

## Quality-by-Design Using a Gaussian Mixture Density Approximation of Biological Uncertainties

N. Rossner, Th. Heine, R. King

*Technische Universität Berlin, Chair of Measurement and Control, Secr. ER  
2-1, D-10623 Berlin, Germany (e-mail: Thomas.Heine@tu-berlin.de).*

**Abstract:** In this contribution the uncertainties of a biological process model are taken into account explicitly to calculate optimal process trajectories. For this purpose, the initial condition and the uncertainties of the model parameters are described by a weighted sum of normal distributions. Such a so-called Gaussian mixture density (GMD) approximation is propagated through the nonlinear process model to calculate a second order approximation of the statistical properties of the planned process trajectory. A Value@Risk primary objective is used to obtain an optimal process design procedure in presence of uncertainties. In an extensive simulation study a descriptive fermentation process model is used to compare the classical trajectory planning with the robust design approaches. Here, different degrees of approximation complexity and the influence of the weighting factor in the Value@Risk dual objective criterion is investigated.

*Keywords:* Quality-by-Design, Gaussian mixture density, value-at-risk, uncertain dynamic systems

### 1. INTRODUCTION

The design of biological processes is often done heuristically even though model-based approaches have shown a much higher potential. However, obtaining a mathematical description of the cell metabolism is a complex and time consuming task. Even the simplest organism contains more than a thousand interacting chemical species which lead to its complex and highly non-linear dynamic behavior. For a model-based process design, interactions and intracellular substances are usually lumped together such that many effects are not resolved in the mathematical description. Additionally, the measurement situation in biological processes is very difficult. Not only the evaluation of the compartments within the cell (DNA, RNA, proteins, etc.) is error-prone but also the substrate measurements are usually uncertain because of inhomogeneities in the fermenter and unknown disturbances in the enzymatic or chemical analysis methods. Summing up these obstacles, the parameters of a chosen model structure show significant uncertainties. These uncertainties carry over to the step of process design. Hence, the outcome of the design has to be associated with a certain probability distribution which should be accounted for. There are several ways to estimate the consequences of uncertain parameters and initial conditions. By calculating the Fisher information matrix (FIM), known measurement noise can be mapped to a normally distributed parameter noise according to the parameter sensitivity of the model [Romero and Navarro (2009); Heine et al. (2008)]. Using a bootstrap analysis [Efron (1979)] the parameter noise can also be described as non-normally distributed. Such a bootstrap result for the maximal specific growth rate  $\mu_{X,\max}$  of a biological process is shown exemplary in Fig. 1 as a gray histogram

\* Financial support from the Deutsche Forschungsgemeinschaft and the Cluster of Excellence Unifying Concepts in Catalysis (EXC 314/1) are gratefully acknowledged.

scaled to a probability density. When only a single normal distribution (1-GMD) is used to describe the histogram of realizations, statistical information is lost due to the symmetric nature of normal densities. This information can be preserved using a Gaussian mixture density consisting of four partial densities (4-GMD). Transferring a set of samples into a Gaussian mixture density is usually done using the expectation-maximization (EM) algorithm [Dempster et al. (1977); Bilmes (1998)], which is very common in image recognition. To use such a description for process design, Sec. 2 of this contribution, presents a method for the propagation of normal and non-normal distributions through a non-linear system via Gaussian mixture densities based on the unscented transformation [Julier and Uhlmann (2003); van der Merwe (2004)].

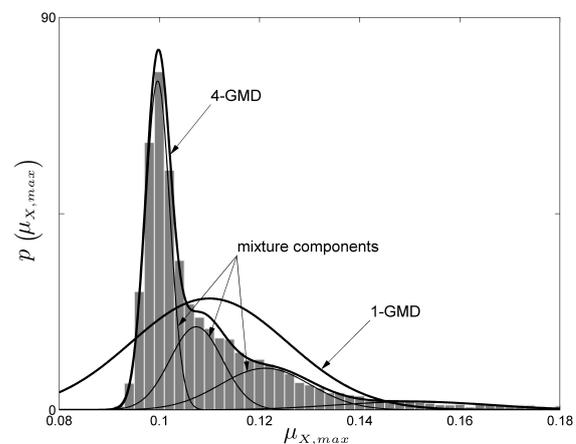


Fig. 1. The parameter distribution from a bootstrap analysis based on 15.000 samples and its approximation using a single normal distribution (1-GMD) and a Gaussian mixture density with four components (4-GMD), respectively.

Sec. 3 shows how the probability information is then incorporated into a design objective for a biological model introducing a Value@Risk formulation. In Sec. 4, three process designs have been calculated. While the first design ignores the uncertainties and uses the nominal parameter values only, the second design will consider the parameter uncertainty as normally distributed. The last design describes the normally distributed parameters using an approximation of 27 mixture components. For this design, it has been investigated, how the process design changes when the primary objective is focusing more and more on the process deviations. The paper concludes with a discussion of the proposed method.

## 2. PREDICTION OF UNCERTAINTIES

There are several different methods for considering uncertainties for process design. Some of these concepts are based on game-theoretic ideas [Chen et al. (1997)], min-max formulations [Lee and Yu (1997)] also with path-constraints [Kühl et al. (2007)],  $H_\infty$ -control [Magni et al. (2001)], or chance constrained programming [Li et al. (2008)]. In this contribution the propagation of a normally distributed input  $\underline{x}$  (e.g. parameters and/or initial condition) through a general non-linear function  $\underline{g}(\underline{x})$  (e.g. primary objective) is calculated in order to reduce deviations in the output of the process. While most methods will approximate the statistical properties of the propagated input  $\underline{g}(\underline{x})$ , a Monte-Carlo simulation enables unbiased estimates if an infinite number of samples is used. Because of the huge numerical burden this method is inapplicable for process design and control even though deterministic sampling techniques reduce the required samples. By contrast, an approximation of first order is a numerically fast method that linearizes the non-linear function  $\underline{g}(\underline{x})$  around the mean of the input  $\bar{\underline{x}}$ . The mean  $\bar{\underline{y}}$  and the covariance  $\underline{\mathbf{C}}_{\underline{y}}$  of the output  $\underline{y} = \underline{g}(\underline{x})$  can then be approximated by mapping the input over the linearized plane

$$\bar{\underline{y}} = \underline{g}(\bar{\underline{x}}), \quad \underline{\mathbf{C}}_{\underline{y}} = \left. \frac{d\underline{g}}{d\underline{x}} \right|_{\bar{\underline{x}}} \cdot \underline{\mathbf{C}}_{\underline{x}} \cdot \left. \frac{d\underline{g}}{d\underline{x}} \right|_{\bar{\underline{x}}}^T. \quad (1)$$

Using a Taylor series expansion of  $\underline{g}(\underline{x})$ , the derivative  $d\underline{g}/d\underline{x}$  in (1) will be calculated around  $\bar{\underline{x}}$  while a statistical linearization would also consider the spread of the random variable  $\underline{\mathbf{C}}_{\underline{x}}$  which leads to better results. However, this approximation is only accurate to first order. For a second order approximation based on the Taylor series expansion also the second derivative  $d^2\underline{g}/d\underline{x}^2$  is needed. This requires a numerical effort of at least  $\mathcal{O}(L_{\underline{x}}^2)$  where  $L_{\underline{x}}$  is the number of uncertain input dimensions. More effective is a statistical approach, known as the unscented transformation. It is a second order approximation of which the effort only grows linearly, requiring  $2L_{\underline{x}} + 1$  function evaluations to form a parabolic approximation. Its basic idea is to describe the normally distributed input  $\underline{x}$  via  $2L_{\underline{x}} + 1$  discrete points located on  $L_{\underline{x}}$  statistically decoupled axis  $\underline{s}_{x_i}$ . These so-called Sigma-points will then be propagated through the non-linear model  $\underline{g}(\underline{x})$ . From the location of these points a  $L_{\underline{x}}$ -dimensional paraboloid can be calculated as a second order approximation of the true unknown solution  $\underline{g}(\underline{x})$ . The mean  $\bar{\underline{y}}$  and covariance  $\underline{\mathbf{C}}_{\underline{y}}$  of the output  $\underline{y} = \underline{g}(\underline{x})$  can then be calculated as

$$\begin{aligned} \bar{\underline{y}} &\approx \frac{h^2 - L_{\underline{x}}}{h^2} \underline{g}(\bar{\underline{x}}) + \frac{1}{2h^2} \sum_{i=1}^{L_{\underline{x}}} (\underline{g}(\bar{\underline{x}} + h\underline{s}_{x_i}) + \underline{g}(\bar{\underline{x}} - h\underline{s}_{x_i})) \quad (2) \\ \underline{\mathbf{C}}_{\underline{y}} &\approx \frac{1}{4h^2} \sum_{i=1}^{L_{\underline{x}}} (\underline{g}(\bar{\underline{x}} + h\underline{s}_{x_i}) - \underline{g}(\bar{\underline{x}} - h\underline{s}_{x_i})) \cdot (\underline{g}(\bar{\underline{x}} + h\underline{s}_{x_i}) - \underline{g}(\bar{\underline{x}} - h\underline{s}_{x_i}))^T \\ &\quad + \frac{h^2 - 1}{4h^4} \sum_{i=1}^{L_{\underline{x}}} (\underline{g}(\bar{\underline{x}} + h\underline{s}_{x_i}) + \underline{g}(\bar{\underline{x}} - h\underline{s}_{x_i}) - 2\underline{g}(\bar{\underline{x}})) \\ &\quad \cdot (\underline{g}(\bar{\underline{x}} + h\underline{s}_{x_i}) + \underline{g}(\bar{\underline{x}} - h\underline{s}_{x_i}) - 2\underline{g}(\bar{\underline{x}}))^T \quad (3) \end{aligned}$$

with an approximation error of third order. The step size  $h$  can be varied, but as proven in [Nørgaard et al. (1998)] the smallest approximation error is expected when  $h = \sqrt{3}$ . The approximation of  $\bar{\underline{y}}$  and  $\underline{\mathbf{C}}_{\underline{y}}$  can be improved when the input is described by several normal distributions leading to a weighted superposition of  $L_{\underline{x}}$ -dimensional paraboloids for the approximation of the true unknown solution  $\underline{g}(\underline{x})$ . The random input  $\underline{x}$  is then approximated with a Gaussian mixture density (GMD) and can hence be written as

$$\underline{\tilde{x}} \sim \sum_{j=1}^M \alpha^{(j)} \mathcal{N}(\bar{\underline{x}}^{(j)}, \underline{\mathbf{C}}_{\underline{x}}^{(j)}), \quad (4)$$

where  $\alpha^{(j)}$  is the weight of the  $j$ th normal density with mean  $\bar{\underline{x}}^{(j)}$  and covariance matrix  $\underline{\mathbf{C}}_{\underline{x}}^{(j)}$ . Such a description of the input  $\underline{x}$  based on  $M$  normal densities also overcomes the restriction of the unscented transformation to normal densities since any arbitrary distribution can be approximated when a sufficient large number  $M$  of normal densities is used [Yun et al. (2008)]. The output  $\underline{y} = \underline{g}(\underline{x})$  will then also be described by a GMD with  $M$  components which is referred to as  $M$ -GMD. The individual means  $\bar{\underline{y}}^{(j)}$  and covariances  $\underline{\mathbf{C}}_{\underline{y}}^{(j)}$  of the  $j$ th partial distribution are calculated using (2) and (3), while the weight of each single density  $\alpha^{(j)}$  remains constant.

### 2.1 Decomposing normal distributions into a GMD

As motivated in the introduction, parameter uncertainties can be identified as non-normally distributed using a bootstrap analysis. The resulting set of samples can then be transferred into a Gaussian mixture density using the EM-algorithm. Even if the system's uncertainties are normally distributed, one can benefit from the GMD prediction by transferring the normal distribution into a Gaussian mixture density according to (4). Such a decomposition of a normal density is presented in the following paragraphs. Even though approximating the normally distributed input  $\underline{x}$  with a sum of normal densities will cause an approximation error (Fig. 2), the prediction of the higher resolved GMD-input  $\underline{\tilde{x}}$  related to the non-linear function  $\underline{g}(\underline{x})$  will be improved as shown in the last section of this contribution. In order to obtain a GMD input  $\underline{\tilde{x}}$  according to (4), the normally distributed input  $\underline{x}$  will be linearly transformed to its standard distribution. The variables related to the standard distribution are indexed with  $\mathbf{I}$ . Therefore,  $\underline{x}_{\mathbf{I}}$  is

$$\begin{aligned} \underline{x} &\mapsto \underline{x}_{\mathbf{I}} : \mathbf{A}^{-1}(\underline{x} - \bar{\underline{x}}) \\ \underline{x}_{\mathbf{I}} &\sim \mathcal{N}(\mathbf{A}^{-1}(\bar{\underline{x}} - \bar{\underline{x}}), \mathbf{A}^{-1} \underline{\mathbf{C}}_{\underline{x}} \mathbf{A}^{-T}) = \mathcal{N}(\mathbf{0}, \mathbf{I}). \quad (5) \end{aligned}$$

The unit covariance matrix  $\mathbf{I}$  will be obtained using the transformation matrix  $\mathbf{A} = \sqrt{\underline{\mathbf{C}}_{\underline{x}}}$ , where  $\sqrt{\underline{\mathbf{C}}_{\underline{x}}}$  can be any matrix square root. Numerically most efficient would be

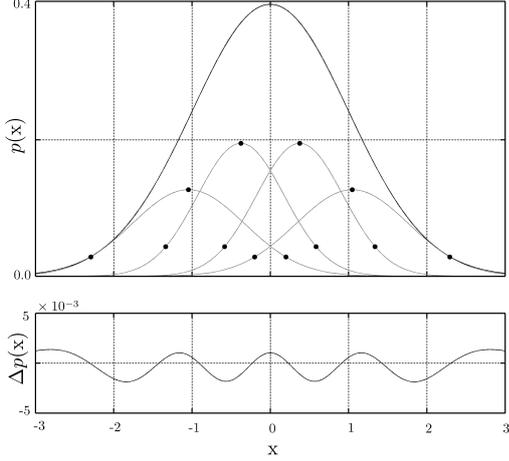


Fig. 2. Top: Standard normal distribution decomposed in a sum of four normal distributions. The Sigma-points are shown as black dots on top of the partial densities in order to show the assignment to their corresponding distribution. Bottom: Approximation error  $\Delta p(x) = p_{\mathcal{N}}(x) - \sum_{i=1}^4 p_i(x)$

the use of the Cholesky-decomposition, most descriptively is probably the use of the Eigenvalue-decomposition since the orthogonal basis is preserved. Anyhow, after the transformation, the desired uncertain variables can be decomposed individually according to Fig. 2 because any correlation is resolved. The shown decomposition results from a least squares optimization on the approximation error  $\Delta p$  with additional constraints on the deviations  $\sigma_{\mathbf{I}}^{(j)}$  in order to prevent wide densities. For all numbers of mixture components the optimization results, i.e.  $\alpha^{(j)}$ ,  $\bar{x}_{\mathbf{I}}^{(j)}$  and  $\sigma_{\mathbf{I}}^{(j)}$ , are stored as a library. The number of decomposing variables, however, should be chosen carefully since the number of overall Gaussian distributions is  $M = \prod_{i=1}^{I_{\mathbf{x}}} m_i$  if  $m_i$  is the number of normal distributions along each uncertain dimension  $i$ . After decomposing certain variables in a reasonable number of univariate distributions, these components then form a multivariate GMD, such that the standardized random variable  $\tilde{x}_{\mathbf{I}}$  can be written as

$$\tilde{x}_{\mathbf{I}} \sim \sum_{j=1}^M \alpha^{(j)} \mathcal{N}\left(\bar{x}_{\mathbf{I}}^{(j)}, \mathbf{C}_{\mathbf{I}}^{(j)}\right). \quad (6)$$

The  $j$ th partial distribution of the GMD has non-zero mean  $\bar{x}_{\mathbf{I}}^{(j)}$  and for its covariance matrix  $\mathbf{C}_{\mathbf{I}}^{(j)} \neq \mathbf{I}$  holds. For the overall GMD density though,  $\bar{x}_{\mathbf{I}} \approx \mathbf{0}$  and  $\mathbf{C}_{\mathbf{I}} \approx \mathbf{I}$  is valid depending on the approximation error  $\Delta p(\underline{x})$  of the decomposition (Fig. 2). Because of the linear nature of the transformation each partial density can be transformed individually and hence the random GMD input reads

$$\tilde{x} \sim \sum_{j=1}^M \alpha^{(j)} \mathcal{N}\left(\mathbf{A} \bar{x}_{\mathbf{I}}^{(j)} + \bar{x}, \mathbf{A} \mathbf{C}_{\mathbf{I}}^{(j)} \mathbf{A}^T\right), \quad \mathbf{A} = \sqrt{\mathbf{C}_{\tilde{x}}}. \quad (7)$$

## 2.2 Robust optimization problem

When considering the input of a dynamical system as being uncertain, the affected process variables will become random as well. In the previous section it was shown, how the effects of a random input on the process values can be

estimated. When planning an optimal trajectory for a production process, an objective  $\Phi$  is formulated. Due to the random nature of the input, i.e.  $x_0$  and  $\theta$ , the formulated objective  $\Phi = g(x_0, \theta)$  will be described by a probability function. Instead of regarding a particular value, the statistical properties of this distribution can be optimized. In (8) the so-called Value@Risk formulation is introduced, where the mean of the objective  $\bar{\Phi}$  is minimized, while contemporary minimizing its squared confidence interval  $\sigma_{\Phi}^2$  (e.g.  $1\sigma$ -interval  $\hat{=} 67\%$ ) by adding it as a penalty term. In case that GMDs are used for the uncertainty prediction the probability density contains information of higher statistical moments which will be considered for the numerical calculation of the confidence region. Anyhow, the importance of the deviation  $\sigma_{\Phi}^2$  towards the mean  $\bar{\Phi}$  of the objective can be adjusted with the factor  $\gamma$ .

$$\mathbf{U}^* = \arg \min_{\mathbf{U}} (\bar{\Phi}(x_0, \mathbf{U}, \theta) + \gamma \sigma_{\Phi}^2)$$

$$\dot{x} = f(t, x(t), u(t), \theta), \quad x(t_0) = x_0$$

$$0 \leq u(t) \leq u_{\max}, \quad \bar{r}_i(x(t)) + \lambda \sigma_{r_i} \leq \underline{0} \quad (8)$$

The non-linear model  $\dot{x} = f(t, x(t), u(t), \theta)$  depends on a set of parameters  $\theta$  as well as on the initial condition  $x_0$  and can be manipulated with a profile  $\mathbf{U} = \{u(t_0), u(t_1), \dots, u(t_k), \dots, u(t_P)\}$  where  $P$  is the number of zero order hold manipulating values in time. The individual rates  $u_k$  have to satisfy a certain input constraint  $u_{\max}$ . Other non-linear process constraints  $r_i$  are formulated in a Value@Risk manner as well in order to consider the time depending change of the process deviations with respect to the constraints. The factor  $\lambda$  determines the safety margin to the actual constraints expressed in  $1\sigma$ -confidence distances.

## 3. ROBUST TRAJECTORY PLANNING FOR A BIOLOGICAL PRODUCTION PROCESS

The previously described methods will be investigated using a simple unstructured biological model.

### 3.1 Unstructured biological model

A typical unstructured model of a biological process can be derived from balance equations whereas the cell is regarded as a black box, which produces a certain product depending on the environmental conditions. Such an unstructured biological model only consists of four states, the biomass  $m_X$ , the mass of available substrate  $m_S$ , the formed product  $m_P$ , and for fed-batch fermentations also the volume of the reactor  $V$ .

$$\begin{aligned} \dot{m}_X &= \mu_X m_X \\ \dot{m}_S &= (-Y_{XS} \mu_X - Y_{main} - Y_{PS} \mu_P) m_X + c_{S,feed} u_S \\ \dot{m}_P &= \mu_P m_X \\ \dot{V} &= u_S \end{aligned} \quad (9)$$

$$\mu_X = \mu_{X,m} \frac{c_S}{K_X + c_S} \frac{K_{XI}}{c_P + K_{XI}}, \quad \mu_P = \mu_{P,m} \frac{c_S}{K_P + c_S + \frac{1}{K_{PI}} c_S^2}$$

The specific growth rate  $\mu_X$  will increase when more substrate is available incorporated by the law of Michaelis-Menten. On the contrary, product formation will induce an inhibition of growth whereby its rate  $\mu_X$  decreases.

The specific production rate  $\mu_P$  will be maximal on a particular low substrate concentration  $c_S$ , which is biologically motivated by the activation of the secondary metabolism during limitations. The assumed initial condition of the process is  $m_X = 0.5\text{g}$ ,  $m_S = 80\text{g}$ ,  $m_P = 0\text{g}$ , and  $V = 8\text{l}$ , while the underlying parameters are shown in Tab. (1).

Table 1. Nominal model parameters  $\underline{\theta}$

$\mu_{X,m} = 0.1$	$\mu_{P,m} = 0.01$	$K_X = 1$
$K_{XI} = 0.1$	$K_P = 0.015$	$K_{PI} = 0.1$
$Y_{XS} = 5$	$Y_{PS} = 0.2$	$Y_{main} = 0.05$

### 3.2 Robust trajectory planning

For the previously described unstructured biological model three different optimal process designs were calculated to maximize the amount of product  $m_P$  at the end of the fermentation  $t_{end} = 100\text{ h}$ . Uncertainties were incorporated by assuming the model parameters  $\underline{\theta}$ , shown in Tab. 1, and the initial biomass  $m_X$  to be normally distributed, uncorrelated, with a standard deviation of 5%. Anyhow, the presented methods are also applicable for an arbitrary random parameter distribution when approximating it with a Gaussian mixture density.

The first process design, the nominal trajectory planning (N-TP), does not consider uncertainties at all but only nominal values for parameters  $\underline{\theta}$  and the initial condition  $x_0$ . Hence, the corresponding objective to minimize is  $\Phi_{N-TP} = -m_P(t_{end})$ . The second design respects the normally distributed uncertainties using a single multivariate normal distribution. This trajectory planning is referred to as 1-GMD-TP. In the third optimal planning, the uncertainties were described as a Gaussian mixture density consisting of 27 multivariate normal distributions (27-GMD). It was obtained by decomposing the univariate normal distributions of  $\mu_{X,m}$ ,  $Y_{PS}$  and  $Y_{main}$  into a sum of 3 normal densities analogously to Fig. 2. All other uncertain parameters as well as the uncertain initial biomass  $m_X(t=0)$  have not been decomposed but regarded as normally distributed. In case of the robust trajectories (1-GMD, 27-GMD), the objective  $\Phi$  also refers to the negative amount of product  $m_P$  at the end of the fermentation  $t_{end}$  but regards the mean product mass  $\bar{m}_P$  and its deviation  $\sigma_\Phi^2$  according to the robust objective (8) presented in the previous section with weighting factor  $\gamma = 1$ . The optimum is searched by variation of the (zero-order-hold) substrate feeding profile  $\mathbf{U}$ , which is restricted to  $u_{max} = 0.05\text{ l/h}$ . Other non-linear constraints to be satisfied are the maximal substrate concentration  $c_{S,max} = 10\text{ g/l}$  and the maximal reactor volume  $V_{max} = 9.5\text{ l}$ , which are incorporated in the non-linear constraint function  $\underline{r}$  according to (8). For a fair comparison with the nominal design, the factor  $\lambda$  was set to 0, such that only the mean values  $\bar{r}_i$  of the constraint function  $\underline{r}$  were regarded.

### 3.3 Optimization results

In the previous section three trajectories, that differ in the degree of how they respect uncertainties, have been introduced. The corresponding optimization results are shown in Fig. 3. The first row shows the nominal trajectory planning (N-TP), while the second and third row refer to

the 1-GMD-TP and the 27-GMD-TP, respectively. In the first column the optimal feeding profiles  $\mathbf{U}^*$  are illustrated. According to the exponential growth of the organism, the substrate feed will likewise increase exponentially through which the substrate concentration in the fermenter is held on the upper bound  $c_{S,max}$  (not shown). In this phase of the process, maximal growth is realized throughout all three designs while no product is formed. Shortly before the 60th hour of the process, the feeding profile will be adjusted such that the optimal substrate concentration for production according to  $\mu_P$  is reached. Since the biomass  $m_X$  can no longer grow under these conditions the feeding profile of the N-TP remains constant, only compensating the loss of substrate due to maintenance and consumption for product formation. In this phase, the 1-GMD-TP and the 27-GMD-TP show some slightly different irregularities in the feeding rate.

The product development over time is shown in the second column. The dashed line is the predicted mean value of product concentration. In case of the robust trajectories also the corresponding  $3\sigma$ -bounds are illustrated as solid lines. Since the N-TP does not consider uncertainties the prediction shows a linear increase of product based on the maximal production rate  $\mu_P$ . However, the real process underlies minor disturbances and if this feeding profile is simulated 10.000 times in a Monte-Carlo study using the assumed normally distributed parameters, the shaded realizations will be obtained, wherein darker areas indicate more simulated processes. In most cases, the uncertainties led to less product, in some cases almost no product was obtained. The robust trajectories predict a smaller mean product concentration since they consider the uncertainties. While the 1-GMD-TP underestimates the process fluctuations by predicting a rather small  $3\sigma$ -confidence region, the 27-GMD-TP is able to describe asymmetric densities and thus gives a realistic uncertainty prediction. However, both robust trajectories decrease the deviations significantly as shown by the 10.000 shaded realizations. Moreover, the mean values (dashed lines) are well predicted.

In the last column the Monte-Carlo realizations of the product concentrations  $c_P(t_{end})$  at the end of the process are illustrated as histograms for the different process designs. The planned amount of product of the N-TP is marked with a black circle, while the predicted probability density is shown for the robust designs. As already indicated by the shaded realizations, the N-TP shows a large spread and beyond that a bad prediction of the mean behavior. By contrast, the 1-GMD-TP is able to give a good prediction of the mean while the variance is a little underestimated due to the skewness of the histogram. This is overcome when using a 27-GMD that is able to describe such a skew distribution. As a result, the predicted probability distribution matches the histogram of realizations almost exactly. However, this is only the case, when the excitation  $u_S$  of the process has already led to a somehow normal distribution of the output. When predicting the uncertainties of the nominal trajectory (Fig. 3, first row), for which the process excitation leads to a very non-normal distribution of the product, the specific 27-GMD prediction, with only 3 further resolved uncertainties, is no longer capable of matching the histogram even though

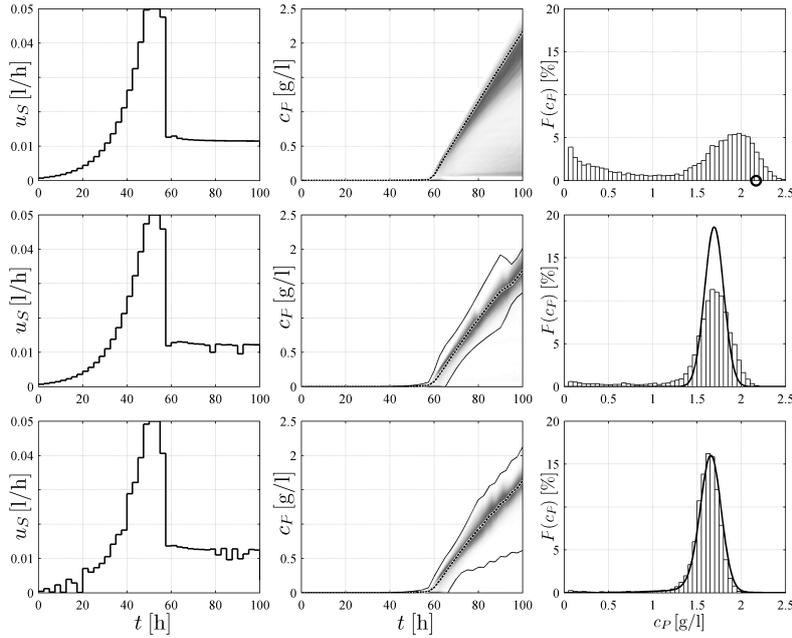


Fig. 3. Comparison of different optimal trajectories when uncertainties are respected in different degrees, refer to text.

its coarse shape could be described as shown in Fig. 4. By contrast, the prediction based on a 1-GMD can only give an approximation of the mean and the variance of the histogram, while higher order statistical information is lost due to the nature of the normal density.

### 3.4 Risk-screening

The robust trajectory plannings (1-GMD-TP, 27-GMD-TP) from the previous section are based on an objective function that also accounts for the spread of the objective using the squared  $1\sigma$ -confidence distance (8). This risk term  $\sigma_{\Phi}^2$  has to be weighted towards the mean value  $\bar{\Phi}$  of the scalar objective function  $\Phi$  with the risk factor  $\gamma$ . The results from the previous section were obtained using  $\gamma = 1$ . The influence of this weighting factor on the

optimization result will be investigated in the following paragraphs based on the 27-GMD-TP.

Fig. 5 shows the optimal substrate feeding profiles  $\mathbf{U}^*(\gamma)$  of the 27-GMD-TP when the risk factor  $\gamma$  is increased beginning from  $\gamma = 0$  in black at the very back of the plot. Even without respecting the risk by using  $\gamma = 0$  the optimal feeding profile  $\mathbf{U}^*$  shows irregularities during the constant feeding phase since the asymmetrical predicted density of the 27-GMD affects the predicted mean of the objective function  $\Phi$ . These irregularities will become slightly more intense with increasing risk factor  $\gamma$ . From risk factor  $\gamma = 3$  onwards, the scattered irregularities begin to merge into larger valleys.

Corresponding to these feeding profiles, Fig. 5 also illustrates the course of the mean product mass  $\bar{m}_P$  (gray dotted lines) as well as the upper and lower  $3\sigma$ -confidence bound as a gray scaled frame around the mean. Starting from  $\gamma = 0$  in black, it is evident that an increase of  $\gamma$  will quickly lead to a smaller spread, while the mean value  $\bar{m}_P$  is almost not affected. Only when the risk factor  $\gamma$  is increased above 3 the mean value starts to drop while the risk does not decrease significantly anymore.

This relation is shown more clearly in Fig. 6, which illustrates the  $1\sigma$ -interval of the objective  $\Phi$  over its mean value  $\bar{\Phi}$  based on different weighting factors  $\gamma$  attached to the points. The optimization results of the 1-GMD-TP are illustrated with asterix, those of the 27-GMD trajectory are shown as black dots. The pareto front of the 1-GMD-TP shows higher mean values  $\bar{\Phi}$  than the one of the 27-GMD design throughout all risk factors  $\gamma$ . This is caused by the asymmetric density of the objective function which cannot, however, be described properly with one normal distribution, only. Hence, the 1-GMD constantly overestimates the output of the process. Moreover, the two pareto fronts illustrate, how to choose the risk factor  $\gamma$

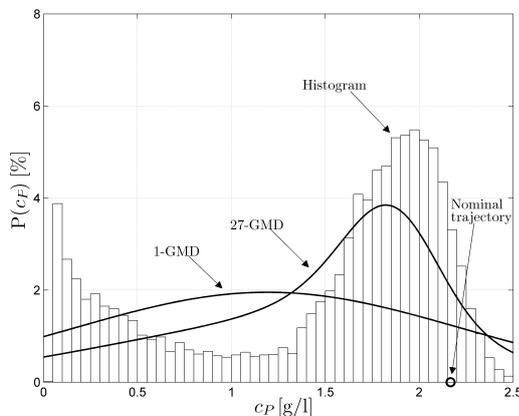


Fig. 4. Prediction of the product concentration when the nominal trajectory is simulated considering the uncertainties with a 1-GMD and a specific 27-GMD, respectively.

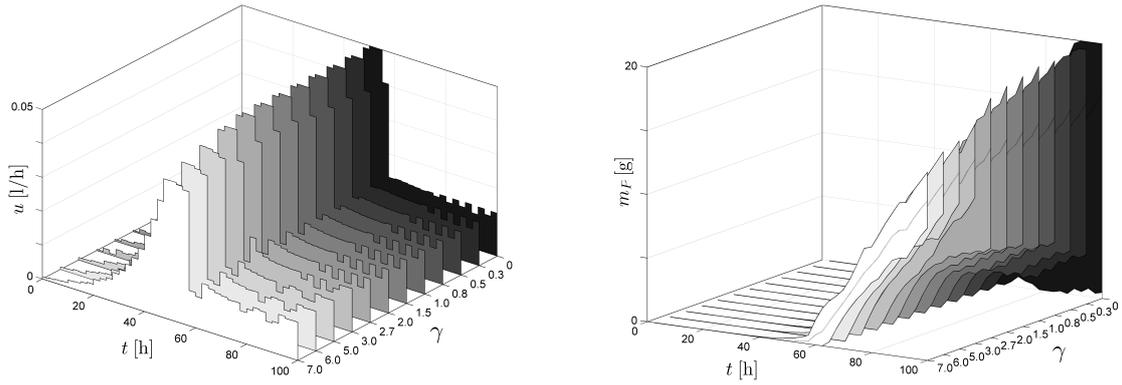


Fig. 5. Left: Optimal feeding profiles  $\mathbf{U}^*$  of the 27-GMD trajectory when different weights  $\gamma$  for the risk term  $\sigma_{\Phi}^2$  in the objective function  $\Phi$  are used. Right: Product history of the 27-GMD trajectory corresponding to the feeding profiles. The gray dotted lines refer to the mean of the product mass  $\bar{m}_P$  while the gray scaled frames indicate the  $3\sigma$ -confidence bound over time.

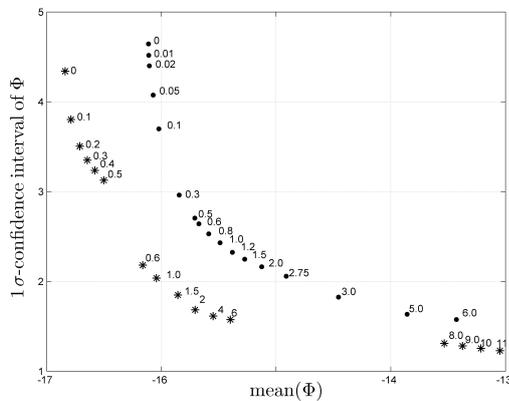


Fig. 6. Pareto front of the optimal results for different weights of the risk  $\gamma$ . The optimal results for the 1-GMD are shown using (\*) asterix. Those of the 27-GMD by (·) dots

according to the given requirements. For instance, even if a high mean value of product is highly preferred, it will make sense to choose  $\gamma = 0.1$  instead of  $\gamma = 0$  because the spread can be reduced by about 20% while the loss of mean value is only about 1% regarding the 27-GMD trajectory. On the other side, when reproducibility is highly preferred,  $\gamma = 3$  would be a reasonable choice, since a further increase of  $\gamma$  costs a lot of mean value while only reducing the risk very modestly.

#### 4. CONCLUSION

The different feeding profiles obtained from different degrees of uncertainty consideration (none, normally distributed, Gaussian mixture densities) are very similar and yet have a big impact on the deviation of the process output. The uncertainty description based on Gaussian mixture densities gives a detailed insight into the uncertainty propagation throughout the process which is essential for finding robust feeding strategies. For that reason, a process design using Gaussian mixture densities is very promising especially for biological systems that are strongly nonlinear and contain many uncertain parameters.

#### REFERENCES

- Bilmes, J.A. (1998). A gentle tutorial to the EM-algorithm. Technical report.
- Chen, H., Scherer, C., and Allgöwer, F. (1997). A game theoretic approach to nonlinear robust receding horizon control of constraint systems. ACC, 3073–3077.
- Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the RSS, 39 Series B (Methodological), 1–38.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. The Annals of Statistics, 7(1), 1–26.
- Heine, T., Kawohl, M., and King, R. (2008). Derivative-free optimal experimental design. Chemical Engineering Science, 63(19), 4873–4880.
- Julier, S. and Uhlmann, J. (2003). Unscented filtering and nonlinear estimation. Proceedings of the IEEE, 92(3).
- Kühl, P., Diehl, M., Milewska, A., Molga, E., and Bock, H. (2007). Robust NMPC for a benchmark fed-batch reactor with runaway conditions. In Assessment and Future Directions of NMPC, 455–464.
- Lee, J. and Yu, Z. (1997). Worst-case formulations of model predictive control for systems with bounded parameters. Automatica, 33(5).
- Li, P., Arellano-Garcia, H., and Wozny, G. (2008). Chance constrained programming approach to process optimization under uncertainty. Automatica, 32, 25–45.
- Magni, L., Nijmeijer, H., and van der Schaft, A. (2001). A receding-horizon approach to the nonlinear H-infinity control problem. Automatica, 37, 429–435.
- Nørgaard, M., Poulsen, N.K., and Ravn, O. (1998). Advances in derivative-free state estimation for nonlinear systems. URL <http://www2.imm.dtu.dk/pubdb/p.php?2706>.
- Romero, J. and Navarro, J. (2009). Improved efficiency in sensitivity calculations for bioreactor models. Computers and Chemical Engineering, 33, 903–910.
- van der Merwe, R. (2004). Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models. Ph.D. thesis, OGI School of Science and Engineering at Oregon Health and Science University.
- Yun, Y., Yun, H., Kim, D., and Kee, C. (2008). A gaussian sum filter approach for DGNSS integrity monitoring. The Journal of Navigation, 61, 687–703.