaning that the confidence interval shall not cover the infeasible interval $p < \beta$. The lower confidence bound can be determined approximatively from the quantiles of a normal distribution or exactly from the percentiles of the inverse cumulative Beta distribution betainv [58, 64, 65] as

$$p_{\text{lb}}(\alpha, N_S, \lfloor\beta_{\text{cor}}N_S\rfloor) = 1 - \text{betainv}\left(1 - \frac{\alpha}{2}, N_S + 1 - \lfloor\beta_{\text{cor}}N_S\rfloor, \lfloor\beta_{\text{cor}}N_S\rfloor\right).$$

Here $\lfloor\beta_{\text{cor}}N_S\rfloor \in \mathbb{N}$ is the number of "satisfied" Bernoulli trials, and β_{cor} represents a corrected (i.e. tightened) satisfaction probability. Since the quantiles are exact and due to the equivalence above, it follows that β_{cor} is the least conservative constraint tightening. The statement (17) in the theorem follows directly from (A.3) and from the requirement $p_{\text{lb}}(\alpha, N_S, \lfloor\beta_{\text{cor}}N_S\rfloor) \geq \beta$. $\qquad\square$

Appendix B. Proof of Theorem 2

Sampling only affects the cost function (6a) and the chance constraints (6c), which means that π' satisfies all other constraints in (6). Due to Theorem 1, π' satisfies the chance constraint (6c) with a confidence level not less than $1 - \alpha$. From this it can be concluded that π' is a feasible point of (6) with a probability not less than $1 - \alpha$. $\qquad\square$

Note that π' is not necessarily an optimal solution to (6) because the sample approximation of the cost function may differ from the exact value of the cost function.

Appendix C. Derivative and Proof of Proposition 1

For simpler presentation, we provide the proof for a scalar function $g(\cdot)$ (i.e. for individual chance constraints) first. The extension to vector-valued functions $\boldsymbol{g}(\cdot)$ (i.e. joint chance constraints) is straightforward and is done last.

The proof follows from the construction as described in the following.

Reformulations

Since the first-order sensitivities are assumed to be given, what remains to be computed are the partial derivatives $\frac{\partial\mathbb{P}[\cdot]}{\partial\overline{X}}$ For this purpose, we derive another approximation of the satisfaction probability (15), which then allows to determine $\frac{\partial\mathbb{P}[\cdot]}{\partial\overline{X}}$. The main idea is to approximate the n_ξ-dimensional integration in (15) by integration over only one random variable, say ξ_1, while keeping the remaining $n_\xi - 1$ variables in $\overline{\boldsymbol{\xi}} := [\xi_2, \ldots, \xi_{n_\xi}]$ fixed to values as determined by sampling. We assumed, without loss of generality, that we integrate with respect to ξ_1. The analysis can be performed analogously for any other random variable ξ_i, $i = 2, \ldots, n_\xi$.

Assume (for now) that for each sample $\overline{\boldsymbol{\xi}}^{[j]}$, a vector defining the integration limits $\boldsymbol{r}^{[j]} := [r_0^{[j]}, r_1^{[j]}, \ldots, r_{n_r^{[j]}+1}^{[j]}]$ (see Fig. 2b) is given. [8] With these integration limits given, Eq. (15) can be approximated[9] as

$$\mathbb{P}\left[g\left(\widetilde{X}, \xi_1, \overline{\boldsymbol{\xi}}\right) \leq \boldsymbol{0}\right] \approx \frac{1}{N_S} \sum_{j=1}^{N_S} \left[\sum_{i=0}^{n_{r^{[j]}}} I_{\mathcal{G}}\left(\hat{r}_i^{[j]}, \overline{\boldsymbol{\xi}}^{[j]}\right) \int_{r_i^{[j]}}^{r_{i+1}^{[j]}} \mu_1(d\xi_1)\right], \tag{C.1}$$

---

[8]The integration limits formally depend on $\widetilde{X}$ and on the sample $\overline{\boldsymbol{\xi}}^{[j]}$, and that the number $n_{r^{[j]}}$ may be different for each sample. See Fig. 2b for further explanations. However, we omit this dependency for shorter notation.

[9]Eq. (C.1) provides a better approximation of (15) than (16) since the entire probability distribution of ξ_1 is taken into account; however, this approximation requires the evaluation of the cumulative probability density function of μ_1, which is for many distributions not available in an analytic form.

where the indicator function is evaluated at the point $(\hat{r}_i^{[j]}, \overline{\boldsymbol{\xi}}^{[j]})$, with $\hat{r}_i^{[j]}$ being an arbitrary point from the interior of the integration interval $(r_i^{[j]}, r_{i+1}^{[j]})$, as e.g. the mid-point $\hat{r}_i^{[j]} = \frac{r_i^{[j]} + r_{i+1}^{[j]}}{2}$ in (19).

Based on the vector of integration bounds $\boldsymbol{r}^{[j]}$, the gradient can be approximated by derivative of (C.1) with respect to $\widetilde{X}$, which gives (19). Note that we used the facts that the order of integration and differentiation can be changed. Since the integration limits $r_i^{[j]}$ in Eq. (C.1) depend on the (known and constant) $\widetilde{X}$, the partial derivatives $\frac{\partial r_i^{[j]}}{\partial \widetilde{X}}$ have to be taken into account, and they are given by (20). The first row in (20) follows from the fact that $r_0^{[j]}$ and $r_{n_{r^{[j]}}+1}^{[j]}$ are constants and defined by the minimum and maximum of the support of ξ_1; the second row in (20) follows from the implicit function theorem.

From Eq. (20) it becomes clear that one has to avoid division by zero, which is the case if $g\left(\widetilde{X}, \xi_1, \overline{\boldsymbol{\xi}}\right) = 0$ has multiple roots with respect to ξ_1. These cases have to be treated and are first discussed for individual chance constraints. The extension to joint chance constraints follows immediately from that.

Integration Limits for Individual Chance Constraints

As illustrated in Fig. 2, the integration limits are defined by the solutions of $g\left(\widetilde{X}, \xi_1, \overline{\boldsymbol{\xi}}^{[j]}\right) = 0$ or, in other words, the values of ξ_1 where the indicator function $I_{\mathcal{G}}\left(r_i^{[j]}, \overline{\boldsymbol{\xi}}^{[j]}\right)$ changes its value when traversing along the direction of ξ_1 for fixed values of $\overline{\boldsymbol{\xi}}^{[j]}$. In the first step, we therefore solve the equations $g_i\left(\widetilde{X}, \xi_1, \overline{\boldsymbol{\xi}}^{[j]}\right) = 0$ for ξ_1 and determine all its (real-valued) roots on the support of the random variable. Due to the PC expansion involving polynomials in ξ_1 of order usually greater than one, the equation is possible to have no unique (real-valued) solution. From Eq. (20) it is clear that roots with multiplicity greater than 1 lead to division by zero and therefore cannot be considered, and the sample has to be discarded.

In the next step, all (real-valued) roots are sorted such that $r_1^{[j]} \leq r_2^{[j]} \leq \cdots \leq r_{n_{r^{[j]}}}^{[j]}$, where $n_{r^{[j]}}$ denotes the number of real-valued roots. As the last step, we introduce $r_0^{[j]}$ (resp. $r_{n_{r^{[j]}}+1}^{[j]}$) corresponding to the lowest (resp. largest) value that ξ_1 can take on its support (which may also be $\pm\infty$). With that, one obtains the vector $\boldsymbol{r}^{[j]} := [r_0^{[j]}, r_1^{[j]}, \ldots, r_{n_{r^{[j]}}+1}^{[j]}]$.

Integration Limits for Joint Chance Constraints

One proceeds similar as for individual chance constraints and determines the real-valued roots on the support of ξ_1 for each equation $g_i(\cdot)$, $i = 1, \ldots, N_g$. As for individual chance constraints, samples having (real-valued) roots with multiplicity greater than one are discarded. In addition, one has to take care of the fact that roots of the different equations may coincide. To avoid such degenerate cases and the ambiguity in applying equation (20), such samples are also discarded. After that, all roots as well as $r_0^{[j]}$ and $r_{n_{r^{[j]}}+1}^{[j]}$ (see individual chance constraints) are collected in vector in ascending order. When evaluating (20), one has to use the function $g_i(\cdot)$ for which $r_i^{[j]}$ is a root.

With that, all elements and equations are established, which completes the proof. $\square$

References

[1] S. J. Qin, T. A. Badgwell, An overview of industrial model predictive control technology, in: AIChE Symposium Series, Vol. 93, American Institute of Chemical Engineers, 1997, pp. 232–256.

[2] J. B. Rawlings, Tutorial overview of model predictive control, IEEE Control Systems Magazine 20 (3) (2000) 38–52.

[3] M. Morari, J. H. Lee, Model predictive control: past, present and future, Computers & Chemical Engineering 23 (4-5) (1999) 667–682.

[4] A. Bemporad, M. Morari, Robust model predictive control: A survey, in: A. Garulli, A. Tesi (Eds.), Robustness in Identification and Control, Springer, Berlin, London, New York, 1999, pp. 207–226.

[5] S. Garatti, M. C. Campi, Modulating robustness in control design: Principles and algorithms, IEEE Control Systems Magazine 33 (2) (2013) 36–51.

[6] M. Vidyasagar, Randomized algorithms for robust controller synthesis using statistical learning theory, Automatica 37 (2001) 1515–1528.

[7] A. Schwarm, M. Nikolaou, Chance-constrained model predictive control, AIChE Journal 45 (1999) 1743–1752.

[8] P. Li, M. Wendt, G. Wozny, Robust model predictive control under chance constraints, Computers & Chemical Engineering 24 (2) (2000) 829–834.

[9] D. H. V. Hessem, O. H. Bosgra, A full solution to the constrained stochastic closed-loop MPC problem via state and innovations feedback and its receding horizon implementation, in: Proc. 42nd IEEE Conference on Decision and Control (CDC), Maui, 2003, pp. 929–934.

[10] P. Li, H. Arellano-Garcia, G. Wozny, Chance constrained programming approach to process optimization under uncertainty, Computers & Chemical Engineering 32 (1) (2008) 25–45.

[11] M. Cannon, B. Kouvaritakis, S. V. Rakovic, Q. Cheng, Stochastic tubes in model predictive control with probabilistic constraints, IEEE Transactions on Automatic Control 56 (1) (2011) 194–200.

[12] J. Primbs, C. Sung, Stochastic receding horizon control of constrained linear systems with state and control multiplicative noise, IEEE Transactions on Automatic Control 54 (2009) 221–230.

[13] D. Bernardini, A. Bemporad, Scenario-based model predictive control of stochastic constrained linear systems, in: Proc. 48th IEEE Conference on Decision and Control (CDC), Shanghai, 2009, pp. 6333–6338.

[14] M. Cannon, B. Kouvaritakis, X. Wu, Model predictive control for systems with stochastic multiplicative uncertainty and probabilistic constraints, Automatica 45 (1) (2009) 167 – 172.

[15] X. Zhang, K. Margellos, P. Goulart, J. Lygeros, Stochastic model predictive control using a combination of randomized and robust optimization, in: Proc. IEEE 52nd Annual Conference on Decision and Control (CDC), 2013, pp. 7740–7745.

[16] L. Blackmore, M. Ono, A. Bektassov, B. C. Williams, A probabilistic particle-control approximation of chance-constrained stochastic predictive control, Robotics, IEEE Transactions on 26 (3) (2010) 502–517.

[17] F. Oldewurtel, C. N. Jones, M. Morari, A tractable approximation of chance constrained stochastic MPC based on affine disturbance feedback, in: Proc. 47th IEEE Conference on Decision and Control (CDC), 2008, pp. 4731–4736.

[18] B. Kouvaritakis, M. Cannon, D. Muñoz-Carpintero, Efficient prediction strategies for disturbance compensation in stochastic MPC, International Journal of Systems Science 44 (7) (2013) 1344–1353.

[19] M. Korda, R. Gondhalekar, F. Oldewurtel, C. N. Jones, Stochastic MPC framework for controlling the average constraint violation, IEEE Transactions on Automatic Control 59 (7) (2014) 1706–1721.

[20] T. Hashimoto, Probabilistic constrained model predictive control for linear discrete-time systems with additive stochastic disturbances, in: Proc. IEEE 52nd Annual Conference on Decision and Control (CDC), 2013, pp. 6434–6439.

[21] M. Farina, L. Giulioni, L. Magni, R. Scattolini, A probabilistic approach to model predictive control, in: Proc. IEEE 52nd Annual Conference on Decision and Control (CDC), 2013, pp. 7734–7739.

[22] G. C. Calafiore, L. E. Ghaoui, On distributionally robust chance-constrained linear programs, Journal of Optimization Theory and Application 130 (1) (2006) 1–22.

[23] A. Geletu, M. Klöppel, H. Zhang, P. Li, Advances and applications of chance-constrained approaches to systems optimisation under uncertainty, International Journal of Systems Science 44 (7) (2013) 1209–1232.

[24] J. O. Royset, E. Polak, Reliability-based optimal design using sample average approximations, Probabilistic Engineering Mechanics 19 (4) (2004) 331–343.

[25] A. Shapiro, Stochastic programming approach to optimization under uncertainty, Mathematical Programming 112 (2008) 183–220.

[26] G. C. Calafiore, M. C. Campi, The scenario approach to robust control design, IEEE Transactions on Automatic Control 51 (5) (2006) 742–753.

[27] M. C. Campi, S. Garatti, The exact feasibility of randomized solutions of uncertain convex programs, SIAM Journal on Optimization 19 (2008) 1211–1230.

[28] G. Calafiore, Random convex programs, SIAM Journal on Optimization 20 (6) (2010) 3427–3464.

[29] G. Schildbach, G. C. Calafiore, L. Fagiano, M. Morari, Randomized model predictive control for stochastic linear systems, in: Proc. American Control Conference (ACC), Montreal, 2012, pp. 417–422.

[30] G. C. Calafiore, L. Fagiano, Robust model predictive control via scenario optimization, IEEE Transactions on Automatic Control 58 (1) (2013) 219–224.

[31] T. Alamo, R. Tempo, A. Luque, On the sample complexity of probabilistic analysis and design methods, in: Perspectives in Mathematical System Theory, Control, and Signal Processing, Springer-Verlag London, London, 2010, pp. 39–50.

[32] T. Alamo, R. Tempo, A. Luque, D. Ramirez, Randomized methods for design of uncertain systems: Sample complexity and sequential algorithms, arXiv: 1110.1892v2.

[33] A. Geletu, M. Klöppel, A. Hoffmann, P. Li, A tractable approximation of non-convex chance constrained optimization with non-Gaussian uncertainties, Engineering Optimization (ahead-of-print) (2014) 1–26.

[34] C. Feng, F. Dabbene, C. M. Lagoa, A kinship function approach to robust and probabilistic optimization under polynomial uncertainty, IEEE Transactions on Automatic Control 56 (7) (2011) 1509–1523.

[35] A. Mesbah, S. Streif, R. Findeisen, R. D. Braatz, Stochastic nonlinear model predictive control with probabilistic constraints, in: Proc. American Control Conference (ACC), Portland, Oregon, 2014, pp. 2413–2419.

[36] L. Imsland, P. Kittilsen, T. S. Schei, Model-based optimizing control and estimation using modelica model, Modeling, Identification and Control 31 (3) (2010) 107–121.

[37] J. Garnier, A. Omrane, Y. Rouchdy, Asymptotic formulas for the derivatives of probability functions and their monte carlo estimations, European Journal of Operational Research 198 (3) (2009) 848–858.

[38] J. A. Paulson, A. Mesbah, S. Streif, R. Findeisen, R. D. Braatz, Fast stochastic model predictive control of high-dimensional systems, Vol. In Press, 2014.

[39] L. Fagiano, M. Khammash, Nonlinear stochastic model predictive control via regularized polynomial chaos expansions, in: Proc. 51st IEEE Conference on Decision and Control (CDC), Maui, 2012, pp. 142–147.

[40] T. Huschto, S. Sager, Stochastic optimal control in the perspective of the wiener chaos, in: Proc. European Control Conference (ECC), Zurich, 2013, pp. 3059–3064.

[41] K.-K. K. Kim, R. D. Braatz, Generalised polynomial chaos expansion approaches to approximate stochastic model predictive control, International Journal of Control 86 (8) (2013) 1324–1337.

[42] N. Wiener, The homogeneous chaos, American Journal of Mathematics 60 (1938) 897–936.

[43] R. Ghanem, P. Spanos, Stochastic Finite Elements - A Spectral Approach, Springer-Verlag, New York, 1991.

[44] D. Xiu, G. E. Karniadakis, The wiener-askey polynomial chaos for stochastic differential equations, SIAM Journal of Scientific Computation

24 (2002) 619–644.

[45] K.-K. K. Kim, D. E. Shen, Z. K. Nagy, R. D. Braatz, Wiener's polynomial chaos for the analysis and control of nonlinear dynamical systems with probabilistic uncertainties, IEEE Control Systems Magazine 33 (5) (2013) 58–67.

[46] T. Ohtsuka, Model structure simplification of nonlinear systems via immersion, IEEE Transactions on Automatic Control 50 (5) (2005) 607–618.

[47] J. Fisher, R. Bhattacharya, Linear quadratic regulation of systems with stochastic parameter uncertainties, Automatica 45 (12) (2009) 2831–2841.

[48] G. Calafiore, M. C. Campi, Uncertain convex programs: randomized solutions and confidence levels, Mathematical Programming 102 (1) (2005) 25–46.

[49] R. H. Cameron, W. T. Martin, The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals, Annals of Mathematics 48 (1947) 385–392.

[50] S. Oladyshkin, W. Nowak, Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion, Reliability Engineering & System Safety 106 (2012) 179–190.

[51] M. Gerritsma, J.-B. Van der Steen, P. Vos, G. Karniadakis, Time-dependent generalized polynomial chaos, Journal of Computational Physics 229 (22) (2010) 8333–8363.

[52] Z. Nagy, R. Braatz, Distributional uncertainty analysis using power series and polynomial chaos expansions, Journal of Process Control 17 (3) (2007) 229–240.

[53] A. Mesbah, S. Streif, R. Findeisen, R. D. Braatz, Active fault diagnosis for nonlinear systems with probabilistic uncertainties, in: Proc. 19th IFAC World Congress, Cape Town, South Africa, 2014, pp. 7079–7084.

[54] M. A. Tatang, W. Pan, R. G. Prinn, G. J. McRae, An efficient method for parametric uncertainty analysis of numerical geophysical models, Journal of Geophysical Research: Atmospheres (1984–2012) 102 (D18) (1997) 21925–21932.

[55] W. Gautschi, R. S. Friedman, J. Burns, R. Darjee, A. Mcintosh, Orthogonal Polynomials: Computation and Approximation, Numerical Mathematics and Scientific Computation Series, Oxford University Press, Oxford, U.K., 2004.

[56] S. Streif, F. Petzke, A. Mesbah, R. Findeisen, R. D. Braatz, Optimal experimental design for probabilistic model discrimination using polynomial chaos, in: Proc. 19th IFAC World Congress, Cape Town, South Africa, 2014, pp. 4103–4109.

[57] M. Campi, S. Garatti, A sampling-and-discarding approach to chance-constrained optimization: Feasibility and optimality, Journal of Optimization Theory and Applications 148 (2) (2011) 257–280.

[58] R. Rossi, Solving stochastic constraint programs via sampling, arXiv: 1110.1892v2.

[59] R. Hannemann, W. Marquardt, Continuous and discrete composite adjoints for the Hessian of the Lagrangian in shooting algorithms for dynamic optimization, SIAM Journal on Scientific Computing 31 (6) (2010) 4675–4695.

[60] T. J. Williams, R. E. Otto, A generalized chemical processing model for the investigation of computer control, Transactions of the American Institute of Electrical Engineers, Part I: Communication and Electronics 79 (5) (1960) 458–473.

[61] A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, C. S. Woodward, SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers, ACM Trans. Math. Softw. 31 (3) (2005) 363–396.
URL http://dx.doi.org/10.1145/1089014.1089020

[62] W. Shi, C. Zhang, Error analysis of generalized polynomial chaos for nonlinear random ordinary differential equations, Applied Numerical Mathematics 62 (12) (2012) 1954–1964.

[63] W. Luo, Wiener chaos expansion and numerical solutions of stochastic partial differential equations, Ph.D. thesis, California Institute of Technology (2006).

[64] M. Hollander, D. A. Wolfe, E. Chicken, Nonparametric Statistical Methods, John Wiley & Sons, 2013.

[65] O. Stan, R. Sirdey, J. Carlier, D. Nace, The robust binomial approach to chance-constrained optimization problems with application to stochastic partitioning of large process networks, Journal of Heuristics 20 (2014) 1–30.