Output feedback stochastic nonlinear model predictive control of a polymerization batch process

Eric Bradford[1], Marcus Reble[2], and Lars Imsland[3]

Abstract— **Nonlinear model predictive control (NMPC) is one of the few methods that can handle multivariate nonlinear control problems while accounting for process constraints. Many dynamic models are however affected by significant stochastic uncertainties that can lead to closed-loop performance problems and infeasibility issues. In this paper we propose a novel stochastic NMPC (SNMPC) algorithm to optimize a probabilistic objective while adhering chance constraints for feasibility in which only noisy measurements are observed at each sampling time. The system predictions are assumed to be both affected by parametric and additive stochastic uncertainties. In particular, we use polynomial chaos expansions (PCE) to expand the random variables of the uncertainties. These are updated using a PCE nonlinear state estimator and exploited in the SNMPC formulation. The SNMPC scheme was verified on a complex polymerization semi-batch reactor case study.**

I. INTRODUCTION

Batch processes play a vital role for the manufacture of high value products in many sectors of the chemical industry, such as pharmaceuticals, polymers, biotechnology, and food. The main reason for the continued use of batch processes is their inherent flexibility to produce multiple products and deal with variations in feedstock, product specifications, and market demand. The control of batch reactors is often challenging due to their frequently highly nonlinear behaviour and operation at states that are not steady states. Therefore, there is an increased acceptance in industry for the application of nonlinear model predictive control (NMPC) to address these challenges [1].

Model predictive control (MPC) is an advanced control method that has been employed to a significant extent in industry due to its ability to deal with multivariate plants and process constraints. MPC solves an optimal control problem (OCP) based on an explicit dynamic model at each sampling time to determine a finite sequence of control actions [2]. Due to various uncertainties however the dynamic system behaviour may be significantly different from the predicted behaviour of the dynamic model. This may lead to constraint violations and a worse control performance. If we assume the uncertainties to lie in a bounded set, robust MPC (RMPC) methods are available to deal with this problem [3]. For robust NMPC, min-max NMPC [4] and tube-based NMPC [5] have been proposed among others. While these approaches can give stability and performance guarantees in the worst-case, the probability of occurrence of the worst-case realization may be very small and hence lead to a too conservative solution [6].

Alternatively stochastic MPC (SMPC) may be employed, for which the uncertainties are given by known probability density functions (pdf). Constraints and objective in this context are addressed in a probabilistic sense, allowing for a pre-defined level of constraint violations in probability and thereby alleviating the previously mentioned problem by trading-off risk with closed-loop performance [6]. SMPC has been largely focused on linear systems [7], such as tube-based SMPC [8], scenario-based SMPC [9] and SMPC using affine-parametrizations [10], while stochastic NMPC (SNMPC) has received relatively little attention.

The main difficulty of SNMPC is propagating continuous stochastic uncertainties through nonlinear equations without being prohibitively computationally expensive. Several methods have been proposed for approximating this case, some of which have been successfully applied to formulate SNMPC approaches: Unscented transformations [11], Polynomial chaos expansions (PCE) [12], Quasi Monte Carlo (MC) [13], Markov Chain MC [14], Gaussian processes [15], Gaussian mixtures (GM) [16], Fokker-Planck [17], linearization [18], and particle filters [19]. The control of systems with discrete stochastic uncertainties has been addressed in [20]. Most SNMPC work is based on full state feedback, but there are several algorithms that have been proposed for output feedback. The unscented transformation work in [11] assumes feedback from the Unscented Kalman filter, in [21] a probabilistic high-gain observer is proposed to be jointly used with a continuous-time SNMPC formulation and lastly [19] use the particle filter equations for both state estimation and uncertainty propagation.

PCE has received a lot of attention for SNMPC, which can be seen as a sampling-based MPC algorithm to approximate both probabilistic constraints and objectives. In [12] PCEs are used to approximate objectives and constraints in expectation. The work by [22] extends the approach to include chance constraints by using Chebyshev's inequality, while in [23] the chance constraints are instead approximated using MC sampling. This leads to computationally more expensive, but less conservative approximations of the chance constraints. PCE are only able to represent time-invariant uncertainties, which causes problems with commonly assumed time-varying uncertainties. This issue has been addressed in [24] for additive noise and in [25] for non-additive noise by using conditional probability rules to essentially deal with time-varying and time-invariant uncertainties separately.

1 E. Bradford and 3 L. Imsland are with the Faculty of Information Technology and Electrical Engineering, Department of Engineering Cybernetics, NTNU, 7491 Trondheim, Norway {eric.bradford,lars.imsland}@ntnu.no

2 M. Reble is with BASF SE, 67056 Ludwigshafen, Germany marcus.reble@basf.com

Apart from SNMPC, PCEs have also been extensively used for nonlinear filtering. Firstly, PCEs can be used as a cheap surrogate to obtain mean and covariance estimates of nonlinear transformations. This has been applied to yield various PCE Kalman filters (KF) for nonlinear estimation problems, including a PCE ensemble KF [26] and a PCE extended KF [27]. Apart from the PCE KFs, several authors have applied Bayes' rule directly to the PCE expansion to attain the posterior distribution of the states. In [28] PCEs are used to propagate uncertainties, from which the moments are fitted to a GM and used in Bayes' rule. Similarly in [29] the same procedure is used, however the update is carried out using linear update laws considering higher order moments. Using sampling the posterior moments of a PCE expansion can be obtained from Bayes' rule and used to fit a posterior PCE expansion by updating the coefficients as shown in [30]. The method in [31] similarly to [30] uses posterior moments from Bayes' rule, but accounts for time-varying additive disturbances and uses the PCE in addition for uncertainty propagation. [32] propose to use PCEs for uncertainty propagation in conjunction with a particle filter.

In this paper we extend the work in [33] that proposes a SNMPC algorithm for output feedback using the nonlinear filter proposed by [30] and a PCE SNMPC algorithm as in [22]. In the previous work additive process noise was ignored due to the issue time-varying uncertainties cause for PCE based methods. To address these issues we extended the approach in [30] to be able to handle additive disturbance noise and in addition formulate an efficient SNMPC algorithm using a sparse Gauss-Hermite (sGH) sampling rule. The framework is verified on an extensive case study of a semi-batch polymerization reaction, for which we directly minimise the required batch time subject to safety and product quality constraints. The paper is comprised of the following sections. In the subsequent section a general problem definition is given and the main algorithm is introduced. In section 3 we give some background on PCE. In section 4 the PCE state estimator is outlined, while in section 5 we introduce the Gauss-Hermite SNMPC formulation. Section 6 defines the case study, while in section 7 the results and discussions of the case study are presented. Lastly, in section 8 conclusions are given.

II. PROBLEM SETUP

The problem to be solved is outlined in this section. Consider a discrete-time system of nonlinear equations with stochastic parameters and additive disturbance noise:

$$\mathbf{x}'_{t+1} = \mathbf{f}'(\mathbf{x}'_t, \mathbf{u}_t) + \mathbf{w}_t, \quad \mathbf{x}'_0 = \mathbf{x}'_0(\boldsymbol{\xi}) \tag{1}$$

$$\mathbf{y}_t = \mathbf{h}'(\mathbf{x}'_t) + \boldsymbol{\nu}_t \tag{2}$$

where t is the discrete time, $\mathbf{x}' = [\mathbf{x}, \boldsymbol{\theta}]^T$ is an augmented state vector, $\mathbf{x} \in \mathbb{R}^{n_\mathbf{x}}$ are the system states, $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$ are parametric uncertainties, $\mathbf{u} \in \mathbb{R}^{n_\mathbf{u}}$ denote the control inputs, $\mathbf{f}'(\mathbf{x}'_t, \mathbf{u}_t) = [\mathbf{f}(\mathbf{x}'_t, \mathbf{u}_t), \boldsymbol{\theta}_t]^T$ are the dynamic equations for the augmented state vector, $\mathbf{f} : \mathbb{R}^{n_\mathbf{x}+n_\theta} \times \mathbb{R}^{n_\mathbf{u}} \rightarrow \mathbb{R}^{n_\mathbf{x}}$ represents the nonlinear dynamic system for the states, $\mathbf{y} \in \mathbb{R}^{n_\mathbf{y}}$ denote the measurements, $\mathbf{h} : \mathbb{R}^{n_\mathbf{x}+n_\theta} \rightarrow \mathbb{R}^{n_\mathbf{y}}$ are the output equations, $\boldsymbol{\nu} \in \mathbb{R}^{n_\mathbf{y}} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{v})$ is additive measurement noise assumed to follow a zero mean multivariate normal distribution with known covariance matrix $\boldsymbol{\Sigma}_\mathbf{v}$, and $\mathbf{w} \in \mathbb{R}^{n_\mathbf{x}+n_\theta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{w})$ is additive disturbance noise assumed to follow a zero mean multivariate normal distribution with known covariance matrix $\boldsymbol{\Sigma}_\mathbf{w}$. The initial condition $\mathbf{x}'_0$ is assumed to follow a known probability distribution represented by a PCE with $\boldsymbol{\xi} \in \mathbb{R}^{n_\mathbf{x}+n_\theta} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. For background information on PCEs refer to section III.

To approximately represent the probability distribution of $\mathbf{x}'_t$ at each discrete time we use PCEs. Let $\mathbf{x}'_t(\boldsymbol{\xi})$ correspond to the PCE of $\mathbf{x}'$ at time t. It is assumed that we are initially given a PCE of $\mathbf{x}'$ denoted by $\mathbf{x}'_0(\boldsymbol{\xi})$ as shown in Eq.1. Usually this initial probability distribution will be broad with a relatively large variance representing the uncertainty of the initial states and uncertain parameters. At each sampling time $t+1$ we measure $\mathbf{y}_{t+1}$ given by Eq.(2). This measurement is then used to determine $\mathbf{x}'_{t+1}(\boldsymbol{\xi})$ by updating the PCE representation of $\mathbf{x}'_t(\boldsymbol{\xi})$ using Bayes' rule. The nonlinear filter using PCEs is outlined in section IV. Given the PCE $\mathbf{x}'_t(\boldsymbol{\xi})$ at each discrete time t, we wish to control the dynamic system defined by Eq.(1) subject to chance constraints and a stochastic objective. To accomplish this we solve a probabilistic finite time-horizon optimal control problem repeatedly in MPC fashion at each time t:

$$\begin{aligned}
&\underset{\mathbf{U}_N}{\text{minimize}} \quad \mathbb{E}(J(N, \mathbf{x}'_t(\boldsymbol{\xi}), \mathbf{U}_N)) \\
&\text{subject to} \\
&\mathbf{x}'_{k+1} = \mathbf{f}'(\mathbf{x}'_k, \mathbf{u}_k) + \mathbf{w}_k \qquad \forall k \in \mathbb{N}_k \\
&\mathbb{P}(g_j(\mathbf{x}'_k, \mathbf{u}_k) \leq 0) \geq 1 - \epsilon \qquad \forall (k,j) \in \mathbb{N}_{k+1} \times \mathbb{N}_g \\
&\mathbb{P}(g_j^N(\mathbf{x}'_N, \mathbf{u}_N) \leq 0) \geq 1 - \epsilon \quad \forall j \in \mathbb{N}_g^N \\
&\mathbf{u}_k \in \mathbb{U}_k \qquad \forall k \in \mathbb{N}_k \\
&\mathbf{x}'_0 = \mathbf{x}'_t(\boldsymbol{\xi})
\end{aligned} \tag{3}$$

where $\mathbb{N}_g = \{1, \ldots, n_g\}$, $\mathbb{N}_g^N = \{1, \ldots, n_g^N\}$, $\mathbb{N}_k = \{0, ..., N-1\}$, $\mathbb{N}_{k+1} = \{1, ..., N\}$, the expectation of $J(N, \mathbf{x}'_t(\boldsymbol{\xi}), \mathbf{U}_N)$ is the objective, N is the time horizon, the probability of the functions $g_j : \mathbb{R}^{n_\mathbf{x}+n_\theta} \times \mathbb{R}^{n_\mathbf{u}} \rightarrow \mathbb{R}$ over all times and $g_j^N : \mathbb{R}^{n_\mathbf{x}+n_\theta} \times \mathbb{R}^{n_\mathbf{u}} \rightarrow \mathbb{R}$ at the final time exceeding 0 should be less than ϵ, the constraints on the inputs are given by $\mathbb{U}_k \subset \mathbb{R}^{n_\mathbf{u}}$, and lastly $\mathbf{U}_N := [\mathbf{u}_0, \ldots, \mathbf{u}_{N-1}]$ represents the control inputs.

The problem in Eq.(3) cannot be solved exactly since it requires the propagation of stochastic uncertainties through nonlinear transformations and involves chance constraints, which require the evaluation of multivariate integrals. Instead, a simplified problem is formulated in section V approximating Eq.(3) using a sparse Gauss-Hermite rule. Overall we propose to use PCEs introduced in section III to represent the probability distributions of the states $\mathbf{x}_t$ and the uncertain parameters $\boldsymbol{\theta}_t$ jointly denoted as $\mathbf{x}'_t$ at each sampling time t. The SNMPC algorithm formulation in section V exploits this uncertainty description to control the dynamic system in Eq.(1), while the measurements from Eq.(2) are utilised to update the PCE presentation of $\mathbf{x}'_t$

recursively as outlined in section IV. The proposed algorithm is summarised below as Algorithm 1.

Algorithm 1: Output feedback PCE SNMPC

Input : $\mathbf{f}'(\mathbf{x}', \mathbf{u})$, $\mathbf{h}(\mathbf{x}')$, $\mathbf{\Sigma_v}$, $\mathbf{\Sigma_w}$, $\mathbf{x}'_0(\boldsymbol{\xi})$

for *each sampling time* $t = 0, 1, 2, \ldots$ **do**

1) Solve PCE SNMPC problem (3) with $\mathbf{x}'_t(\boldsymbol{\xi})$ and obtain optimal control actions.
2) Apply the first control action to the plant.
3) Measure $\mathbf{y}_{t+1}$.
4) Apply PCE filter to update $\mathbf{x}'_t(\boldsymbol{\xi})$ to $\mathbf{x}'_{t+1}(\boldsymbol{\xi})$.

end

III. BACKGROUND: PCE

The polynomial chaos expansion (PCE) scheme will be briefly outlined in this section, for more information refer to [22], [34], [35]. In this work PCEs are used as an efficient means to represent random variables. It can be shown that a random variable γ with finite second order moments can be expressed as a convergent series expansion:

$$\gamma(\boldsymbol{\xi}) = \sum_{j=0}^{\infty} a_j \phi_{\boldsymbol{\alpha}_j}(\boldsymbol{\xi}) \tag{4}$$

where $\boldsymbol{\xi} \in \mathbb{R}^{n_{\boldsymbol{\xi}}} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ is a $n_{\boldsymbol{\xi}}$-dimensional random variable following a standard normal distribution with zero-mean and unit variance, a_j denote expansion coefficients and $\phi_{\boldsymbol{\alpha}_j} = \prod_{i=1}^{n_{\boldsymbol{\xi}}} \phi_{\alpha_{j_i}}(\xi_i)$ are multivariate polynomials with $\phi_{\alpha_{j_i}}(\xi_i)$ being univariate polynomials of ξ_i of degree α_{j_i}.

The univariate polynomials are chosen to satisfy an orthogonality property according to the pdf of $\boldsymbol{\xi}$, which in the case of standard Gaussian random variables are given by Hermite polynomials. Hermite polynomials He with degree j in terms of ξ_i can be expressed as:

$$He_j(\xi_i) = (-1)^j \exp\left(\frac{1}{2}\xi_i^2\right) \frac{d^j}{d\xi_i^j} \exp\left(-\frac{1}{2}\xi_i^2\right) \tag{5}$$

These orthogonal polynomials have the useful property:

$$\langle \phi_i, \phi_j \rangle = \int \phi_i(\boldsymbol{\xi}) \phi_j(\boldsymbol{\xi}) p(\boldsymbol{\xi}) d\boldsymbol{\xi} = \delta_{ij} \langle \phi_i^2 \rangle \tag{6}$$

where δ_{ij} is the Kronecker delta and $p(\boldsymbol{\xi})$ is the pdf of $\boldsymbol{\xi}$.

To approximate $\gamma(\boldsymbol{\xi})$ for practical reasons the PCE in Eq.(4) needs to be truncated:

$$\gamma(\boldsymbol{\xi}) = \sum_{0 \leq |\boldsymbol{\alpha}_j| \leq m} a_j \phi_{\boldsymbol{\alpha}_j}(\boldsymbol{\xi}) = \mathbf{a}^T \boldsymbol{\phi}(\boldsymbol{\xi}) \tag{7}$$

where $\boldsymbol{\phi}(\cdot) = [\phi_1(\cdot), \ldots, \phi_L(\cdot)]^T$ contains the multivariate polynomials of the expansion, m denotes the order of truncation, and $|\boldsymbol{\alpha}_j| = \sum_{i=1}^{n_{\boldsymbol{\xi}}} \alpha_{j_i}$. The truncated series consists of $L = \frac{(n_{\boldsymbol{\xi}}+m)!}{n_{\boldsymbol{\xi}}!m!}$ terms and $\mathbf{a} \in \mathbb{R}^L$ represents a vector of coefficients of these terms.

From Eq.(7) we have a PCE representation of γ parametrized by $\boldsymbol{\xi}$. PCE may also represent multivariate random variables as we require for $\mathbf{x}'$. Let a multivariate stochastic variable be given by $\boldsymbol{\gamma}(\boldsymbol{\xi}) = [\gamma_1(\boldsymbol{\xi}), \ldots, \gamma_{n_\gamma}(\boldsymbol{\xi})]^T \in \mathbb{R}^{n_\gamma = n_{\boldsymbol{\xi}}}$ with coefficients collected in $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_{n_\gamma}]$, which is parametrized in terms of standard normal variables $\boldsymbol{\xi}$ with the same dimension. The properties of $\boldsymbol{\gamma}(\boldsymbol{\xi})$ are dependent on the coefficients $\mathbf{A}$ of the expansion. Let each component of $\boldsymbol{\gamma}(\boldsymbol{\xi})$ be given by a truncated PCE with the same order of truncation and the same number of terms L, then the moments of $\boldsymbol{\gamma}(\boldsymbol{\xi})$ have a closed-form expression in terms of the PCE coefficients. The statistical moments of $\boldsymbol{\gamma}$ can be defined as:

$$M_{\mathbf{r}}(\mathbf{A}) = \int \prod_{i=1}^{n_{\boldsymbol{\xi}}} \gamma_i^{r_i}(\boldsymbol{\xi}) p(\boldsymbol{\xi}) d\boldsymbol{\xi} \tag{8}$$

where $\mathbf{r} \in \mathbb{R}^{n_{\boldsymbol{\xi}}}$ is a vector defining the moments with $k = \sum_{i=1}^{n_{\boldsymbol{\xi}}} r_i$ being the overall order.

The moments of the PCE expansion with the definition in Eq.(8) are [29]:

$$M_{\mathbf{r}}(\mathbf{A}) = \int \prod_{i=1}^{n_{\boldsymbol{\xi}}} (\mathbf{a}_i^T \boldsymbol{\phi}(\boldsymbol{\xi}))^{r_i} p(\boldsymbol{\xi}) d\boldsymbol{\xi} \tag{9}$$

IV. PCE STATE ESTIMATOR

The general outline for the PCE filter was taken from [30], [36], however these works do not consider additive disturbance noise. Let $D_t = \{\mathbf{y}_1 \ldots, \mathbf{y}_t\}$ be the measurements collected up to time t and $\mathbf{y}_t$ the most recent measurement. The state estimation step concerns the update of the states given the noisy measurements available. In our algorithm the uncertainties $\mathbf{x}'_t$ are given by PCEs. In particular, let $\mathbf{x}'_{t-1}(\boldsymbol{\xi})$ refer to the previously estimated PCE. Bayes's rule can be employed to update $\mathbf{x}'$ from $\mathbf{x}'_{t-1}|D_{t-1}$ to $\mathbf{x}'_t|D_t$ recursively as follows:

$$p(\mathbf{x}'_t|D_t) = \frac{p(\mathbf{x}'_t|D_{t-1}) p(\mathbf{y}_t|\mathbf{x}'_t, D_{t-1})}{p(\mathbf{y}_t|D_{t-1})} \tag{10}$$

Next we define each term on the RHS of Eq.(10), which are dependent on the dynamic and measurement equation.

1) $p(\mathbf{x}'_t|D_{t-1})$: Prior distribution of $\mathbf{x}'_t$ given the previous measurements D_{t-1}, which can be expressed as:

$$p(\mathbf{x}'_t|D_{t-1}) = \int p(\mathbf{x}'_t|\mathbf{x}'_{t-1}) p(\mathbf{x}'_{t-1}|D_{t-1}) d\mathbf{x}'_{t-1} \tag{11}$$

where $p(\mathbf{x}'_t|\mathbf{x}'_{t-1}) = \mathcal{N}(\mathbf{x}'_t|\mathbf{f}'(\mathbf{x}'_{t-1}, \mathbf{u}_{t-1}), \mathbf{\Sigma_w})$ is a multivariate normal pdf with mean given by the dynamics defined in Eq.(1) and the covariance by the disturbance noise evaluated at $\mathbf{x}_t$. It should be noted that without disturbance noise $p(\mathbf{x}'_t|D_{t-1}) = \int \delta(\mathbf{x}'_t - \mathbf{f}'(\mathbf{x}'_{t-1}, \mathbf{u}_{t-1})) p(\mathbf{x}'_{t-1}|D_{t-1}) d\mathbf{x}'_{t-1}$.

2) $p(\mathbf{y}_t|\mathbf{x}'_t, D_{t-1})$: The pdf of the current measurement $\mathbf{y}_t$ being observed given $\mathbf{x}'_t$, which can be given as follows:

$$p(\mathbf{y}_t|\mathbf{x}'_t, D_{t-1}) = \mathcal{N}(\mathbf{y}_t|\mathbf{h}(\mathbf{x}'_t), \mathbf{\Sigma_v}) \tag{12}$$

3) $p(\mathbf{y}_t|D_{t-1})$: Total probability of observation $\mathbf{y}_t$ given previous measurements can be expressed as:

$$p(\mathbf{y}_t|D_{t-1}) = \int p(\mathbf{y}_t|\mathbf{x}'_t, D_{t-1}) p(\mathbf{x}'_t|D_{t-1}) d\mathbf{x}'_t \tag{13}$$

If we take both sides of Eq.(10) times $\prod_{j=1}^{n_{\xi}}(x'_{tj})^{r_j}$ and integrate over both sides with respect to $\mathbf{x}'_t$ we obtain:

$$M_{\mathbf{r}}^{+} = \frac{\int \prod_{j=1}^{n_{\xi}}(x'_{tj})^{r_j} p(\mathbf{y}_t|\mathbf{x}'_t, D_{t-1}) p(\mathbf{x}'_t|D_{t-1}) d\mathbf{x}'_t}{p(\mathbf{y}_t|D_{t-1})} \quad (14)$$

where from Eq.(10) $M_{\mathbf{r}}^{+} = \int \prod_{j=1}^{n_{\xi}}(x'_{tj})^{r_j} p(\mathbf{x}'_t|D_t) d\mathbf{x}'_t$ and $k = \sum_{j=1}^{n_{\xi}} r_j$. Now $M_{\mathbf{r}}^{+}$ refers to the various k-th order moments of the updated distribution of $\mathbf{x}'_t$, $p(\mathbf{x}'_t|D_t)$.

Next we will deal with the approximation of the RHS of Eq.(14). Evaluating the various multivariate integrals required analytically is difficult and we therefore apply sampling. We are given a PCE expansion at time $t-1$, $\mathbf{x}'_{t-1}(\boldsymbol{\xi})$, corresponding to the pdf $p(\mathbf{x}'_{t-1}|D_{t-1})$ in Eq.(11). This distribution can be readily sampled, since $\boldsymbol{\xi}$ follows a standard normal distribution. Apart from sampling $\boldsymbol{\xi}$, we also require samples of the disturbance $\mathbf{w}$ to deal with the integrals over both $\mathbf{x}_{t-1}$ and $\mathbf{x}_t$. It should be noted that Gauss-Hermite rules were not used for the sampling, since these showed poor convergence due to the high nonlinearity of the likelihood functions. Latin hypercube sampling was applied instead with the inverse normal cumulative transformation to improve the convergence rate over MC [37]:

$$\alpha = \frac{1}{N_s} \sum_{s=1}^{N_s} \mathcal{N}(\mathbf{y}_t|\mathbf{h}(\mathbf{x}'^{(s)}_t), \boldsymbol{\Sigma}_{\mathbf{v}}) \quad (15)$$

where α is the sample estimate of $p(\mathbf{y}(t)|D_{t-1})$, $\mathbf{x}'^{(s)}_t = \mathbf{f}'(\mathbf{x}_{t-1}(\boldsymbol{\xi}^{(s)}), \mathbf{u}_{t-1}) + \mathbf{w}^{(s)}$, N_s is the sample size, $\boldsymbol{\xi}^{(s)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\mathbf{w}^{(s)} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{w}})$ are the sample points.

Now using the sample estimate in Eq.(15) and applying a further sample estimate to Eq.(14) we obtain:

$$M_{\mathbf{r}}^{(s)+} = \frac{\sum_{s=1}^{N_s} \prod_{j=1}^{n_{\xi}} (x'_{tj}(\boldsymbol{\xi}^{(s)}))^{r_j} \mathcal{N}(\mathbf{y}_t|\mathbf{h}(\mathbf{x}'^{(s)}_t), \boldsymbol{\Sigma}_{\mathbf{v}})}{\alpha N_s} \quad (16)$$

where $M_{\mathbf{r}}^{(s)+}$ is an approximation of the RHS of Eq.(14).

To update $\mathbf{x}'_{t-1}(\boldsymbol{\xi})$ we match the moments found in Eq.(16) with those of the PCE $\mathbf{x}'_t(\boldsymbol{\xi})$, which are a function of its coefficients as shown in Eq.(9). The PCE is then fitted by solving a nonlinear least-squares optimization problem:

$$\hat{\mathbf{A}}_t = \arg\min_{\mathbf{A}_t} \sum_{k \leq m} ||M_{\mathbf{r}}^{+}(\mathbf{A}_t) - M_{\mathbf{r}}^{(s)+}||_2^2 \quad (17)$$

where $k = \sum_{j=1}^{n_{\xi}} r_j$ was defined above as the order of the moments and hence m defines the total order of moments we want to match. $M_{\mathbf{r}}^{+}(\mathbf{A}_t)$ is parametrized by $\mathbf{A}_t$ as shown in Eq.(9). The estimated coefficients $\hat{\mathbf{A}}_t$ then define the updated PCE $\mathbf{x}'_t(\boldsymbol{\xi})$ as required for Algorithm 1.

V. SPARSE GAUSS HERMITE SNMPC

In this section we outline a SNMPC formulation to approximately solve the OCP stated in Eq.(3). The problem of controlling a dynamic equation system given by Eq.(1) with initial conditions represented by PCEs and additive time-varying disturbance noise has been addressed in [24]. In this section we propose to use a similar approach with some changes. In [24] the initial conditions and parametric uncertainties are propagated by using PCEs, while the disturbance noise is propagated using linearization. We use a sGH rule instead of PCEs, which is significantly cheaper at estimating the mean and variance of nonlinear transformations. The GH sampling rule is particularly well-suited for uncertainties described by PCEs, since these are parametrized by standard normal distributed variables. In addition, it can be seen in [38] that for the same sample size the accuracy between GH sampling rules and PCE approximations is nearly identical.

In this work we use a sGH quadrature rule proposed in [39] to approximate mean and variance of the constraint and objective functions, which create a deterministic sample of $\boldsymbol{\xi}$ with corresponding weights. The sGH approximations of expectation and variances of a function $q(\boldsymbol{\xi})$ are:

$$\mathbb{E}[q(\boldsymbol{\xi})] \approx \mu_q = \sum_{q=1}^{N_q} w_q q(\boldsymbol{\xi}_q) \quad (18)$$

$$\mathbb{E}\left[(q(\boldsymbol{\xi}) - \mu_q)^2\right] \approx \sigma_q^2 = \sum_{q=1}^{N_q} w_q \left(q(\boldsymbol{\xi}_q) - \mu_q\right)^2 \quad (19)$$

where $\boldsymbol{\xi}_q$ and w_q are given by the sGH quadrature rule with overall N_q points. μ_q and σ_q^2 are the sGH mean and variance approximation respectively.

The above approximation is utilised to account for the contribution of the initial PCE $\mathbf{x}'_t(\boldsymbol{\xi})$, while the time-varying uncertainty $\mathbf{w}$ is accounted for by using linearisation. This is necessary, since $\mathbf{w}$ is time-varying and hence each $\mathbf{w}_t$ is a separate random variable leading otherwise to a too high dimension for sGH. Employing the law of total expectation we can deal with the uncertainties using sGH for $\mathbf{x}'_t(\boldsymbol{\xi})$ and $\mathbf{w}$ using linearization, which can be stated as [31]:

$$\mathbb{E}[q(\mathbf{x}'_t, \mathbf{w})] = \mathbb{E}_{\mathbf{x}'_t}[\mathbb{E}_{\mathbf{w}}[q(\mathbf{x}'_t, \mathbf{w})|\mathbf{x}'_t]] \quad (20)$$

$$\begin{aligned}\mathrm{Var}[q(\mathbf{x}'_t, \mathbf{w})] = \mathbb{E}_{\mathbf{x}'_t}[\mathrm{Var}_{\mathbf{w}}[q(\mathbf{x}'_t, \mathbf{w})|\mathbf{x}'_t]] + \\ \mathrm{Var}_{\mathbf{x}'_t}[\mathbb{E}_{\mathbf{w}}[q(\mathbf{x}'_t, \mathbf{w})|\mathbf{x}'_t]]\end{aligned} \quad (21)$$

By approximating the inner expectation and variance over $\mathbf{w}$ using linearisation, we arrive at:

$$\mathbb{E}[q(\mathbf{x}'_t, \mathbf{w})] \approx \mathbb{E}_{\mathbf{x}'_t}[q(\mathbf{x}'_t, \boldsymbol{\mu}_{\mathbf{w}})] \quad (22)$$

$$\begin{aligned}\mathrm{Var}[q(\mathbf{x}'_t, \mathbf{w})] \approx \mathbb{E}_{\mathbf{x}'_t}[\mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{w}}\mathbf{Q}^T] + \\ \mathrm{Var}_{\mathbf{x}'_t}[q(\mathbf{x}'_t, \boldsymbol{\mu}_{\mathbf{w}})]\end{aligned} \quad (23)$$

where $\boldsymbol{\mu}_{\mathbf{w}}$ denotes the mean of $\mathbf{w}$, $\mathbf{Q} = \frac{\partial q}{\partial \mathbf{w}}|_{\mathbf{x}'_t, \boldsymbol{\mu}_{\mathbf{w}}}$ and $\boldsymbol{\Sigma}_{\mathbf{w}}$ is the covariance of $\mathbf{w}$.

The remaining expectation and variances can then be approximated by creating samples of $\mathbf{x}'_t$ using the sGH rule. Chance constraints are multivariate integrals that are very difficult to estimate online. We use Chebychev‘s inequality to robustly reformulate the chance constraints in terms of only the mean and variance of the constraint function. Let γ be a random variable with a finite variance, then [22]:

$$\kappa_{\epsilon}\sqrt{\sigma_{\gamma}^2} + \hat{\gamma} \leq 0, \ \kappa_{\epsilon} = \sqrt{(1-\epsilon)/\epsilon} \Rightarrow \mathbb{P}(\gamma \leq 0) \geq 1 - \epsilon \quad (24)$$

where $\epsilon \in (0, 1) \subset \mathbb{R}$ is the probability that γ exceeds 0, $\hat{\gamma}$ and σ_{γ}^2 are the mean and variance of γ respectively.

Next we state the finite-horizon stochastic OCP problem that approximates Eq.(3) using the above results. For the SNMPC algorithm we use linearization to account for and propagate the uncertainty of $\mathbf{w}$ similar to an extended Kalman filter based NMPC algorithm [1]. This is carried out for each sample generated by the sGH rule and the overall mean and variance of the objective and constraint function are then found using the law of total expectation. The chance constraints are then further approximated using Chebychev's inequality in Eq.(24). First we create a sGH quadrature sample design with N_q points, $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{N_q}\}$ each with corresponding weights w_q. Assuming we are at time t and are hence given a PCE representation $\mathbf{x}'_t(\boldsymbol{\xi})$, the SNMPC problem can be stated as follows:

$$
\begin{aligned}
&\underset{\mathbf{U}_N}{\text{minimize}} \quad \sum_{i=1}^{N_q} w_q J(N, \mathbf{x}'_t(\boldsymbol{\xi}_i), \mathbf{U}_N) \\
&\text{subject to} \\
&\boldsymbol{\mu}^{(i)}_{\mathbf{x}',k+1} = \mathbf{f}(\boldsymbol{\mu}^{(i)}_{\mathbf{x}',k}, \mathbf{u}_k) && \forall (k,i) \in \mathbb{N}_k \times \mathbb{N}_q \\
&\boldsymbol{\Sigma}^{(i)}_{\mathbf{x}',k+1} = \mathbf{F}^{(i)}_k \boldsymbol{\Sigma}^{(i)}_{\mathbf{x}',k} \mathbf{F}^{(i)T}_k + \boldsymbol{\Sigma}_{\mathbf{w}} && \forall (k,i) \in \mathbb{N}_k \times \mathbb{N}_q \\
&\mu^{(k)}_{g_j} + \kappa_\epsilon \sigma^{(k)}_{g_j} \leq 0 && \forall (k,j) \in \mathbb{N}_{k+1} \times \mathbb{N}_g \\
&\mu_{g^N_j} + \kappa_\epsilon \sigma_{g^N_j} \leq 0 && \forall j \in \mathbb{N}_g \\
&\mathbf{u}_k \in \mathbb{U}_k && \forall k \in \mathbb{N}_k \\
&\boldsymbol{\mu}^{(i)}_{\mathbf{x}',0} = \mathbf{x}'_t(\boldsymbol{\xi}_i) && \forall i \in \mathbb{N}_q
\end{aligned}
\tag{25}
$$

where $\boldsymbol{\mu}^{(i)}_{\mathbf{x}',k}$ and $\boldsymbol{\Sigma}^{(i)}_{\mathbf{x}',k+1}$ represent the mean and covariance of $\mathbf{x}'_k$ for sample i, $\mathbf{F}^{(i)}_k = \frac{\partial \mathbf{f}'}{\partial \mathbf{x}'}|_{\boldsymbol{\mu}^{(i)}_{\mathbf{x}',k}, \mathbf{u}_k}$ denotes the linearised dynamic equation system at time k for sample i, $\mu_{g^{(k)}_j} = \sum_{i=1}^{N_q} w_i g_j(\boldsymbol{\mu}^{(i)}_{\mathbf{x}',k}, \mathbf{u}_k)$ is the mean of the constraint $g^{(k)}_j$, $\sigma^2_{g^{(k)}_j} = \sum_{i=1}^{N_q} w_i \mathbf{G}^{(i)}_{g^{(k)}_j} \boldsymbol{\Sigma}^{(i)}_{\mathbf{x}',k} \mathbf{G}^{(i)T}_{g^{(k)}_j} + \sum_{i=1}^{N_q} w_i \left(g_j(\boldsymbol{\mu}^{(i)}_{\mathbf{x}',k}, \mathbf{u}_k) - \mu^2_{g^{(k)}_j} \right)^2$ is the variance of constraint $g^{(k)}_j$, $\mathbf{G}^{(i)}_{g^{(k)}_j} = \frac{\partial g_j}{\partial \mathbf{x}'}|_{\boldsymbol{\mu}^{(i)}_{\mathbf{x}',k}, \mathbf{u}_k}$ is the Jacobian matrix for the constraint $g^{(k)}_j$ for sample i, $\mu_{g^N_j} = \sum_{i=1}^{N_q} w_i g^N_j(\boldsymbol{\mu}^{(i)}_{\mathbf{x}',N}, \mathbf{u}_N)$ is the mean of constraint g^N_j, $\sigma^2_{g^N_j} = \sum_{i=1}^{N_q} w_i \mathbf{G}^{(i)T}_{g^N_j} \boldsymbol{\Sigma}^{(i)}_{\mathbf{x}',k} \mathbf{G}^{(i)}_{g^N_j} + \sum_{i=1}^{N_q} w_i \left(g_j(\boldsymbol{\mu}^{(i)}_{\mathbf{x}',k}, \mathbf{u}_k) - \mu^2_{g^N_j} \right)^2$ is the variance of constraint g^N_j and lastly $\mathbf{G}^{(i)T}_{g^N_j} = \frac{\partial g^N_j}{\partial \mathbf{x}'}|_{\boldsymbol{\mu}^{(i)}_{\mathbf{x}',N}, \mathbf{u}_N}$ is the Jacobian matrix for constraint g^N_j for sample i.

Solving Eq.(25) at each time t gives the required control inputs for Algorithm 1.

VI. CASE STUDY

The algorithm outlined in section II is employed for the control of a semi-batch polymerization reaction involving the production of polyol from propylene oxide (PO) in a shrinking horizon fashion. A schematic of the process is shown in Fig. 1. A complex model for this process has been proposed in [40], which was employed in [41] for NMPC and in [42] for multi-stage NMPC. The computational times reported in these papers is relatively high at around 30 seconds to minutes. This is due to the model being highly nonlinear and requiring a separate balance equation for each polymer with a specific chain length. To reduce computational times we therefore simplified the model using the method of moments [43] to derive differential equations for the average molecular weight. Further, we disregard the balance equations for the unsaturated proportion of the polymer and assume that there is no water or methanol present in the reactor initially. In the aforementioned work perfect temperature control was assumed. We added a heat balance in this work due to the importance of temperature control for safety reasons. The simplified ordinary equation system can be stated as follows:

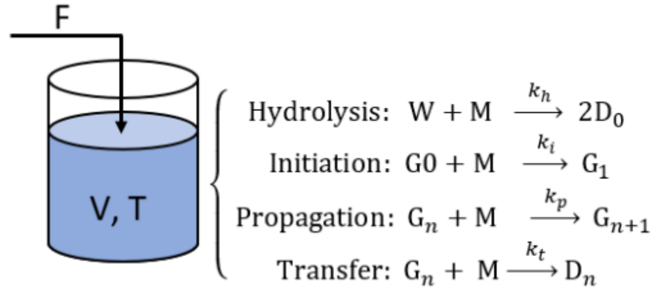

Fig. 1. F is the monomer feedrate, V and T are the volume and temperature of the liquid in the reactor respectively, W is water, M is the monomer, D_n and G_n are the dormant and active product chains with length n respectively.

$$\dot{m} = F MW_{\text{PO}}, \quad m(0) = m_0(\boldsymbol{\xi}) \tag{26a}$$

$$\dot{T} = \frac{(-\Delta H_p) k_p n_C PO}{V m C_{pb}} - \frac{UA(T - T_C)}{m C_{pb}} - \frac{F MW_{\text{PO}} C_{pf}(T - T_f)}{m C_{pb}}, \quad T(0) = T_0(\boldsymbol{\xi}) \tag{26b}$$

$$\dot{PO} = F - \frac{n_C (k_p + k_t) PO}{V}, \quad PO(0) = PO_0(\boldsymbol{\xi}) \tag{26c}$$

$$\dot{\gamma}_1 = \frac{k_p n_C PO}{V}, \quad \gamma_1(0) = \gamma_{1 0}(\boldsymbol{\xi}) \tag{26d}$$

where m[g] is the liquid mass in the reactor, F[mol/s] is the feed rate of the monomer, T[K] is the temperature of the reactor, PO[mol] is the amount of monomer and γ_1 is the first moment and hence the average molecular weight of the polymer chains, $MW_{\text{PO}} = 58.08$g/mol is the molecular weight of PO, ΔH_p is the enthalpy of the propagation reaction, $k_p = A_p \exp(-E_{Ap}/RT)$, n_C is the amount of catalyst, V is the volume of the liquid in the reactor, C_{pb} and C_{pf} are the heat capacities of the bulk liquid and the monomer feed respectively, $k_t = A_t \exp(-E_{At}/RT)$ and T_C[K] is the cooling water temperature. $m_0(\boldsymbol{\xi})$, $T_0(\boldsymbol{\xi})$, $PO_0(\boldsymbol{\xi})$, and $\gamma_{1 0}(\boldsymbol{\xi})$ are the initial PCE expansions of the states m, T, PO, and γ_1 respectively.

The missing parameter values including the vaporliquid equilibrium equations and temperature correlations can be found in [40]. The aim of the control algorithm is to minimize the required remaining batch time (t_f[s]) to achieve a specified number average molecular weight

(NAMW[g/mol]) of the final product of 350g/mol and ensuring that the amount of monomer (PO) at the end of the batch does not exceed 1000ppm. The definitions NAMW can be found in [40], which requires the zeroth moment γ_0[mol]. Due the assumptions made this zeroth moment is a constant in the defined problem and given in Tab. I. For safety reasons the reactor temperature T[K] is constrained to remain below 420K. The control variables are the monomer feed rate $F[\text{mol/s}]$ to the reactor and the cooling water temperature T_C[K]. The chance of constraint violation was set to 0.05. Apart from the uncertainty of the states, the variables pre-exponential coefficient of the propagation kinetic constant ($A_p[\text{m}^3/\text{mol/s}]$) and the heat transfer coefficient (UA[W/K]) were assumed to be uncertain and added to the state vector as shown in Eq.(1). Measurements during the reaction are the pressure (P[bar]) of the reactor, the temperature (T[K]) of the reactor, and the amount of monomer (PO[mol]). The discretization of Eq.(26) was carried-out utilising orthogonal collocation. The resulting optimization problems for both the sGH SNMPC problem and the PCE state estimator were solved using Casadi [44] in conjunction with IPOPT [45]. The control problem to be solved is specified in Tab. I.

TABLE I
SPECIFICATIONS OF CONTROL PROBLEM

Augmented state ($\mathbf{x}'$)	m[g], PO[mol], T[K], γ_1[mol]
Disturbance noise	$\mathbf{\Sigma_w} = \text{diag}(0, 0, 1, 2, 50, 200)$
Outputs ($\mathbf{y}$)	P[bar], T[K], PO[mol]
Output noise	$\mathbf{\Sigma_v} = \text{diag}(0.25, 0.001, 1000)$
Inputs (u)	F[mol], T_C[K]
Uncertainties	$A_p[\text{m}^3/\text{mol/s}]$, n_C[mol]
Objective	minimize t_f[s]
Path constraints	$T[K] - 420 \leq 0$
1st end constraint	$350 - \text{NAMW[g/mol]} \leq 0$
2nd end constraint	$\text{PO[ppm]} - 1000 \leq 0$
Probability	$\epsilon = 0.05$
Input constraints	$0 \leq F[\text{mol/s}] \leq 10$, $298.15 \leq T_C[\text{K}]$
sGH SNMPC	sGH accuracy = 2, sGH manner = 1
$\mathbf{x}'$-PCE	PCE order = 2
PCE filter	Samples = 4000, Moments considered = 4
Discretization	N = 8, Degree = 5
Initial PCE $m_0(\boldsymbol{\xi})$	1537710
Initial PCE $PO_0(\boldsymbol{\xi})$	10000 + 1000ξ_2
Initial PCE $T_0(\boldsymbol{\xi})$	378.15 + 4ξ_3
Initial PCE $\gamma_{1\,0}(\boldsymbol{\xi})$	10000 + 500ξ_4
Initial PCE $A_{P\,0}(\boldsymbol{\xi})$	8504 + 1000ξ_5
Initial PCE $UA_0(\boldsymbol{\xi})$	40000 + 4000ξ_6
Reactor specs.	V_R = 17m^3, $n_C = 1000$mol, $\gamma_0 = 10000$mol

VII. RESULTS AND DISCUSSIONS

In this section we verify Algorithm 1 by applying it to the previously specified control problem. First we run the algorithm on a specific realization of the initial uncertainties on $\mathbf{x}'$ as specified by the initial PCE in Tab. I, which is as follows: $\mathbf{x}' = [1537710, 12371, 375.2, 9855, 10100, 36301]^T$. The disturbances $\mathbf{w}$ and measurement noise $\mathbf{v}$ are randomly sampled. The results of this run are shown in Fig. 2.

The two graphs in the first row of Fig. 2 show the evolution of the pdf of the two uncertain parameters A_p and UA from its PCE. Firstly, it can be seen that both parameters are significantly better approximated at the final time than initially through the measurement updates. In particular, the distribution of A_p starts out at around 7500, but thereafter rapidly approaches its true value shown by the vertical black line at around 10000. Nonetheless some bias remains towards a lower value, however with some probability to take its true value or higher. UA on the other hand does not have such a bias and converges quickly to its true value, but the distribution remains relatively broad due to influence of the disturbance and measurement noise. The next two rows show the trajectories of the 4 states as a continuous blue line, while the black crosses and error bars represent the state estimates with a 95% confidence interval of the PCE expansion. The mass and monomer have near exact state estimates due to the assumed low disturbance noise. The monomer has a slight deviation initially but quickly converges to the true trajectory once the first measurement becomes available. The ppm of the monomer at the final time was found to be 916 and hence 84 less than required. This can be explained by the underestimation of A_p, which means the SNMPC will run longer to ensure a sufficiently low ppm is reached. Temperature has some significant uncertainty, which is accounted for in the SNMPC algorithm to not violate the constraint to remain below 420K. Lastly, the first moment over-shoots the constraint required to reach 350mol/g NAMW and instead reaches a NAMW of 385mol/g. This can again be explained by the underestimation of A_p and hence increasing the batch time to reach the required NAMW and ppm in at least 95% of possible cases. This can also be seen in the graph in the last row, where at first the sampling time decreases due to less uncertainty but then increases again to ensure the end-point constraints. In the 4th row the control inputs are shown, which are as expected. First the feed rate of the monomer is set to the maximum before it is set to zero to ensure a low ppm a the end of the batch process. The reaction rate is highest initially at which point the cooling temperature is at its lower bound and after which only moderate cooling temperatures are required to prevent constraint violations of the reactor temperature.

Next we ran 100 MC simulations of the control problem. This is compared to 100 MC simulations of a nominal NMPC algorithm using the mean value of the PCE expansion as state estimate from the PCE state estimator, but ignoring the shape of the probability distribution otherwise. This is done to show the importance of accounting for the inherent uncertainty in the problem. The results of these MC simulations are highlighted in Fig. 3. The first graph illustrates that while the nominal NMPC method manages in only 49% of cases to realize the required NAMW, the SNMPC approach due its increased conservativeness manages to fulfil the NAMW constraint in 99% of the simulations. Similarly the second graphs shows that the SNMPC approach manages to reduce the amount of monomer below the required threshold of 1000 ppm in 97% of realizations, while the nominal NMPC adheres this end-point constraint in only 52% of the simulations. The next graph shows that this increased robustness comes at the price of on-average longer batch times. The nominal NMPC took on average 4800 seconds, while the

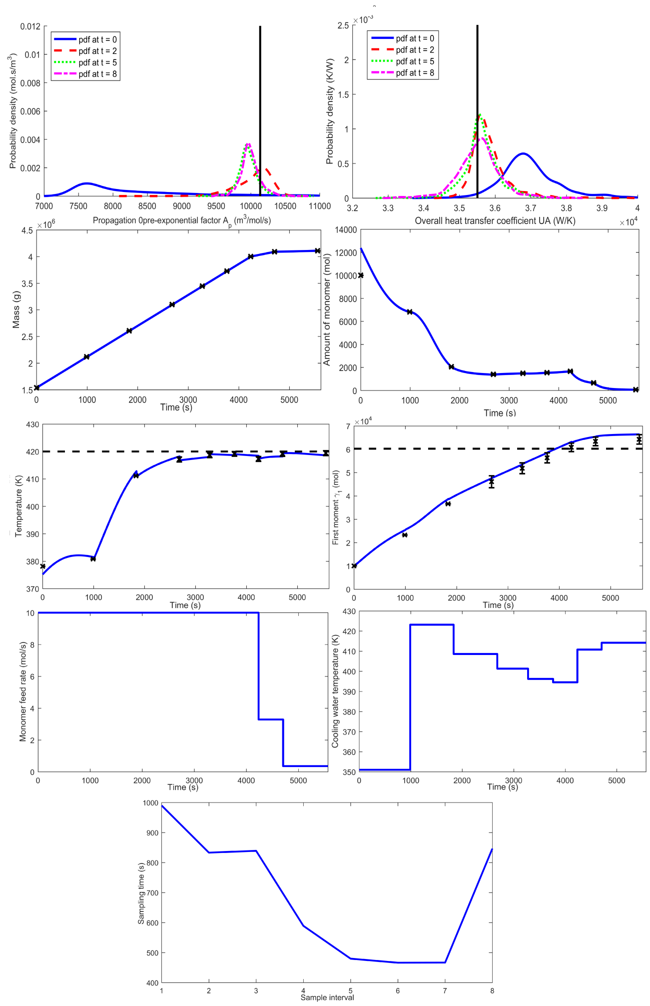

Fig. 2. First row shows the pdfs at $t = 0, 2, 5, 8$ for A_p and UA, 2nd and 3rd row show the trajectories of the states with the corresponding state estimates as black-crosses with 95% confidence intervals, 4th row illustrates the control inputs and the last row the changes in the sampling time.

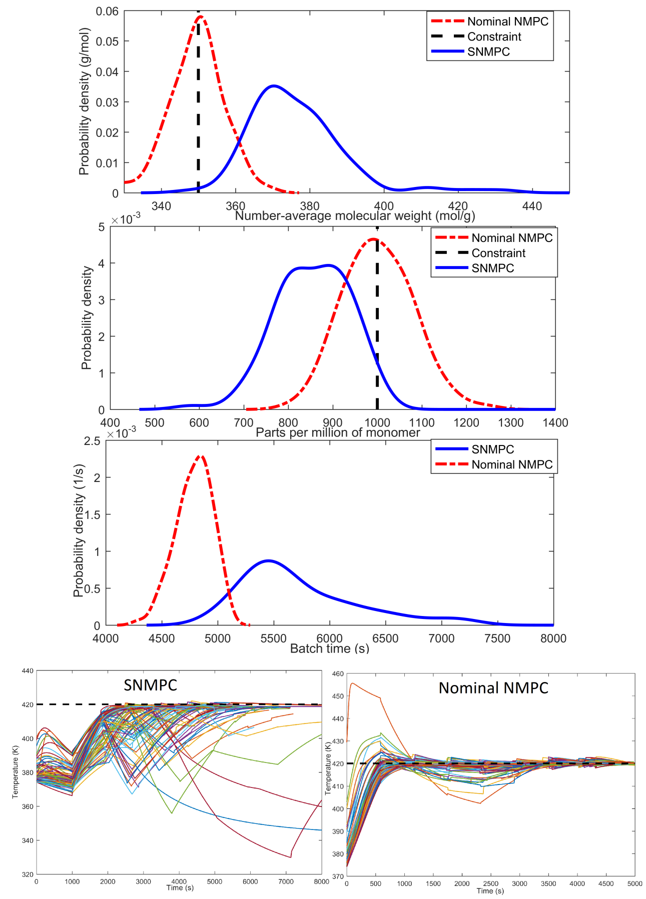

Fig. 3. Probability densities of NAMW, ppm of monomer, and batch time at final time and temperature trajectories of SNMPC and nominal NMPC based on 100 MC simulations

SNMPC had average batch times of 9300 seconds. The final graphs illustrate significantly better temperature control of the SNMPC approach compared to the nominal NMPC. In particular at the beginning ignoring the uncertainty of the overall heat transfer coefficient UA leads to constraint violations of up to 30K, while later on not accounting for the disturbance also leads to constraint violations for nominal NMPC method. The SNMPC method on the other hand shows close to no constraint violations.

VIII. CONCLUSIONS

In conclusion, we have presented a new algorithm for output feedback SNMPC that utilises a PCE nonlinear state estimator to approximate the probability distribution of states and uncertain parameters at each sampling time. This PCE representation is then exploited in a SNMPC formulation to account for both the initial value uncertainty using a sGH sampling rule and additive disturbance noise employing linearisation. Objectives and constraints were based on general nonlinear functions. A semi-batch reactor case study verified that the SNMPC framework is able to control the system with large initial uncertainties and additive disturbance noise. The PCE nonlinear state estimator is shown to able to accurately update the distribution of the states and uncertain parameters, while considering the uncertainty information from the distribution to avoid constraint violations. Furthermore, it was shown that ignoring the uncertainty informations leads to 50% constraint violations of the end-point constraints and large overshoots of the temperature path-constraint.

ACKNOWLEDGMENTS

Eric Bradford gratefully appreciates BASF for hosting his placement over the course of the Marie-Curie project and the helpful input and discussions for this work. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 675215.

REFERENCES

[1] Z. K. Nagy and R. D. Braatz, "Robust nonlinear model predictive control of batch processes," *AIChE Journal*, vol. 49, no. 7, pp. 1776–1786, 2003.

[2] J. M. Maciejowski, *Predictive control: with constraints*. Pearson education, 2002.

[3] A. Bemporad and M. Morari, "Robust model predictive control: A survey," in *Robustness in identification and control*. Springer, 1999, pp. 207–226.

[4] H. Chen, C. W. Scherer, and F. Allgower, "A game theoretic approach to nonlinear robust receding horizon control of constrained systems," in *Proceedings of the 1997 American Control Conference*, vol. 5. IEEE, 1997, pp. 3073–3077.

[5] D. Q. Mayne, E. C. Kerrigan, E. J. Van Wyk, and P. Falugi, "Tubebased robust nonlinear model predictive control," *International Journal of Robust and Nonlinear Control*, vol. 21, no. 11, pp. 1341–1353, 2011.

[6] A. Mesbah, "Stochastic model predictive control: An overview and perspectives for future research," *IEEE Control Systems*, vol. 36, no. 6, pp. 30–44, 2016.

[7] M. Farina, L. Giulioni, and R. Scattolini, "Stochastic linear model predictive control with chance constraintsa review," *Journal of Process Control*, vol. 44, pp. 53–67, 2016.

[8] M. Cannon, B. Kouvaritakis, S. V. Rakovic, and Q. Cheng, "Stochastic tubes in model predictive control with probabilistic constraints," *IEEE Transactions on Automatic Control*, vol. 56, no. 1, pp. 194–200, 2011.

[9] G. Schildbach, L. Fagiano, C. Frei, and M. Morari, "The scenario approach for stochastic model predictive control with bounds on closed-loop constraint violations," *Automatica*, vol. 50, no. 12, pp. 3009–3018, 2014.

[10] D. H. Van Hessem and O. H. Bosgra, "A conic reformulation of model predictive control including bounded and stochastic disturbances under state and input constraints," in *Proceedings of the 41st IEEE Conference on Decision and Control*, vol. 4. IEEE, 2002, pp. 4643–4648.

[11] E. Bradford and L. Imsland, "Economic Stochastic Model Predictive Control Using the Unscented Kalman Filter," *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 417–422, 2018. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2405896318320196

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

[13] E. Bradford and L. Imsland, "Expectation constrained stochastic nonlinear model predictive control of a batch bioreactor," *Computer Aided Chemical Engineering*, vol. 40, pp. 1621–1626, 2017.

[14] J. M. Maciejowski, A. L. Visintini, and J. Lygeros, "NMPC for complex stochastic systems using a Markov chain Monte Carlo approach," in *Assessment and Future Directions of Nonlinear Model Predictive Control*. Springer, 2007, pp. 269–281.

[15] E. Bradford and L. Imsland, "Stochastic Nonlinear Model Predictive Control Using Gaussian Processes," in *2018 European Control Conference (ECC)*. IEEE, 2018, pp. 1027–1034.

[16] F. Weissel, M. F. Huber, and U. D. Hanebeck, "Stochastic nonlinear model predictive control based on Gaussian mixture approximations," in *Informatics in Control, Automation and Robotics*. Springer, 2009, pp. 239–252.

[17] E. A. Buehler, J. A. Paulson, and A. Mesbah, "Lyapunov-based stochastic nonlinear model predictive control: Shaping the state probability distribution functions," in *2016 American Control Conference (ACC)*. IEEE, 2016, pp. 5389–5394.

[18] M. Cannon, D. Ng, and B. Kouvaritakis, "Successive linearization NMPC for a class of stochastic nonlinear systems," in *Nonlinear Model Predictive Control*. Springer, 2009, pp. 249–262.

[19] M. A. Sehr and R. R. Bitmead, "Particle model predictive control: Tractable stochastic nonlinear output-feedback MPC," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 15 361–15 366, 2017.

[20] P. Patrinos, P. Sopasakis, H. Sarimveis, and A. Bemporad, "Stochastic model predictive control for constrained discrete-time Markovian switching systems," *Automatica*, vol. 50, no. 10, pp. 2504–2514, 2014.

[21] T. Homer and P. Mhaskar, "Output-Feedback Lyapunov-Based Predictive Control of Stochastic Nonlinear Systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 2, pp. 571–577, 2018.

[22] A. Mesbah, S. Streif, R. Findeisen, and R. D. Braatz, "Stochastic nonlinear model predictive control with probabilistic constraints," in *2014 American Control Conference*. IEEE, 2014, pp. 2413–2419.

[23] S. Streif, M. Karl, and A. Mesbah, "Stochastic nonlinear model predictive control with efficient sample approximation of chance constraints," *arXiv preprint arXiv:1410.4535*, 2014.

[24] V. A. Bavdekar and A. Mesbah, "Stochastic nonlinear model predictive control with joint chance constraints," *IFAC-PapersOnLine*, vol. 49, no. 18, pp. 270–275, 2016.

[25] J. A. Paulson and A. Mesbah, "An efficient method for stochastic optimal control with joint chance constraints for nonlinear systems," *International Journal of Robust and Nonlinear Control*, 2017.

[26] J. Li and D. Xiu, "A generalized polynomial chaos based ensemble Kalman filter with high accuracy," *Journal of computational physics*, vol. 228, no. 15, pp. 5454–5469, 2009.

[27] E. D. Blanchard, A. Sandu, and C. Sandu, "A polynomial chaos-based Kalman filter approach for parameter estimation of mechanical systems," *Journal of Dynamic Systems, Measurement, and Control*, vol. 132, no. 6, p. 61404, 2010.

[28] P. Dutta and R. Bhattacharya, "Nonlinear estimation of hypersonic state trajectories in Bayesian framework with polynomial chaos," *Journal of guidance, control, and dynamics*, vol. 33, no. 6, pp. 1765–1778, 2010.

[29] ——, "Nonlinear estimation with polynomial chaos and higher order moment updates," in *American Control Conference (ACC), 2010*. IEEE, 2010, pp. 3142–3147.

[30] R. Madankan, P. Singla, T. Singh, and P. D. Scott, "Polynomial-chaos-based Bayesian approach for state and parameter estimations," *Journal of Guidance, Control, and Dynamics*, vol. 36, no. 4, pp. 1058–1074, 2013.

[31] V. A. Bavdekar and A. Mesbah, "A polynomial chaos-based nonlinear Bayesian approach for estimating state and parameter probability distribution functions," in *American Control Conference (ACC), 2016*. IEEE, 2016, pp. 2047–2052.

[32] R. Pandurangan, A. Chaudhuri, and S. Gupta, "The use of polynomial chaos for parameter identification from measurements in nonlinear dynamical systems," *ZAMMJournal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 95, no. 12, pp. 1372–1392, 2015.

[33] E. Bradford, M. Reble, A. Bouaswaig, and L. Imsland, "Economic stochastic nonlinear model predictive control of a semi-batch polymerization reaction," in *Dynamics and Control of Process Systems, including Biosystems - 12th DYCOPS*. IFAC, 2019, p. submitted.

[34] D. Xiu and G. E. Karniadakis, "Modeling uncertainty in flow simulations via generalized polynomial chaos," *Journal of computational physics*, vol. 187, no. 1, pp. 137–167, 2003.

[35] M. Eldred and J. Burkardt, "Comparison of Non-Intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Quantification," in *47th AIAA Aerospace Sciences Meeting including The New Horizons Forum and Aerospace Exposition*, 2009, p. 976. [Online]. Available: http://arc.aiaa.org/doi/10.2514/6.2009-976

[36] T. Mühlpfordt, J. A. Paulson, R. D. Braatz, and R. Findeisen, "Output feedback model predictive control with probabilistic uncertainties for linear systems," in *American Control Conference (ACC), 2016*. IEEE, 2016, pp. 2035–2040.

[37] M. Stein, "Large sample properties of simulations using Latin hypercube sampling," *Technometrics*, vol. 29, no. 2, pp. 143–151, 1987.

[38] J. H. Panchal, S. R. Kalidindi, and D. L. McDowell, "Key computational modeling issues in integrated computational materials engineering," *Computer-Aided Design*, vol. 45, no. 1, pp. 4–25, 2013.

[39] B. Jia, M. Xin, and Y. Cheng, "Sparse-grid quadrature nonlinear filtering," *Automatica*, vol. 48, no. 2, pp. 327–341, 2012.

[40] Y. Nie, L. T. Biegler, C. M. Villa, and J. M. Wassick, "Reactor modeling and recipe optimization of polyether polyol processes: Polypropylene glycol," *AIChE Journal*, vol. 59, no. 7, pp. 2515–2529, 2013.

[41] T. Y. Jung, Y. Nie, J. H. Lee, and L. T. Biegler, "Model-based on-line optimization framework for semi-batch polymerization reactors," *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 164–169, 2015.

[42] H. Jang, J. H. Lee, and L. T. Biegler, "A robust NMPC scheme for semi-batch polymerization reactors," *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 37–42, 2016.

[43] Y. Nie, L. T. Biegler, C. M. Villa, and J. M. Wassick, "Reactor modeling and recipe optimization of ring-opening polymerization: Block copolymers," *Industrial & Engineering Chemistry Research*, vol. 53, no. 18, pp. 7434–7446, 2013.

[44] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADi: a software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, pp. 1–36, 2018.

[45] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.