Available online at www.sciencedirect.com
ScienceDirect
Journal of Process Control 17 (2007) 229–240
JOURNAL OF PROCESS CONTROL
www.elsevier.com/locate/jprocont

Distributional uncertainty analysis using power series and polynomial chaos expansions

Z.K. Nagy [a,*], R.D. Braatz [b]

[a] Loughborough University, Department of Chemical Engineering, Loughborough, Leics LE11 3TU, United Kingdom
[b] Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 6180, United States

Received 23 July 2006; accepted 3 October 2006

Abstract

This paper provides an overview of computationally efficient approaches for quantifying the influence of parameter uncertainties on the states and outputs of nonlinear dynamical systems with finite-time control trajectories, focusing primarily on computing probability distributions. The advantages and disadvantages of various uncertainty analysis approaches, which use approximate representations of the full nonlinear model using power series or polynomial chaos expansions, are discussed in terms of computational cost and accuracy in computing the shape and tails of the state and output distributions. Application of the uncertainty analysis methods to a simulation study is used to provide advice as to which uncertainty analysis methods to select for a particular application. In particular, the results indicate that first-order series analysis can be accurate enough for the design of real-time robust feedback controllers for batch processes, although it is cautioned that the accuracy of such analysis should be confirmed *a posteriori* using a more accurate uncertainty analysis method. The polynomial chaos expansion is well suited to robust design and control when the objectives are strongly dependent on the shape or tails of the distributions of product quality or economic objectives.

Keywords: Distributional robustness analysis; Polynomial chaos expansion; Robust control; Nonlinear control; Optimal control

1. Introduction

A comprehensive uncertainty analysis of mechanistic models is crucial, especially when these models are used in the optimal control of processes, which normally occurs close to safety and performance constraints. The model-based computation of optimal control policies for batch and semibatch processes is of increasing importance due to the industrial interest in increasing productivity in the specialty chemicals, polymers, pharmaceuticals, and other industries [1–3]. However, uncertainty almost always exists in chemical systems and its disregard can lead to the loss of all of the benefits of using optimal control [4]. This motivates the development of techniques to quantify the influence of parameter uncertainties on the process states and outputs, or take uncertainties into account in the design of processes or control systems [5]. There is an increasing need to develop generic formalisms that facilitate the quantification of uncertainty and its effect on model-based predictions. Quantitative estimates obtained from robustness analysis can be used to evaluate the probability of extreme events such as failures, to estimate the confidence in model predictions, or to design efficient schemes for model or data refinement, for example by deciding whether more laboratory experiments are needed to provide better parameter [6,7]. Uncertainty propagation has a key role in robust optimization where appropriate back-off terms, which guarantee that important constraints, such as those that arise from safety or environmental considerations, are satisfied under uncertain conditions.

* Corresponding author. Tel.: +44 1509 222516; fax: +44 1509 223923.
E-mail address: z.k.nagy@lboro.ac.uk (Z.K. Nagy).

0959-1524/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2006.10.008

Uncertainty analysis consists of the characterization of uncertainty in model parameters or inputs based on their probability density functions (pdfs) and then propagating these pdfs through the model equations to obtain the pdfs of the model outputs. A number of techniques developed by various communities (including climate modelling, hurricane prediction, and batch process control) are available for uncertainty propagation. The propagation of uncertainties via traditional Monte Carlo methods based on standard or Latin Hypercube sampling is valid for a wide range of problems, however, it is very computationally expensive since it requires a large number of simulations using the full model. Although this caveat is becoming less significant for analysis purposes, it still can be prohibitive if the propagation has to be performed in real-time, such as in robust optimizing control. Another more intrinsic disadvantage of the classical Monte Carlo method is that it does not provide a manageable representation of the predicted process. Therefore there is a need to investigate computationally-efficient alternative techniques for uncertainty propagation, which also provide a simplified mathematical representation of the process. One alternative approach to the Monte Carlo procedure is the perturbation method, or sensitivity analysis [8]. This method represents the prediction as a perturbation around its nominal value, usually associated with the mean or expected value of the uncertain deterministic problem. Sensitivity analysis usually provides reliable predictions only when the associated perturbations are small, and it is not used to predict the shape of the whole distribution. An alternative way of considering the uncertainty propagation problem is based on probability theory, and consists of representing the random variables and model in terms of convergent expansions according to a framework, which is similar to the deterministic approximation approaches.

The paper investigates computationally-efficient methods for estimating the entire distribution of states and outputs of generic batch processes. One class of methods is based on first- and second-order power series [9] and can be considered as a combination of the sensitivity analysis and probability theory techniques. The first-order method is a very computationally efficient way to propagate uncertainties but does not capture the shape and tails of nonGaussian distributions; for strongly nonlinear processes the number of terms in the power series expansion has to be increased, requiring higher computational cost. An alternative approach, belonging to the category of probability theory methods, is based on the stochastic response surface approach (SRS), which models the performance function/model output, as a sum of elementary functions (bases) of stochastic input parameters. The efficient application of the SRS approach requires the proper choice of the elementary functions and the number and location of the sampling points. This paper will consider a variant of the SRS approach based on the polynomial chaos expansion (PCE) [10–13], which uses Hermite polynomial bases and provides a closed-form functional approximation of the mathematical model from a significantly lower number of model simulations than those required by conventional methods. The power series and polynomial chaos expansion are suitable for studying the uncertainty propagation in open-loop or closed-loop systems, and can be applied for uncertainties in model parameters, initial conditions, or inputs. The techniques are compared with the Monte Carlo method and applied to compute the distributions of the states and outputs for the batch crystallization of an inorganic chemical subject to uncertainties in the nucleation and growth kinetics. The simulation study is used to draw some more general recommendations about which uncertainty analysis methods to select for a particular application.

2. Distributional uncertainty analysis

2.1. Uncertainty description

Consider the class of finite time (batch) processes described by the generic ODE:

$$\dot{x}(t) = f(x(t), u(t);\ \theta) \tag{1}$$

with the state vector $x \in \mathbb{R}^{n_x}$ with initial values x_0, the vector of control inputs $u \in \mathbb{R}^{n_u}$, the uncertain parameter vector $\theta \in \mathbb{R}^{n_\theta}$, and the vector function f which is continuous with respect to its elements. This paper considers the quantification of the effect of parametric uncertainties on the accuracy of the model predictions, which can be used to assess whether more experimental data are needed to further refine the parameters. It is assumed that model development is detailed enough to capture the main structural characteristics of the system, so that uncertainties in the model structure can be neglected.

Uncertainty quantification can be decomposed in three main steps: (i) characterisation of the parameter uncertainty, (ii) propagation of uncertainty, and (iii) uncertainty management or evaluation and exploitation of the information obtained. The uncertainties in the parameters can be represented in terms of standard normal variables or hard bounds on the parameters. Define $\hat{\theta}$ as the nominal model parameter vector of dimension $(n \times 1)$, $\delta\theta$ as the perturbation about $\hat{\theta}$, and the model parameter vector for the real system as

$$\theta = \hat{\theta} + \delta\theta. \tag{2}$$

Hard bounds on the uncertainty in the parameter is characterized by the generalized ball

$$\mathscr{P} \triangleq \left\{ \theta \in \mathbb{R}^{n_\theta} \middle| \left\| \theta - \hat{\theta} \right\| \leqslant 1 \right\} \tag{3}$$

defined by using an appropriate norm $\|\cdot\|$ in $\mathbb{R}^{n_\theta}$ such as a scaled Hölder p-norm $(1 \leqslant p \leqslant \infty)$, given by $\|\theta\| = \|\mathbf{W}_\theta^{-1}\theta\|_p$, with the invertible weighting matrix $\mathbf{W}_\theta \in \mathbb{R}^{n_\theta \times n_\theta}$. This generalized description of the uncertainty set includes the case of a confidence hyper-ellipsoid $\varepsilon_\theta = \left\{ \theta : \left(\theta - \hat{\theta}\right)^{\mathrm{T}} \mathbf{V}_\theta^{-1} \left(\theta - \hat{\theta}\right) \leqslant r^2(\alpha) \right\}$, [14], for a Gaussian random variable vector θ with expected value $\varepsilon(\theta) = \hat{\theta}$, the $(n_\theta \times n_\theta)$ positive-definite variance–covariance matrix

$\mathbf{V}_\theta$, and the scalar r which is the chi-squared distribution function with n_θ degrees of freedom ($\chi^2_{n_\theta}(\alpha)$) for a chosen confidence level α:

$$\mathscr{P}_{\text{ellipsoid}}(\alpha) \triangleq \left\{\theta \in \mathbb{R}^{n_\theta} \middle| \left\|(1/r(\alpha))\mathbf{V}_\theta^{-1/2}\left(\theta - \hat{\theta}\right)\right\|_2 \leqslant 1\right\}. \tag{4}$$

The generalized ball described by (2) also includes the case of known lower and upper bounds θ_l, θ_u on the uncertain parameters, leading to an uncertainty hyper-box:

$$\begin{aligned}\mathscr{P}_{\text{box}} &\triangleq \{\theta \in \mathbb{R}^{n_\theta} | \theta_l \leqslant \theta \leqslant \theta_u\} \\ &= \left\{\theta \in \mathbb{R}^{n_\theta} \middle| \left\|\frac{1}{2}\text{diag}(\theta_u - \theta_l)\left(\theta - \frac{1}{2}(\theta_u + \theta_l)\right)\right\|_\infty \leqslant 1\right\}.\end{aligned} \tag{5}$$

Although the focus on this manuscript is on uncertainties in parameters, any generic method for analyzing uncertainties in parameters can be extended, with minor modifications, to analyze the effects of simultaneous uncertainties in parameters, initial conditions, and inputs (as discussed by [6]).

2.2. Power series expansion based approaches for worst-case and distributional robustness analysis

Define $\hat{\psi}$ as a state or output of a process with the nominal model parameters $\hat{\theta}$, ψ as its value for the perturbed model parameter vector θ, and the difference $\delta\psi = \psi - \hat{\psi}$. The worst-case robustness approach [4,9] writes $\delta\psi$ as a power series in $\delta\theta$:

$$\delta\psi = L\delta\theta + \frac{1}{2}\delta\theta^{\mathrm{T}}\mathbf{M}\delta\theta + \cdots, \tag{6}$$

where the jacobian $L \in \mathbb{R}^{n_\theta}$, and hessian $\mathbf{M} \in \mathbb{R}^{n_\theta \times n_\theta}$ are:

$$L(t) = \left(\frac{\partial\psi(t)}{\partial\theta}\right)_{\hat{\theta}}, \tag{7}$$

$$\mathbf{M}(t) = \left(\frac{\partial^2\psi(t)}{\partial\theta^2}\right)_{\hat{\theta}}. \tag{8}$$

The elements of the time-varying sensitivity vector $L(t)$ and matrix $\mathbf{M}(t)$ can be computed using finite differences or by integrating the model's differential–algebraic equations augmented with an additional set of differential equations known as sensitivity equations [15]

$$\dot{L} = \mathbf{J}_x L + \mathbf{J}_\theta \tag{9}$$

with the matrixes $\mathbf{J}_x = \mathrm{d}f/\mathrm{d}x \in \mathbb{R}^{n_2 \times n_x}$ and $\mathbf{J}_\theta = \mathrm{d}f/\mathrm{d}\theta \in \mathbb{R}^{n_x \times n_\theta}$.

When a first-order series expansion is used, analytical expressions of the worst-case deviation in the performance index ($\delta\psi_{\text{w.c.}}$) can be computed and the analysis can be performed with low computational cost [16]. In the case of an ellipsoidal uncertainty description, the worst-case deviation is defined by

$$\partial\psi_{\text{w.c.}}(t) = \max_{\|\mathbf{W}_\theta\delta\theta\| \leqslant 1} |L(t)\delta\theta| \tag{10}$$

which has an analytical solution

$$\partial\psi_{\text{w.c.}}(t) = \left(r(\alpha)L(t)\mathbf{V}_\theta L^{\mathrm{T}}(t)\right)^{1/2}. \tag{11}$$

$$\partial\theta_{\text{w.c.}}(t) = \frac{(r(\alpha))^{1/2}}{\left(L(t)\mathbf{V}_\theta L^{\mathrm{T}}(t)\right)^{1/2}}\mathbf{V}_\theta L^{\mathrm{T}}(t). \tag{12}$$

A probability density function (pdf) for the model parameters is needed to compute the pdf of the performance index. More than 90% of the available algorithms to estimate parameters from experimental data [14] produce a multivariate normal distribution:

$$f_{\text{p.d.}}(\theta) = \frac{1}{(2\pi)^{n_\theta/2}\det(\mathbf{V}_\theta)^{1/2}}\exp\left(\frac{1}{2}\left[\left(\theta - \hat{\theta}\right)^{\mathrm{T}}\mathbf{V}_\theta^{-1}\left(\theta - \hat{\theta}\right)\right]\right). \tag{13}$$

When a first-order series expansion is used to relate $\delta\psi$ and $\delta\theta$, then the estimated pdf of ψ is

$$f_{\text{p.d.}}(\psi) = \frac{1}{V_\psi^{1/2}\sqrt{2\pi}}\exp\left(-\left(\psi - \hat{\psi}\right)^2/(2V_\psi)\right), \tag{14}$$

where the variance of ψ is

$$V_\psi = L\mathbf{V}_\theta L^{\mathrm{T}}. \tag{15}$$

The distribution is a function of time since the nominal value for ψ and the vector of sensitivities L are functions of time.

2.2.1. Propagation of probability distribution using the Monte Carlo method or by contour mapping

In the Monte Carlo method for uncertainty analysis, a large number of parameters are generated based on the multidimensional probability distribution function of the parameter space and are propagated through the nonlinear model, and the probability distribution functions of the performance index or model states/outputs are constructed from the histograms resulted from the nonlinear dynamic simulations. This approach is illustrated schematically in Fig. 1 for a system with three uncertain parameters in which the uncertainty is given by a multivariate normal distribution with confidence hyper-ellipsoid shown. Points are taken randomly form the hyper-ellipsoid and propagated through the nonlinear model, which can be any system of ordinary differential equations (ODE), differential algebraic equations (DAE), partial differential equations (PDE), or the combination of such equations. The main disadvantage of the classical Monte Carlo techniques is that an accurate construction of the output distribution requires a very large number of parameter samples and computer-intensive simulations, to accurately model the tails of the distribution, which correspond to parameters with low probability of occurrence. In the case illustrated in Fig. 1, although 10,000 sets of parameters were simulation, a significant portion of the parameter ellipsoid was not well sampled, leading to errors in representing the tails of the output distribution. Although these parameter

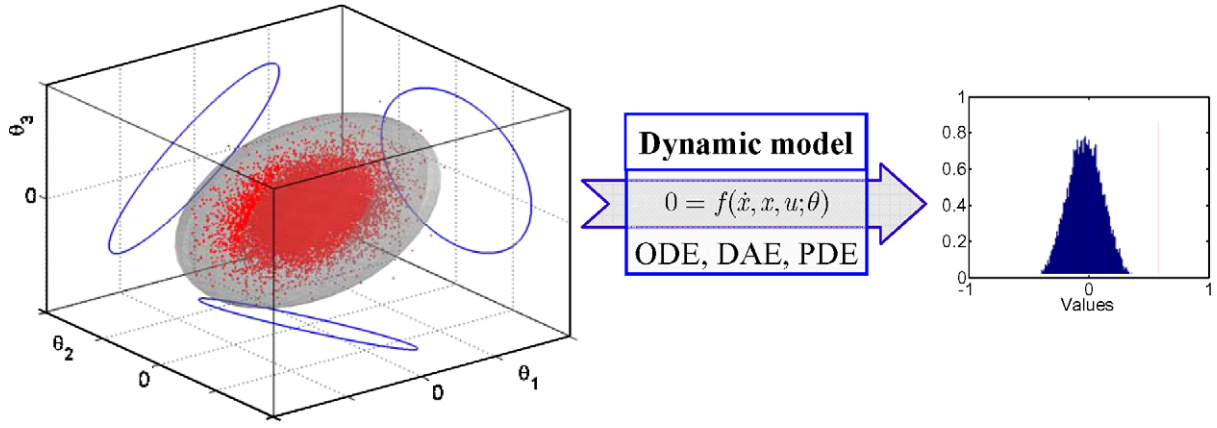

Fig. 1. Schematic representation of the uncertainty propagation through the full nonlinear model using the Monte Carlo method.

values have a low probability to occur, such parameters may have significant effects on the performance index, with large economic or safety consequences if their effect is not taken into consideration during process or controller design based on a model with uncertainties.

An approach that has much lower computational expense than performing exhaustive Monte Carlo simulations is to map only the contours of the uncertainty hyper-ellipsoid obtained for different α levels, from the solution of the worst-case analysis problem (10), based on the power series expansion representation of the model. This approach is illustrated schematically in Fig. 2 for a case with three uncertain parameters. With this approach the mapping of the α levels is performed from the n_θ dimensional space characterized by a chi-squared distribution of n_θ degrees of freedom ($\chi^2_{n_\theta}(\alpha)$) to an n_ψ dimensional space in which the same α-levels are characterized by a chi-squared distribution with n_ψ degrees of freedom ($\chi^2_{n_\psi}(\alpha)$), with usually $n_\psi = 1$. Hence the probability mapping between the two spaces characterized by different degrees of freedom can be captured by multiplying the obtained worst-case deviations ($\delta\psi_{\text{w.c.}}$) with the ratio $\left(\chi^2_{n_\psi}(\alpha)/\chi^2_{n_\theta}(\alpha)\right)^{1/2}$. This method is exact for models that are first-order in the parameters. For higher-order power series expansions the contour mapping approach is less accurate but much less computationally expensive than the approach described next.

2.2.2. Uncertainty propagation using the Monte Carlo method with second-order power series expansion

For increased accuracy of the estimated shape of the pdf, the second-order series expansion can also be used in a classical or Latin Hypercube type Monte Carlo simulation. In this approach the model is approximated by the second-order power series expansion given by (6) where the first- and second-order sensitivities are calculated along the time-varying control trajectory. The second-order sensitivities, **M**, can be calculated using a finite difference approximation from the first-order sensitivities, requiring only one or two simulations (per parameter) using the nonlinear model augmented with first-order sensitivity equations. In the next step, the Monte Carlo method is applied to the second-order power series expansion with fixed (but time-varying) L and **M**. This revised Monte Carlo method is orders-of-magnitude computationally more efficient than applying the brute-force Monte Carlo method to the original nonlinear model. The computational cost of this second-order approximate Monte Carlo method is higher than using the analytic pdf for the first-

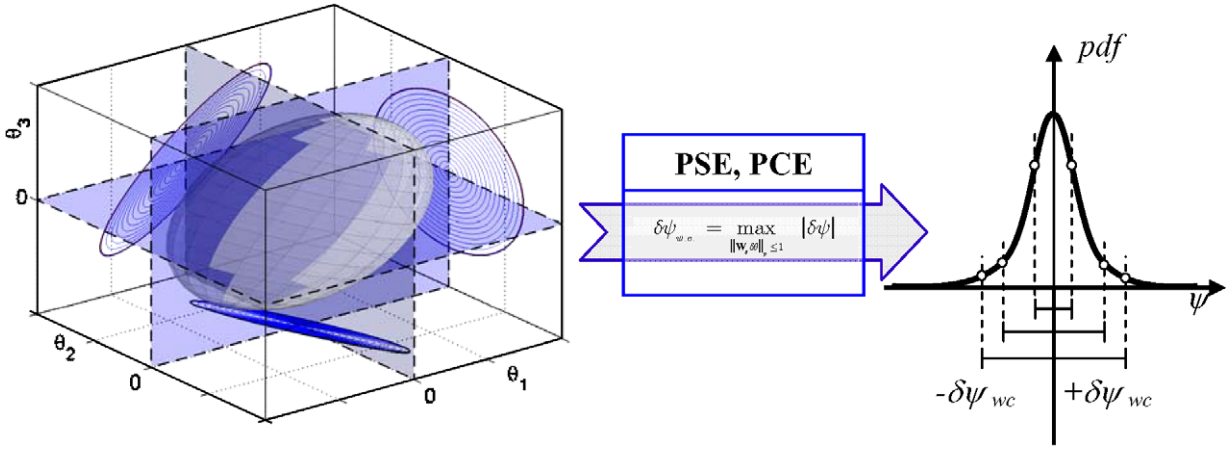

Fig. 2. Distributional robustness analysis based on the contour mapping technique.

order power series approximation but provides a much better estimate of the shape of the output distribution.

2.3. Uncertainty analysis using polynomial chaos expansions

An alternative to the power series based techniques is to use polynomial chaos expansions (PCEs). If the parameter uncertainties are described in terms of standard normal random variables, the PCE [17] can describe the model output ψ as an expansion of multidimensional Hermite polynomial functions of the uncertain parameters θ. For other types of random variables, either different polynomial bases (e.g., Laguerre for the gamma distribution, Legendre for the uniform distribution, etc.) or an appropriate transformation can be used [18]. Using the Hermite bases in the PCE, the output can be expressed in terms of the standard random normal variables $\{\theta_i\}$ using an expansion of order d:

$$\psi^{(d)} = \underbrace{a_0^{(d)}\Gamma_0}_{\text{constant}} + \underbrace{\sum_{i_1=1}^{n_\theta} a_{i_1}^{(d)}\Gamma_1(\theta_{i_1})}_{\text{first order terms}} + \underbrace{\sum_{i_1=1}^{n_\theta}\sum_{i_2=1}^{i_1} a_{i_1 i_2}^{(d)}\Gamma_2(\theta_{i_1},\theta_{i_2})}_{\text{second order terms}} + \underbrace{\sum_{i_1=1}^{n_\theta}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} a_{i_1 i_2 i_3}^{(d)}\Gamma_3(\theta_{i_1},\theta_{i_2},\theta_{i_3})}_{\text{third order terms}} + \cdots, \quad (16)$$

where n_θ is the number of parameters, the $a_{i_1}^{(d)}, a_{i_1 i_2}^{(d)}, a_{i_1 i_2 i_3}^{(d)}, \ldots$ are deterministic coefficients in $\mathbb{R}$ to be estimated, and the multidimensional Hermite polynomials of degree $m = i_1, i_2, \ldots, i_{n_\theta}$, $\Gamma_m(\theta_{i_1}, \ldots, \theta_m)$ are

$$\Gamma_m(\theta_{i_1}, \ldots, \theta_m) = (-1)^m \mathrm{e}^{1/2\theta^{\mathrm{T}}\theta} \frac{\partial^m \mathrm{e}^{-1/2\theta^{\mathrm{T}}\theta}}{\partial\theta_1 \cdots \partial\theta_m}. \quad (17)$$

The polynomial chaos terms are random variables, since they are functions of the random variables, and terms of different order are orthogonal to each other (with respect to an inner product defined in Gaussian measures as the expected value of the product of the two random variable, i.e., $\mathscr{E}[\Gamma_i\Gamma_j] = 0$ for $\Gamma_i \neq \Gamma_j$). In addition, polynomial chaos terms of the same order but with a different argument list are also orthogonal ($\mathscr{E}[\Gamma_m(\{\theta\}_i)\Gamma_m(\{\theta\}_j)] = 0$, $i \neq j$). In PCE any form of polynomial could be used but the properties of orthogonal polynomials make the uncertainty analysis more efficient. For example, calculating the expected value of both sides of (16) results in the expected value of ψ being simply $\mathscr{E}[\psi^{(d)}] = a_0^{(d)}\Gamma_0$. The calculation of other statistical measures is also significantly simplified using the properties of orthogonality. The orthogonal polynomials are derived from the probability distribution of the parameters using the orthogonality condition:

$$\int_0 f_{\text{d.f.}}(\theta)\Gamma_i(\theta)\Gamma_j(\theta) = \delta_{ij} \quad \text{where } \delta_{ij} = 1 \text{ if } i = j, \delta_{ij} = 0 \text{ if } i \neq j. \quad (18)$$

Since $\Gamma_0(\theta) = 1$ the first-order Hermite polynomial can be calculated from $\int_{\theta_1} f(\theta_1)\Gamma_1(\theta_1)(1) = 0$, and the procedure can be repeated to obtain all terms in the PCE. The number of coefficients in the PCE depends on the number of uncertain parameters and the order of expansion (e.g., there are 6 coefficients for two parameters and a second-order PCE and 15 coefficients for a fourth-order PCE, whereas for four uncertain parameters there are 15 coefficients for a second-order PCE and 70 for a fourth-order PCE), however, for most engineering applications it is not necessary to use higher order than three or four. The polynomial chaos expansion is convergent in the mean-square sense [18], therefore the coefficients in the PCE are calculated using least-squares minimization considering sample input/output pairs from the model, so that the best fit is achieved between the PCE and the nonlinear model (or experimental data). Because all parameters are considered as standard random variables, for a more accurate determination of the coefficients $a_{i_1}^{(d)}, a_{i_1 i_2}^{(d)}, a_{i_1 i_2 i_3}^{(d)}, \ldots$ the probability distribution of $\{\theta_i\}$ has to be considered in the selection of the sampling points. In principle there are two main methods for computing the coefficients of the PCE: (i) the probabilistic collocation method (PCM) [12,13], and (ii) the regression method with improved sampling (RMIS) [11]. Both methods are weighted residual schemes, which differ in the way in which the sampling points are chosen. The methods use the principle of collocation, which imposes that ψ is exact at a set of chosen collocation points, thus making the residual between the output of PCE and complex nonlinear model at those points equal to zero. In PCM, the number of collocation points is set equal to the number of unknown coefficients, which are found by solving a set of linear equations generated from the outputs from the original simulation model. In RMIS, additional collocation points are selected to improve the accuracy of the computed coefficients. The two methods also differ in how the collocation points are chosen. In PCM, the collocation points are selected from the roots of the orthogonal polynomial of a degree one higher than the order of the PCE [19,20]. Which roots are selected affects the accuracy of the approximation. In this work the probability collocation method was used to calculate the coefficients. An iterative approach to select the order of the PCE was implemented (see Fig. 3), with automatic derivation of the Hermite polynomials using an algorithm based on ORTHPOL [21].

Generally, the approaches based on PCE are similar to the power series expansion-based techniques, since they also utilize a simpler representation of the simulation model, which can be used to compute the pdf of the outputs (either directly via the Monte Carlo method or via the contour mapping approach based on the solution of the worst-case optimization). The PCE and PSE can be used to analytically compute statistical measures, such as the mean, variance, or higher order moments of the outputs.

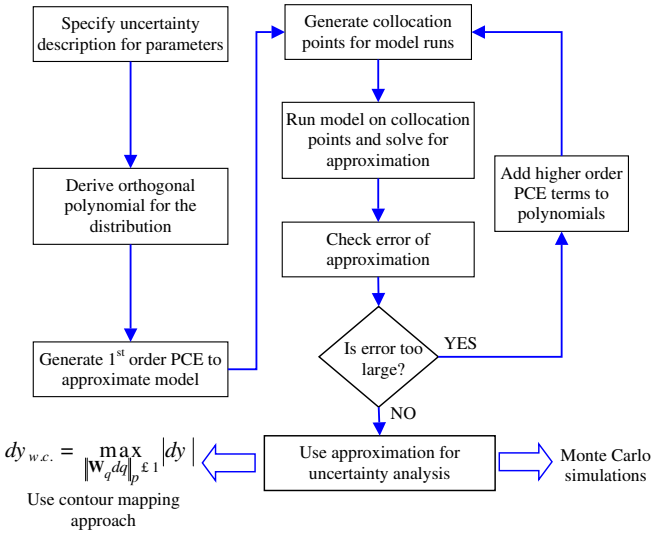

Fig. 3. Flowchart of the algorithm used to obtain the PCE approximation of the nonlinear model for distributional robustness analysis.

3. Application to a simulated batch crystallization process

Crystallization from solution is an industrially important unit operation due to its ability to provide high purity separation. The control of the crystal size distribution (CSD) can be critically important for efficient downstream operations (such as filtration or drying) and product quality (e.g., bioavailability, tablet stability, dissolution rate). Most studies on the optimal control of batch crystallizers focus on computing the temperature profile that optimizes some property of the CSD [22]. The problem of computing the optimal temperature profile can be formulated as a nonlinear optimization problem, which is then solved using general-purpose optimization algorithms. The commonly used approach to describe the temperature trajectory is to discretize the batch time and consider the temperatures at every discrete time k as the optimization variables. In this case the optimal control problem can be written in the following form:

$$\underset{T(k)}{\text{optimize}}\ J \tag{19}$$

subject to:

$$f(x) = \begin{bmatrix} B \\ G\mu_0 \\ 2G\mu_1 \\ 3G\mu_2 \\ 4G\mu_3 \\ -\rho_c k_v 3G\mu_2 \\ G\mu_{\text{seed},0} \\ 2G\mu_{\text{seed},1} \\ 3G\mu_{\text{seed},2} \end{bmatrix} \tag{20}$$

$$T_{\min}(k) \leqslant T(k) \leqslant T_{\max}(k),$$
$$R_{\min}(k) \leqslant \frac{\mathrm{d}T(k)}{\mathrm{d}t} \leqslant R_{\max}(k),$$
$$C_{\text{final}} \leqslant C_{\text{final,max}}, \tag{21}$$

where the objective function J is a function of the states, and usually is a representative property of the final CSD. The equality constraints (20) are the model equations, with initial conditions given in [23], where μ_i is the ith moment ($i = 0,\ldots,4$) of the total crystal phase (resulted from growth from seed and nucleation), $\mu_{\text{seed},j}$ is the jth moment ($j = 0,\ldots,3$) corresponding to the crystals grown from seed, C is the solute concentration, T is the temperature, r_0 is the crystal size at nucleation, k_v is the volumetric shape factor, and ρ_c is the density of the crystal. The rate of crystal growth (G) and the nucleation rate (B), respectively, are given by [24]:

$$G = k_g S^g, \tag{22}$$
$$B = k_b S^b \mu_3, \tag{23}$$

where $S = (C - C_{\text{sat}})/C_{\text{sat}}$ is the relative supersaturation, and $C_{\text{sat}} = C_{\text{sat}}(T)$ is the saturation concentration. The model parameter vector consists of the kinetic parameters of growth and nucleation:

$$\theta^{\mathrm{T}} = [g, k_g, b, k_b], \tag{24}$$

with nominal values [16]:

$$\hat{\theta}^{\mathrm{T}} = [1.31, \exp(8.79), 1.84, \exp(17.38)], \tag{25}$$

with the uncertainty description of the form (4) characterized by the covariance matrix [7]:

$$\mathbf{V}_\theta^{-1} = \begin{bmatrix} 102873 & -21960 & -7509 & 1445 \\ -21960 & 4714 & 1809 & -354 \\ -7509 & 1809 & 24225 & -5198 \\ 1445 & -354 & -5198 & 1116 \end{bmatrix}. \tag{26}$$

In the inequality constraints (21), $T_{\min}$, $T_{\max}$, $R_{\min}$, and $R_{\max}$ are the minimum and maximum temperatures and temperature ramp rates, respectively, during the batch. The first two inequality constraints ensure that the temperature profile stays within the operating range of the crystallizer. The last inequality constraint ensures that the solute concentration at the end of the batch C_{final} is smaller than a certain maximum value $C_{\text{final,max}}$ set by the minimum yield required by economic considerations.

Typical crystal size distribution (CSD) parameters of interest are the nucleation to seed mass ratio ($J_{\text{n.s.r.}}$), coefficient of variation ($J_{\text{c.v.}}$), and weight mean size of the crystals ($J_{\text{w.m.s.}}$):

$$J_{\text{n.s.r.}} = (\mu_3 - \mu_{\text{seed},3})/\mu_{\text{seed},3} \tag{27}$$
$$J_{\text{c.v.}} = (\mu_2\mu_0/(\mu_1)^2 - 1)^{1/2} \tag{28}$$
$$J_{\text{w.m.s.}} = \mu_4/\mu_3 \tag{29}$$

The optimal temperature trajectory in Fig. 4 used to compare the uncertainty analysis methods was computed by solving the optimal control problem (19)–(21) for $J = J_{\text{n.s.r.}}$ with the nominal parameter $\hat{\theta}$, which led to a decrease of more than 20% in the nucleation to seed mass ratio compared to the linear cooling profile. The distributional uncertainty analysis approaches based on the power series approximation and polynomial chaos expansion, respectively, were evaluated in comparison to Monte Carlo (MC) simulations based on the full nonlinear model. 80,000 random parameter sets with mean $\hat{\theta}_i$ and covariance $\mathbf{V}_\theta$ were generated from the multivariate distribution using the Cholesky decomposition of the covariance matrix. With the random parameter vectors, Monte Carlo simulation was performed using the full dynamic model. Then, the frequency that the simulation output from the Monte Carlo simulation fell in the confidence interval obtained for a certain α level with the power series approach was computed for all states and outputs. The results obtained with the first-order analysis approach for the output variables are shown in Fig. 5. The accuracy of the first-order approach is very good with a slightly decreasing tendency for large α values. This can be explained by the cumulative effect of the truncation error due to the first-order approach (which increases when α increases) and the tail effect of the Monte Carlo simulation (as seen on Fig. 1). Although there is a slight decrease in the accuracy of the first-order approach for large α, in the case of certain process outputs the first-order expansion was quite accurate for estimating the size of confidence intervals on the CSD objectives. As seen next, the first-order expansion is not accurate for estimating the shape of the pdfs on the CSD objectives.

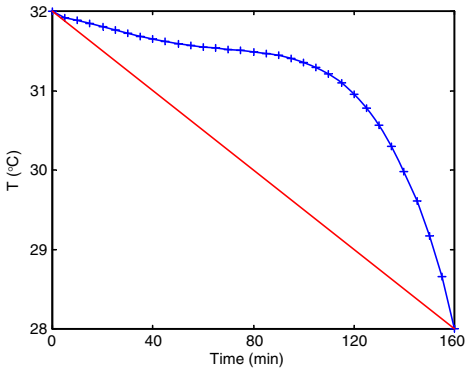

Fig. 4. Linear cooling and optimal temperature profile for $J = J_{\text{n.s.r.}}$.

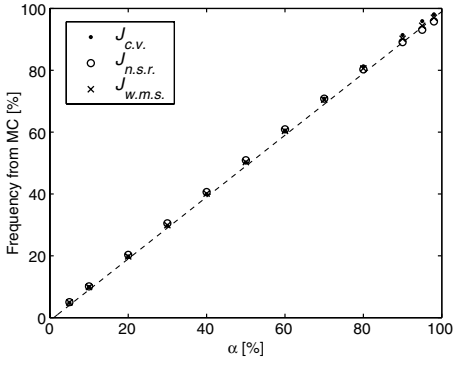

Fig. 5. Comparison between first-order power series analysis and the Monte Carlo method applied to the full nonlinear model. The dashed line corresponds to the ideal case.

For a better comparison of the uncertainty analysis methods, the entire pdfs of the states and outputs are computed based on the frequency histograms for all model states and outputs at the end of the batch. The histograms were obtained by splitting the range of the values in equal-sized bins (called classes), and then counting the number of points that fall into each bin. The probability distribution functions were calculated from the frequency histograms of the simulations, as the count in the class divided by the number of observations times the class width. For this normalization, the area (or integral) under the histogram is equal to one, and the resulted normalized histogram resembles the real pdf. Fig. 6 shows the histograms obtained via the Monte Carlo method applied to the full nonlinear model, using 80,000 parameter sets. The resulted distributions are not Gaussian, showing the effect of the nonlinearity on the uncertainty propagation. Some of the states are closer to the form of the normal distribution, e.g., μ_3, whereas others, e.g., μ_0, show large deviations from a Gaussian distribution. Although the Monte Carlo method provides a good estimate for the shape of the nonlinear distributions, it requires a large number of sampling points and very high computational cost (see Table 1), which makes it inappropriate for on-line applications such as real-time robust feedback control and inconvenient for off-line applications such as robust batch recipe design.

The first-order PSE approach estimates the output distributions (Fig. 7) at very low computational cost (Table 1). Since the first-order PSE approach uses a linear approximation of the nonlinear model, the output distributions are Gaussian, but still provide a good approximation of the real distributions for most of the states and outputs. This approximation is accurate enough, for example, for robust controller design, or minimum variance control, where the objective is to design a controller that decreases the variations around the nominal performance index. It is an acceptable approximation that, if the variance of the Gaussian approximation of the nonlinear distribution is decreased, a similar effect will be achieved for the real distribution.

There certainly exist some finite-time processes in which first-order analysis will not be accurate enough. Also, accurate quantification of the risk of producing batches with poor product characteristics and accurate assessment of the performance of fault diagnostic systems requires a more accurate quantification of the output pdfs. In this case the higher order power series expansion or polynomial chaos expansion give more accuracy while keeping the

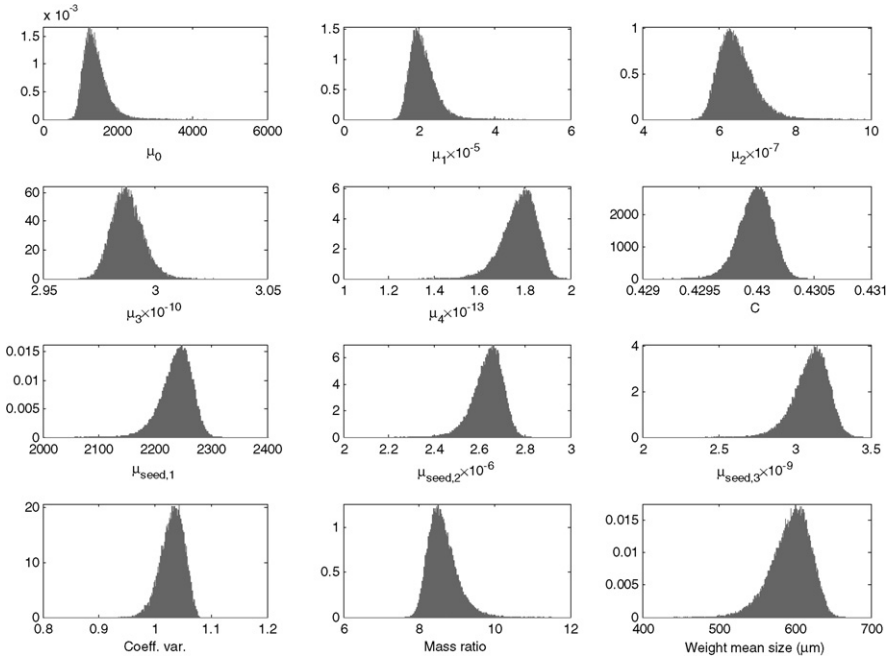

Fig. 6. Probability distribution functions at the end of the batch obtained via the Monte Carlo method applied to the full nonlinear model.

Table 1
Computational cost of the various approaches to compute the distribution for all process outputs during the entire batch (on a P4-1.4 GHz computer)

Method	Computational time
Monte Carlo simulation with full dynamic model (80,000 points)	8 h
First-order approach (50 α levels)	1 s
Monte Carlo applied to second-order series expansion model (80,000 points)	4 min
Second-order polynomial chaos expansion	20 s

computational cost low. Applying the Monte Carlo method to the second-order PSE results in a more accurate approximation of the nonGaussian distribution (Fig. 8) with relatively low computational cost (Table 1). In our simulations the sensitivity jacobian L was calculated by integrating the sensitivity Eq. (9) which were 32 ($n_x \times n_\theta$) ODEs. The computational cost was reduced by computing the matrices $\mathbf{J}_x = \mathrm{d}f/\mathrm{d}x$ and $\mathbf{J}_\theta = \mathrm{d}f/\mathrm{d}\theta$ analytically (considering $r_0 = 0$):

$$A = \begin{bmatrix} 0 & 0 & 0 & B/\mu_3 & 0 & bB/(C-C_s) & 0 & 0 & 0 \\ G & 0 & 0 & 0 & 0 & gG\mu_0/(C-C_s) & 0 & 0 & 0 \\ 0 & 2G & 0 & 0 & 0 & 2gG\mu_1/(C-C_s) & 0 & 0 & 0 \\ 0 & 0 & 3G & 0 & 0 & 3gG\mu_2/(C-C_s) & 0 & 0 & 0 \\ 0 & 0 & 0 & 4G & 0 & 4gG\mu_3/(C-C_s) & 0 & 0 & 0 \\ 0 & 0 & -3\rho_c k_v G & 0 & 0 & -3\rho_c k_v gG\mu_2/(C-C_s) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & gG\mu_{\mathrm{seed},0}/(C-C_s) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2gG\mu_{\mathrm{seed},1}/(C-C_s) & 2G & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3gG\mu_{\mathrm{seed},2}/(C-C_s) & 0 & 3G & 0 \end{bmatrix} \tag{30}$$

$$\mathbf{J}_\theta = \begin{bmatrix} 0 & 0 & B\ln(S) & B/k_b \\ \mu_0 G\ln(S) & \mu_0 G/k_g & 0 & 0 \\ 2\mu_1 G\ln(S) & 2\mu_1 G/k_g & 0 & 0 \\ 3\mu_2 G\ln(S) & 3\mu_2 G/k_g & 0 & 0 \\ 4\mu_3 G\ln(S) & 4\mu_3 G/k_g & 0 & 0 \\ -3\rho_c k_v \mu_2 G\ln(S) & -3\rho_c k_v \mu_2 G/k_g & 0 & 0 \\ \mu_{\mathrm{seed},0} G\ln(S) & \mu_{\mathrm{seed},0} G/k_g & 0 & 0 \\ 2\mu_{\mathrm{seed},1} G\ln(S) & 2\mu_{\mathrm{seed},1} G/k_g & 0 & 0 \\ 3\mu_{\mathrm{seed},2} G\ln(S) & 3\mu_{\mathrm{seed},2} G/k_g & 0 & 0 \end{bmatrix}. \tag{31}$$

To obtain the matrix of second-order sensitivities, $\mathbf{M}$, forward finite differences was used, which requires only one additional simulation with the augmented model for each parameter. For the Monte Carlo simulations for a fixed control trajectory, L and $\mathbf{M}$ are computed off-line and Monte Carlo simulations are performed using the PSE, which is computationally much more efficient than using the full integration of the nonlinear model expressed by the set of ODEs.

As with the PSE, the PCE can be used with the Monte Carlo method, or contour mapping which solves of the optimization

$$\delta\psi_{\mathrm{w.c.}}(\alpha) = \max_{\|\mathbf{W}_\theta \delta\theta\| \leqslant 1} \delta\psi^{\mathrm{PCE}}, \tag{32}$$

for different values of α, where $\delta\psi^{\mathrm{PCE}}$ is calculated for the PCE. The order of the PCE was determined to be two using the iterative approach in Fig. 3 (including third-order terms did not increase the accuracy of the approximation significantly). Since there are four uncertain parameters, the sec-

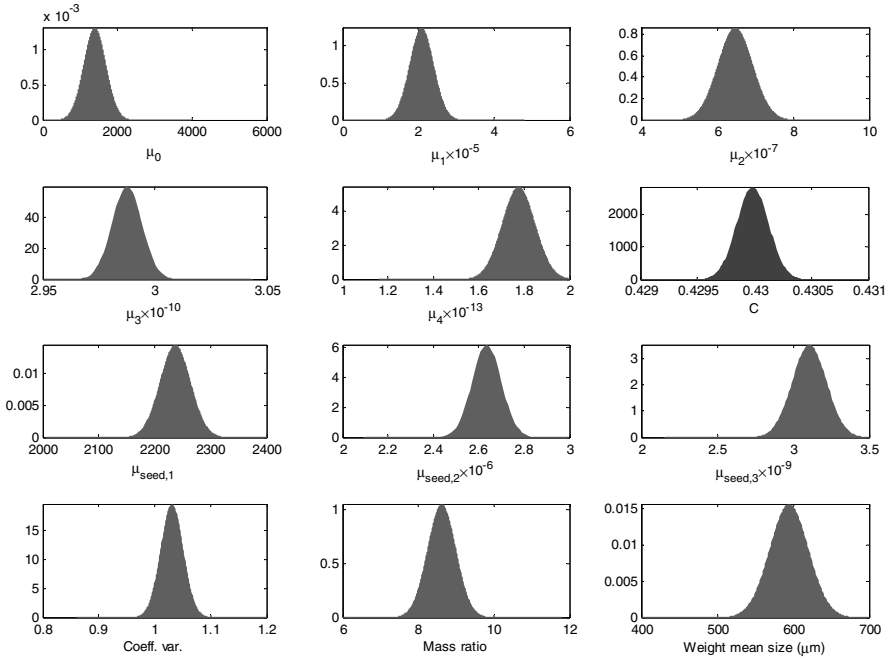

Fig. 7. Probability distribution functions at the end of the batch obtained from the first-order power series expansion (PSE).

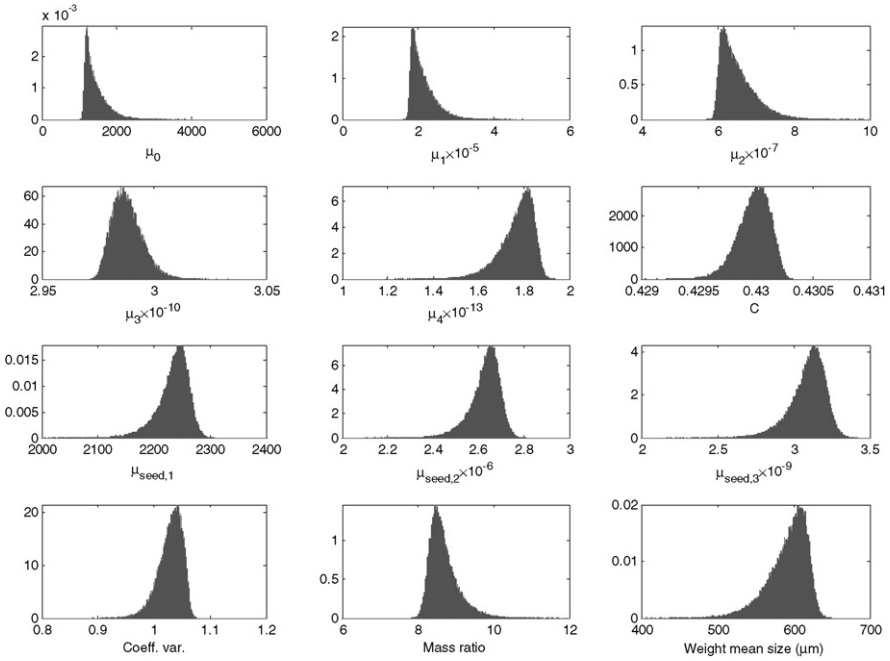

Fig. 8. Probability distribution functions at the end of the batch obtained via the Monte Carlo method applied to the second-order PSE.

ond-order PCE required the determination of 15 coefficients (according to $N_c^{(2)} = 1 + 2n_\theta + n_\theta(n_\theta - 1)/2$). The second-order four-dimensional PCE is

$$\psi^{(2)}(\theta) = \sum_{i=0}^{14} a_i \Gamma_i, \tag{33}$$

where the Hermite polynomials are in Table 2. The coefficients a_i were computed using the probabilistic collocation method with the collocation points obtained from the roots of the third-order Hermite polynomials [19,20]. The optimization problem (32) was convex for the four-dimensional second-order polynomial as the objective function and so was easy to solve (see Fig. 9). The second-order PCE very accurately captured the shape of the pdfs at the end of the batch (compare Figs. 6 and 9), resulting in a more accurate estimation of the effects of parameter uncertainties than the second-order PSE at a lower computational cost.

The pdfs of the moment μ_0, which shows the strongest deviation from the Gaussian distribution, are plotted for

Table 2
The Hermite polynomials for the second-order four-dimensional PCE

ith Polynomial chaos	Γ_i	Order of polynomial chaos
0	1	0
1	θ_1	1
2	θ_2	1
3	θ_3	1
4	θ_4	1
5	$\theta_1^2 - 1$	2
6	$\theta_1\theta_2$	2
7	$\theta_1\theta_3$	2
8	$\theta_1\theta_4$	2
9	$\theta_2^2 - 1$	2
10	$\theta_2\theta_3$	2
11	$\theta_2\theta_4$	2
12	$\theta_3^2 - 1$	2
13	$\theta_3\theta_4$	2
14	$\theta_4^2 - 1$	2

the first-order PSE and second-order PCE in Fig. 10 with the pdf from the Monte Carlo method applied to the full nonlinear model. As noted before, while the first-order PSE gave acceptable accuracy (which is even better for the rest of the states and outputs) for most purposes, a more accurate estimate of the tails of the distribution can have significant economic or safety implications, especially in the case of robust design. If the performance specification was in terms of μ_0, Fig. 10 illustrates that, if there was a high penalty on exceeding the specification limits, then a design based on the first-order PSE would lead to a more conservative result with respect to the lower specification limit, whereas it underpredicts the probability of events with large values of μ_0 which could have safety implications in particular applications, such as pharmaceutical crystallization. The second-order PCE based approach captures very well the shape and tails of the distribution.

Unlike sensitivity analysis, these uncertainty analysis approaches provide the variation of the whole distribution for all states and outputs along the entire batch. This is illustrated in Fig. 11, which shows the pdfs for the performance index $J_{\text{w.m.s.}}$ with the various uncertainty analysis methods. As noted before, the power series and PCE robustness analysis methods have low computational costs, therefore they can be used easily for the assessment and efficient design of robust control systems. The simplified dependency of these expansions on the parameters lead to simpler optimization of the parameters than obtained by direct robust controller synthesis, which usually leads to nonconvex problem formulations. Robust controller design can be formulated as a classical minmax optimization or as a multiobjective optimization problem where one objective accounts for the nominal performance and the other (usually a measure of the variance of the distribution) accounts for robustness [25,26]. Both uncertainty analysis methods can be used to efficiently calculate the robustness term. Additionally, since the PCE approach provides a means for accurate estimation of the shape of distribution, it can form the basis of a procedure for designing controllers to shape the distribution, which would provide more flexibility in addressing uncertainty than worst-case or minimum variance control. This is especially motivated by the asymmetry in the economic cost functionals of processes operated near safety, environmental, and quality constraints; for example, designing a controller to shape the distribution in relation to the cost functional rather than merely reducing variance would more closely map controller design to economic objectives.

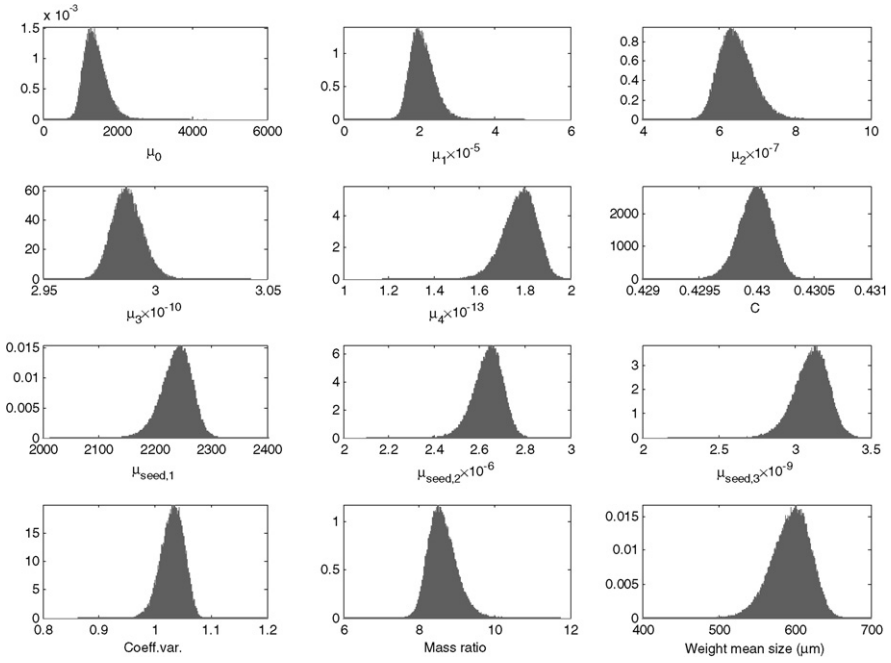

Fig. 9. Probability distribution functions at the end of the batch obtained via the Monte Carlo method applied to the second-order polynomial chaos expansion.

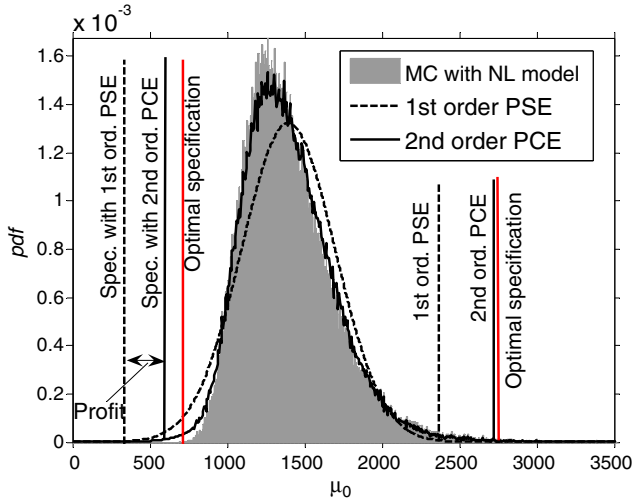

Fig. 10. Potential effect of better distribution shape estimation on the profit/safety in the design under uncertainties. Comparison of the performance between the Monte Carlo (MC) method applied to the full nonlinear (NL) model, and the first-order power series expansion (PSE) and second-order polynomial chaos expansion (PCE) methods.

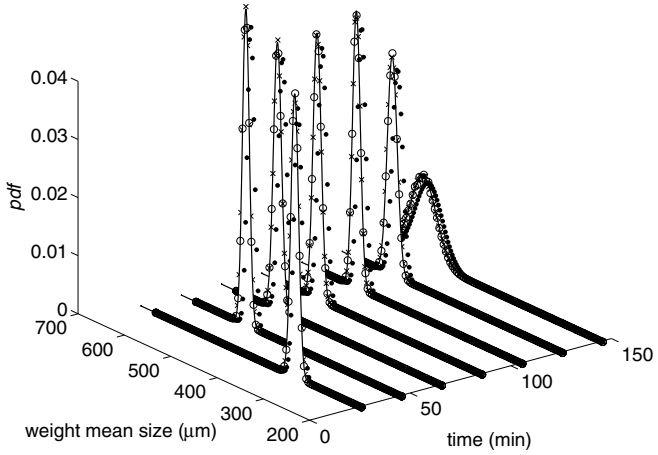

Fig. 11. Variation of the distribution of the weight mean size along the batch, computed using various uncertainty analysis methods. The solid line is for the Monte Carlo method applied to the full nonlinear model with 80,000 sets of parameter, the dots are for the first-order power series method, the circles are for the Monte Carlo method applied to the second-order power series expansion, and the x-marks are for the polynomial chaos expansion method.

4. Conclusions

This paper provides an overview of several computationally-efficient robustness analysis approaches, focusing primarily on computing probability distributions on states and outputs. The approaches are based on the approximate representation of the full process model using power series or polynomial chaos expansions, and provide a qualitative and quantitative estimation of the effect of parameter uncertainties on the states and output variables along a batch. The computational cost of the robustness analysis methods are significantly lower than the classical Monte Carlo method applied to the full nonlinear model. Application of the algorithms to a simulated batch cooling crystallization process provided the recommendations: (1) for some systems, the first-order series analysis is accurate enough for the design of real-time robust feedback controllers, but the accuracy of the first-order expansion should be confirmed *a posteriori* using Monte Carlo simulation or a higher order PCE method, and (2) even with a relatively low-order approximation, the polynomial chaos expansions can provide a very good approximation of the shape and tails of the output and states distribution for batch processes, providing a generally applicable approach for uncertainty propagation in robust batch control and design.

References

[1] M.D. Barrera, L.B. Evans, Optimal design and operation of batch processes, Chem. Eng. Commun. 82 (1989) 45–66.
[2] D.W.T. Rippin, Simulation of single- and multiproduct batch chemical plants for optimal design and operation, Comput. Chem. Eng. 7 (1983) 137–156.
[3] D. Bonvin, Optimal operation of batch reactors: a personal view, J. Process Control 8 (1998) 355–368.
[4] D.L. Ma, S.H. Chung, R.D. Braatz, Worst-case performance analysis of optimal batch control trajectories, AIChE J. 45 (1999) 1469–1476.
[5] B. Srinivasan, D. Bonvin, E. Visser, S. Palanki, Dynamic optimization of batch processes: II. Role of measurements in handling uncertainty, Comput. Chem. Eng. 27 (2003) 27–44.
[6] D.L. Ma, R.D. Braatz, Worst-case analysis of finite-time control policies, IEEE Trans. Control Syst. Technol. 9 (2001) 766–774.
[7] S.M. Miller, J.B. Rawlings, Model identification and control strategies for batch cooling crystallizers, AIChE J. 40 (1994) 1312–1327.
[8] A. Saltelli, K. Chan, E.M. Scott, Sensitivity Analysis, Wiley, 2000.
[9] Z.K. Nagy, R.D. Braatz, Worst-case and distributional robustness analysis of finite-time control trajectories for nonlinear distributed parameter systems, IEEE Trans. Control Syst. Technol. 11 (2003) 694–704.
[10] S.S. Isukapalli, Uncertainty analysis of Transport-Transformation Models, Ph.D. thesis, The State University of New Jersey, 1999.
[11] S.S. Isukapalli, A. Roy, P.G. Georgopoulos, Stochastic response surface methods (SRSMs) for uncertainty propagation: application to environmental and biological systems, Risk Anal. 18 (1998) 351–363.
[12] W. Pan, M.A. Tatang, G.J. McRae, R.G. Prinn, Uncertainty analysis of indirect radiative forcing by anthropogenic sulfate aerosols, J. Geophys. Res. 103 (1998) 3815–3823.
[13] W. Pan, M.A. Tatang, G.J. McRae, R.G. Prinn, Uncertainty analysis of direct radiative forcing by anthropogenic sulfate aerosols, J. Geophys. Res. 102 (1997) 21916–21924.
[14] J.V. Beck, K.J. Arnold, Parameter Estimation in Engineering and Science, Wiley, New York, 1977.
[15] M. Caracotsios, W.E. Stewart, Sensitivity analysis of initial value problems with mixed ODEs and algebraic equations, Comput. Chem. Eng. 9 (1985) 359–365.
[16] J.B. Rawlings, S.M. Miller, W.R. Witkowski, Model identification and control of solution crystallization processes: A review, Ind. Eng. Chem. Res. 32 (1993) 1275–1296.
[17] N. Wiener, The homogeneous chaos, Am. J. Math. 60 (1938) 897–936.
[18] R.G. Ghanem, P.D. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer-Verlag, New York, 1991.
[19] M.A. Tatang, P. Wenwei, R.G. Prinn, G.J. McRae, An efficient method for parametric uncertainty analysis of numerical geophysical models, J. Geophys. Res. 12 (1997) 21932–21952.

[20] M. Webster, M.A. Tatang, G.J. McRae, Application of the Probabilistic Collocation Method for an Uncertainty Analysis of a Simple Ocean Model, MIT Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology, Cambridge, <http://web.mit.edu/afs/athena.mit.edu/org/g/globalchange/www/reports.html>, Report Series No. 4., 1996.
[21] W. Gautschi, Algorithm 726: ORTHPOL–A Package of routines for generating orthogonal polynomials and gauss type quadrature rules, ACM Trans. Math. Software 20 (1994) 21–62.
[22] R.D. Braatz, Advanced control of crystallization processes, Annu. Rev. Control 26 (2002) 87–99.
[23] S.H. Chung, D.L. Ma, R.D. Braatz, Optimal seeding in batch crystallization, Can. J. Chem. Eng. 77 (1999) 590–596.
[24] J. Nyvlt, O. Sohnel, M. Matuchova, M. Broul, The Kinetics of Industrial CrystallizationChemical Engineering Monographs, vol. 19, Elsevier, Amsterdam, 1985.
[25] Z.K. Nagy, R.D. Braatz, Open-loop and closed-loop robust optimal control of batch processes using distributional worst-case analysis, J. Process Control 14 (2004) 411–422.
[26] Z.K. Nagy, R.D. Braatz, Robust nonlinear model predictive control of batch processes, AIChE J. 49 (2003) 1776–1786.