

Supporting the Shift towards Continuous Pharmaceutical Manufacturing by Condition Monitoring

René Schenkendorf^{1,2}

Abstract—Over the last decade there has been an increased interest in the pharmaceutical industry to shift the manufacturing process of drugs from batch to continuous operation. The continuous manufacturing of pharmaceuticals provides significant benefits, e.g. savings in cost, time and materials - to name but a few. The implementation of a continuous manufacturing strategy, however, is challenging. To gain profit from a continuous process one has to ensure its proper operation, i.e. a long time-span until the next prospective unscheduled downtime. Thus, the installed operation units have to be: 1) robust against disturbances by engineering design principles and by advanced fault tolerant control schemes, respectively; and 2) the condition of the operation units has to be monitored reliably to trigger, in case of need, appropriate intervention strategies in a timely manner. In this paper, the focus is on the monitoring aspect. Here, a model-based fault detection and identification framework is implemented, which selects the most data-supported model candidate from a set of predefined model hypotheses including the reference model (normal behavior) as well as failure models. In addition, to enable an improved diagnosis the system under study can be steered deliberately based on the proposed concept resulting into an active fault diagnosis framework. Preliminary results are demonstrated by an academic three-tank system.

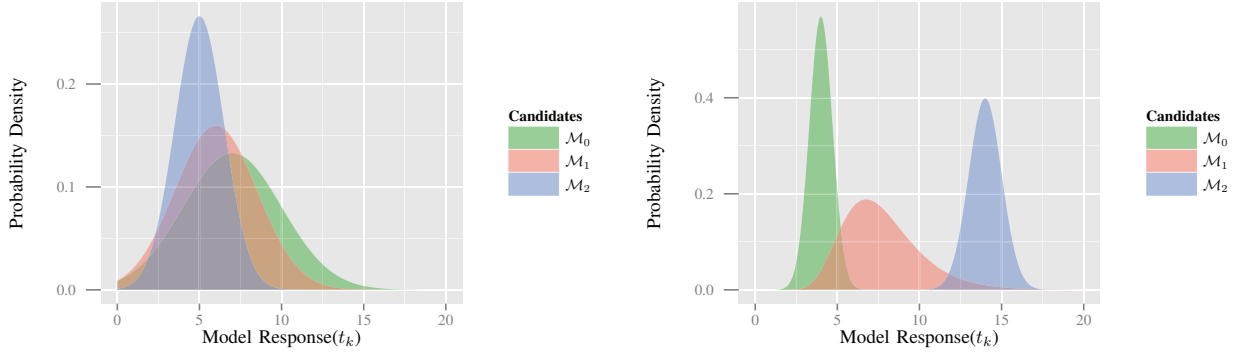
I. INTRODUCTION

Continuous pharmaceutical manufacturing (CPM) is a key technology to increase the Overall Asset Effectiveness (OAE) and to decrease the Throughput Time (TPT) in the production of pharmaceuticals [1]. Beside these processing performance indicators also regulatory requirements benefit from CPM. The individual process runs in the case of batch manufacturing require individual testing and recording of quality standards. In consequence, testing becomes a costly add-on step in pharmaceutical manufacturing. In CPM, at the same time, quality monitoring can be implemented as an inherent and economic part of the overall process control framework. The recent advances in the Process Analytical Technology (PAT) [2], i.e. in-line and on-line analytics, contribute in novel and innovative control and monitoring concepts. However, the shift of batch towards continuous operations in pharmaceutical manufacturing is challenging. In order to gain the theoretical savings in cost, time, and materials the implemented continuous processes have to be: i) designed deliberately by first engineering principles [3], ii) controlled robustly [4], [5] (e.g. by applying Fault Tolerant Control (FTC) methods), and iii) monitored reliably [6],

[7], [8] (e.g. by using Fault Detection and Isolation (FDI) or Fault Detection and Diagnosis (FDD) principles). In this paper, the focus is on the monitoring aspect primarily. In detail, a model-based FDI framework is proposed selecting the most data-supported model candidate from a predefined model set including the reference model (normal behavior) and failure models, too. Here, a suitable distance measure quantifying the differences of the model candidates as well as the distance to the measurement data while taking into account the stochastic nature of the underlying problem (e.g. due to measurement and process noise) is mandatory. In case that all relevant quantities, i.e. simulation results and measurement data, can be described by their entire probability density function (PDF), the Kullback-Leibler (KL) divergence might be an ideal distance measure [9]. In many practical situations, however, only a limited information of the relevant PDFs are available, i.e. characteristic moments of the PDF as expected values and variances. For nonlinear systems these statistical quantities can be approximated by the Kalman Filter (KF) framework. Here, the quality of the fault identification correlates significantly with the approximation power of the KF. Over the last two decades it has been demonstrated that the so-called Unscented Kalman Filter (UKF) [10] provides an excellent trade-off between computational demands and approximation power [11]. In this study, a square-root UKF version is implemented due to its numerical robustness [12] and its reported performance in chemical engineering applications [13]. Based on this set-up, conditional model probabilities can be determined for all analyzed model candidates. In addition, to enable an improved diagnosis the system under study can be steered deliberately based on the proposed concept resulting into an active fault diagnosis framework [14], [15]. Moreover, this approach can be easily extended to include non-probabilistic uncertainties [16] (e.g. zonotope / set membership ideas) and to address critical state constraints [17] as well. The remainder of this paper is structured as follows. In Section II some comments about the used distance measure are given. In Section III the basics of the KF framework are summarized including a detailed description of the applied square-root UKF version. The expected benefit of an active FDI is highlighted in Section IV. The overall concept of the proposed active fault diagnosis framework is demonstrated by a three-tank system in Section V. Finally, the conclusions are drawn in Section VI.

¹Institute of Energy and Process Systems Engineering, Technical University of Braunschweig, 38106 Braunschweig, Germany
r.schenkendorf@tu-braunschweig.de

²Center of Pharmaceutical Engineering (PVZ), 38106 Braunschweig, Germany



(a) Undesired situation in model-based FDI.

(b) Desired situation in model-based FDI.

Fig. 1: Probability density functions of three potential model outcomes at time point t_k .

II. DISTANCE MEASURE

When applying model-based fault identification strategies the inherent uncertainties in simulation results (due to uncertain model parameters and initial conditions) and gathered data have to be considered. In practice, there might be scenarios at which different model candidates cannot be distinguished (Fig. 1a) as well as situations at which a proper model selection / fault identification is straightforward - even in the case of significant uncertainties (Fig. 1b). To distinguish both cases and to enable a concept which aims at a reliable fault detection an appropriate distance measure has to be derived which takes these uncertainties into account. Abstracted from information theory the KL distance quantifies the differences of two PDFs. In principle for each model candidate, \mathcal{M}_i , from the set of available model hypotheses, $M \triangleq \{\mathcal{M}_0, \mathcal{M}_1, \dots, \mathcal{M}_m\}$, a PDF of the model outcomes might be calculated. For example, assuming two PDFs, p_i and p_j , the KL distance is expressed by:

$$d_{KL}(p_i||p_j) = \int_{R^n} p_i(x) \ln \frac{p_i(x)}{p_j(x)} dx \quad (1)$$

Here, it has to be stressed that the KL distance, typically, is not symmetric ($d_{KL}(f||g) \neq d_{KL}(g||f)$). Alternatively, a symmetric distance measure can be derived according to:

$$D_s(p_i, p_j) = d_{KL}(p_i||p_j) + d_{KL}(p_j||p_i) \quad (2)$$

In case of two Gaussian distributions Eq. 2 simplifies to the Kullback's total measure of information [18]:

$$D_s^g(p_i, p_j) = \frac{1}{2} \left[\frac{(\sigma_i^2 - \sigma_j^2)^2}{(\sigma^2 + \sigma_i^2)(\sigma^2 + \sigma_j^2)} + (E[y_i] - E[y_j])^2 \cdot \left(\frac{1}{\sigma^2 + \sigma_i^2} + \frac{1}{\sigma^2 + \sigma_j^2} \right) \right], \quad (3)$$

where σ_i^2 (σ_j^2) and $E[y_i]$ ($E[y_j]$) are the variance and the expected value of the model output of \mathcal{M}_i (\mathcal{M}_j), respectively. σ^2 represents the variance of the measurement data and is assumed to be known. When combining the Kullback's total

measure of information with model probabilities, $\Pi(\mathcal{M}_j)$, the maximum change in Shannon's Entropy (SE) [19] can be derived:

$$\Delta(SE) = \sum_{i=1}^m \sum_{j=i+1}^m \Pi(\mathcal{M}_i) \Pi(\mathcal{M}_j) D_s^g(p_i, p_j) \quad (4)$$

Here, the model probabilities are determined by utilizing the Bayes' theorem:

$$\Pi(\mathcal{M}_i|y_d^+) = \frac{p_{y_d^+|\mathcal{M}_i} \Pi(\mathcal{M}_i|y_d^-)}{\sum_{j=1}^m p_{y_d^+|\mathcal{M}_j} \Pi(\mathcal{M}_j|y_d^-)}; \quad \forall i = 1, \dots, m \quad (5)$$

where y_d^+ represents the latest data and y_d^- the previous measurement data sample.

Up to this point, this section provides necessary tools for: i) determining the likelihood of the model candidates (Eq. 5), ii) measuring the differences between model hypotheses (Eqs. 2 & 3), and iii) calculating the maximum change in SE (Eq. 4), i.e. the increase in discrimination power when new data/information are available. Obviously, to determine these quantities numerically some statistics of the simulation results have to be provided. Here, the square-root UKF is tailor-made in terms of computational demand and approximation power [20].

III. KALMAN FILTER

Assuming a discrete-time nonlinear system in its state-space form the set of governing equations reads as:

$$x_{k+1} = f(x_k, u_k) + w_k \quad (6)$$

$$y_{k+1} = h(x_{k+1}) + v_{k+1}, \quad (7)$$

where $x \in R^{n_x}$ denotes the state variable vector, $y \in R^{n_y}$ means the observed model output vector. (Note that this work assumes $n_y = 1$ but the proposed framework can be extended to multi-output cases as well.) The measurement noise and the process noise are considered by w_k and v_{k+1} , respectively. The system input is expressed by $u \in R^{n_u}$.

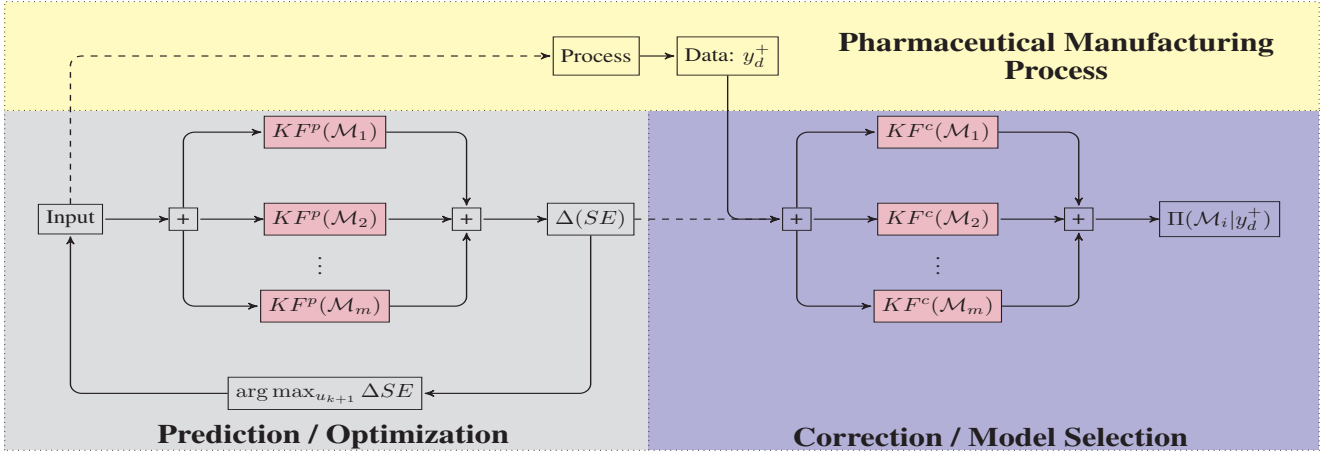


Fig. 2: Fault Identification Work-flow: The prediction step of the Kalman Filter, KF^p , is used to identify an optimal fault revealing input. This optimal input is applied to the real system some informative data, y_d^+ , used within the Kalman correction step, KF^c . Based on the resulting statistics the conditioned model probabilities are derived, i.e. fault-related model candidates are selected.

The two vector functions, $f(\cdot)$ & $h(\cdot)$, are known as the state and the output function, respectively. For nonlinear problems as in this study, it has been shown in many benchmark studies that the UKF outperforms the Extended Kalman Filter (EKF) in terms of approximation power. Compared to CPU-intensive Particle Filters the UKF generates and propagates deliberately chosen Sample Points (i.e. position \mathcal{X}_i and weights W_i) to quantify relevant statistics of the prediction and correction step within the KF framework. To improve the numerical stability additionally, the so-called square-root KF (SR-UKF) was introduced providing guaranteed positive definite co-variance matrices. Algorithmically, the prediction and correction steps of SR-UKF are given in Alg. 1. Here, qr means the QR decomposition, $cholupdate$ refers to Cholesky factor updating, and the “/”-operator indicates an efficient least squares approach utilizing QR decomposition with pivoting. In doing so, the expected values, $\hat{x}^{(-)}$, as well as the co-variance matrices, $P^{(-)} = S^{(-)}S^{(-)T}$, can be derived and used to select the most plausible model candidate and to design an informative, fault-revealing system input as described in the next section. As a side note, the conditional probability densities, $p_{y_d^+|\mathcal{M}_i}$ in Eq. 5, are approximated according to:

$$p_{y_d^+|\mathcal{M}_i} \approx \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{1}{2\sigma_i^2}(y_d^+ - \hat{y}_k^+(\mathcal{M}_i))^2} \quad (8)$$

IV. ACTIVE FAULT DIAGNOSIS

In the field of fault detection and diagnosis one distinguishes between passive and active strategies, i.e. no intervention and applying auxiliary input signals, respectively. In this study, the proposed concept contributes to both. The active FDI technique, however, needs some further explanation. The derived auxiliary inputs aim at exciting the process under study deliberately to facilitate the actual fault detection and identification performance. In practice,

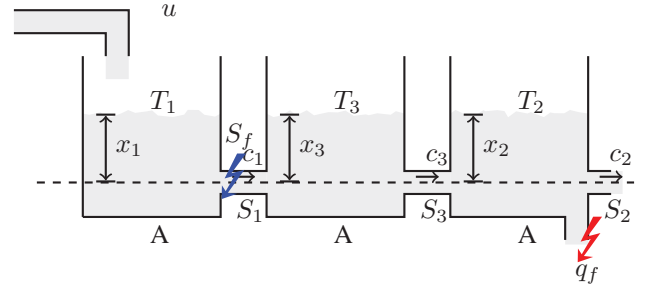


Fig. 3: Scheme of the three-tank system (adapted from [21]) with the two considered fault scenarios: i) leakage at tank T_2 indicated by ζ and ii) pipe obstruction (e.g. by an obstacle or fouling processes) between T_1 and T_3 indicated by ζ .

the calculation of such an informative fault-revealing input can be expressed as an optimisation / optimal control problem. One of the common methods for solving optimal control problems is the so-called direct sequential approach [22]. Here, the idea is to parametrize a control vector (e.g. auxiliary input) in combination with the numerical integration of the model equations. Here, these forward simulations are part of prediction step of the KF and lead to the determination of the expected maximum change in SE (Eq. 4). For instance, assuming a one step ahead predictive time step an optimal constant auxiliary input can be derived according to:

$$\begin{aligned} u^* &= \arg \max_{u_{k+1}} \Delta SE \\ s.t. & \hat{x}_{k+1}^-(\mathcal{M}_i), \forall i \in M \\ & P_{k+1}^-(\mathcal{M}_i), \forall i \in M \\ & u_{k+1} \in \mathcal{U} \end{aligned} \quad (9)$$

Algorithm 1 SR-UKF

```

1: procedure INITIALIZE
2:    $\hat{x}_0 = E[x_0]$ 
3:    $S_0 = chol\{E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]\}$ 
4: procedure SIGMA POINT GENERATION
5:    $\hat{x}_0 = E[x_0]$ 
6:    $S_0 = chol\{E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]\}$ 
7: procedure PREDICTION STEP ( $KF^p$ )
8:    $\mathcal{X}_{k-1} = \hat{x}_{k-1} + (S_{k-1})^T \mathcal{X}_{ini}$ 
9:    $\mathcal{X}_{k|k-1}^* = f(\mathcal{X}_{k-1})$ 
10:   $\hat{x}_k^- = \sum_{i=0}^{n+1} W_i^{(m)} \mathcal{X}_{i,k|k-1}^*$ 
11:   $S_k^- = qr \left\{ \left[ \sqrt{W_1} \left( \mathcal{X}_{1:n+1,k|k-1}^* - \hat{x}_k^- \right) \sqrt{Q} \right] \right\}$ 
12:   $S_k^- = cholupdate \{S_k^-, \mathcal{X}_{0,k} - \hat{x}_k^-, W_0\}$ 
13:   $\mathcal{X}_{k-1} = \hat{x}_{k-1} + (S_k^-)^T \mathcal{X}_{ini}$ 
14:   $\mathcal{Y}_{k|k-1} = h[\mathcal{X}_{k-1}]$ 
15:   $\hat{y}_k^- = \sum_{i=0}^{n+1} W_i^{(m)} \mathcal{Y}_{i,k|k-1}$ 
16:   $S_{y_k} = qr \left\{ \left[ \sqrt{W_1} (\mathcal{Y}_{1:n+1,k} - \hat{y}_k^-) \sqrt{R} \right] \right\}$ 
17:   $S_{y_k} = cholupdate \{S_{y_k}, \mathcal{Y}_{0,k} - \hat{y}_k^-, W_0\}$ 
18:   $P_{x_k, y_k} = \sum_{i=0}^{n+1} W_i^{(c)} [\mathcal{X}_{i,k|k-1} - \hat{x}_k^-] [\mathcal{Y}_{i,k|k-1} - \hat{y}_k^-]^T$ 
19: procedure CORRECTION STEP ( $KF^c$ )
20:   $K_k = (P_{x_k, y_k} / S_{y_k}^T) / S_{y_k}$ 
21:   $\hat{x}_k = \hat{x}_k^- + K_k (y_k - \hat{y}_k^-)$ 
22:   $U = K_k S_{y_k}$ 
23:   $S_k = cholupdate \{S_k^-, U, 1\}$ 

```

Here, it should be stressed that the model probabilities (Eq. (5)) impacts the optimization problem significantly, i.e. the optimized input sequence focus on the most likely model candidates and attempts to separate these. The overall work-flow of the proposed concept for fault identification is illustrated in Fig. 2.

Cross-sectional area of individual tanks	$A = 0.0154 \text{ m}^2$
Cross-sectional area of connecting pipes	$S_i = 5 \times 10^{-5} \text{ m}^2,$ $i = 1, 2, 3$
Cross-sectional area reduction factor	$S_f \in \{1, 0.5\}$
Leakage flow	$q_f \in \{0, 0.002\} \text{ m}^3/\text{s}$
Gravitational acceleration	$g = 9.81 \text{ m/s}^2$
Nondimensional outflow coefficients	$c_i = 1, i = 1, 2, 3$

TABLE I: Parameters of the three-tank model.

V. DEMONSTRATION

For the purpose of demonstration an academic three-tank system (Fig. 3) is used. Despite its simplicity, this system shows some common characteristics of continuous

pharmaceutical manufacturing processes: i) continuous operation and ii) functional connection of sub-systems forming a manufacturing process chain. Here, the tanks T_1 , T_2 , and T_3 are connected while the liquid inlet flow, u , enters T_1 and a potential leakage flow, q_f , is defined for T_2 . Based on the mass conservation law and the Torricelli's law following set of differential equations can be derived:

$$\begin{aligned}
 \dot{x}_1 &= \frac{-c_1 S_1 S_f \text{sign}(x_1 - x_3) \sqrt{2g|x_1 - x_3|} + u}{A} \\
 \dot{x}_2 &= \frac{-c_3 S_3 \text{sign}(x_2 - x_3) \sqrt{2g|x_2 - x_3|}}{A} \\
 &\quad - \frac{c_2 S_2 \sqrt{2gx_2} - q_f}{A} \\
 \dot{x}_3 &= \frac{c_1 S_1 S_f \text{sign}(x_1 - x_3) \sqrt{2g|x_1 - x_3|}}{A} \\
 &\quad - \frac{c_3 S_3 \text{sign}(x_3 - x_2) \sqrt{2g|x_3 - x_2|}}{A}
 \end{aligned} \tag{10}$$

Thus, the dynamic of liquid levels is given by the system states x_1 , x_2 , and x_3 . The corresponding model parameters are summarized in Tab. I. In what follows, two different fault scenarios are analyzed: i) \mathcal{M}_1 , tank leakage at T_2 ($q_f = 0.0025$) and ii) \mathcal{M}_2 , pipe obstruction between T_1 and T_3 ($S_f = 0.5$). The fault-free reference model, \mathcal{M}_0 completes the set of potential model candidates. Starting with \mathcal{M}_0 the faults are induced at second 120. Challenged with simulated measurement data (sample frequency of 10 sec) the proposed concept aims to identify the correct model hypothesis, i.e. to detect the fault as soon as possible. As shown in Fig. 4a & 4c, the passive as well as the active FDI strategy identifies the leakage. As expected a nominal input sequence ($u(t_k) = 0.75$) is less informative and fault-revealing compared to an optimal input sequence ($u(t_k) \in [0.25, 2.50]$). The leakage, \mathcal{M}_1 , is rapidly identified in case of the active FDI scenario. Similar results can be derived for the pipe obstruction, \mathcal{M}_2 , as shown in Fig. 4b & 4d. Here, the temporary selection of \mathcal{M}_1 can be explained by a short match of related simulation results and incoming measurement data (Fig. 5a). The associated optimal input sequence (Fig. 5b), however, reveals the pipe obstruction in the long run correctly. Here, too, the active FDI setting enables an early failure detection by the optimized input sequence.

VI. CONCLUSIONS

This paper presents a framework for passive as well as active FDI operation based on a SR-UKF and information theory principles. By these actions, the inherent uncertainties in modeling and gathered measurement data are addressed comprehensively. By a simulation study it has been shown that the induced faults can be detected (passive & active FDI) reliably. As expected, the active FDI enables a better detection rate and may help to mitigate undesired consequences in terms of plant damages and pollution of the environment by early interventions. Future work will focus

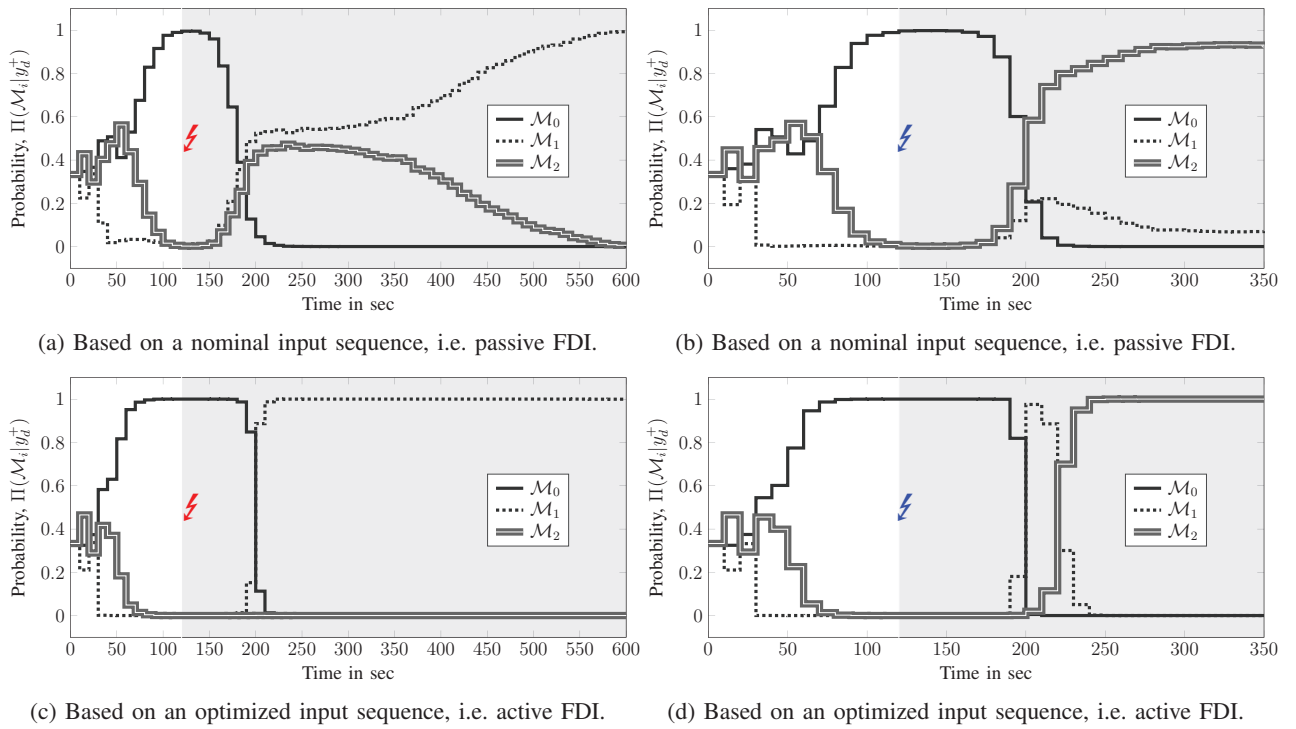


Fig. 4: Model probabilities starting with an uninformative a-priori set-up, i.e. $\Pi(\mathcal{M}_i) = 1/3$, $i = 1, 2, 3$. The faults, leakage ⚡ and pipe obstruction ⚡, are induced at second 120.

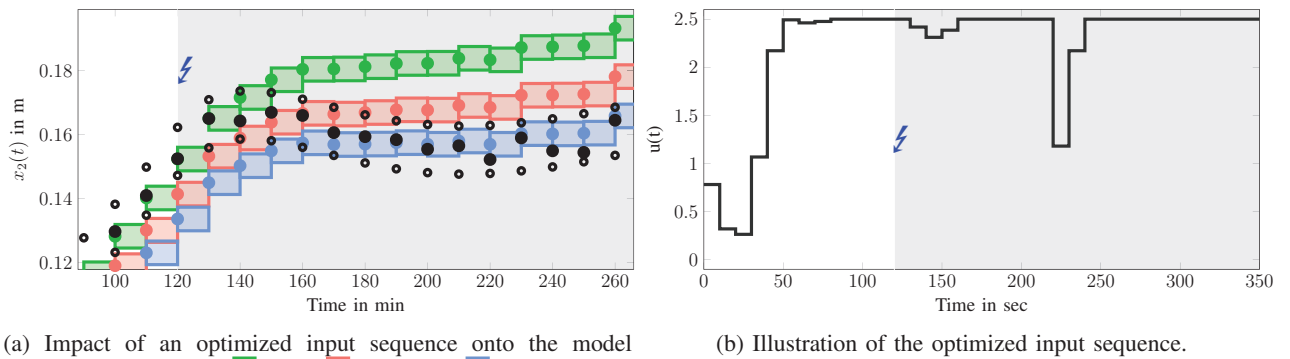


Fig. 5: Illustration of the improved fault diagnosis (5a) in consequence of the fault-revealing optimized input sequence (5b) based on the proposed active FDI concept.

on relevant, dedicated processing equipment in CPM, e.g. continuous flow reactors. In particular, following aspects have to be considered for a concept expected to be of practical relevance:

- real time optimization to enable an active FDI implementation
- identification of multiple faults in parallel
- novelty detection in case of untrained/non-modeled faults.

In conjunction with the incorporation of non-probabilistic uncertainties (e.g. zonotope / set membership concepts) we are quite sure that we can provide workable as well as flexible tools supporting the shift towards CPM in the near future.

ACKNOWLEDGMENT

The author thanks V. N. Emenike and X. Xie for helpful discussions.

REFERENCES

- [1] W. Bisson, "Novartis approach to continuous manufacturing (api and drug product)," in *IFPAC 10 - Cortona, Italy*, 2010.
- [2] S. Bordawekar, A. Chanda, A. M. Daly, A. W. Garrett, J. P. Higgins, M. A. LaPack, T. D. Maloney, J. Morgado, S. Mukherjee, J. D. Orr, G. L. Reid III, B.-S. Yang, and H. W. Ward II, "Industry perspectives on process analytical technology: Ttool and applications in api manufacturing," *Organic Process Research & Development*, 2015.
- [3] V. N. Emenike and U. Krewer, "Model-based optimal design of continuous flow reactors for the synthesis of active pharmaceutical ingredients," in *ProcessNet-Jahrestagung und 32. DECHEMA-Jahrestagung der Biotechnologen 2016*, submitted.
- [4] A. Mesbah, A. N. Ford Versypt, X. Zhu, and R. D. Braatz, "Nonlinear model-based control of thin-film drying for continuous pharmaceutical manufacturing," *Ind. Eng. Chem. Res.*, vol. 53, no. 18, pp. 7447–7460, 2013.
- [5] A. Mesbah, J. A. Paulson, R. Lakerveld, and R. D. Braatz, "Plant-wide model predictive control for a continuous pharmaceutical process," in *2015 American Control Conference, Palmer House Hilton, Chicago, IL, USA*, 2015.
- [6] R. Isermann, "Model-based fault-detection and diagnosis - status and applications," *Annu. Rev. Control*, vol. 29, no. 1, pp. 71–85, 2005.
- [7] S. Kristen, C. Paphonwit, and R. D. Braatz, "Perspectives on fault detection and identification in industrial systems," in *9th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, Paris, France*, 2015.
- [8] M. Vileiniskis, R. Remenyte-Rescott, D. Rama, and J. Andrews, "Fault detection and diagnostics of a three-phase separator," *Journal of Loss Prevention in the Process Industries*, vol. 41, pp. 215–230, 2016.
- [9] K. P. Burnham and D. R. Anderson, *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*. Springer, 2002.
- [10] S. J. Julier and J. K. Uhlmann, "A general method for approximating nonlinear transformations of probability distributions," Dept. of Engineering Science, University of Oxford, Tech. Rep., 1996.
- [11] D. Simon, *Optimal state estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley-Interscience, 2006.
- [12] R. van der Merwe, "Sigma-point kalman filters for probabilistic inference in dynamic state-space models," Ph.D. dissertation, OGI School of Science & Engineering at Oregon Health & Science University, 2004.
- [13] M. Mangold, A. Bck, R. Schenkendorf, C. Steyer, A. Voigt, and K. Sundmacher, "Two state estimators for the barium sulfate precipitation in a semi-batch reactor," *Chemical Engineering Science*, vol. 64, no. 4, pp. 646 – 660, 2009, 3rd International Conference on Population Balance Modelling.
- [14] A. Mesbah, S. Streif, R. Findeisen, and R. D. Braatz, "Active fault diagnosis for nonlinear systems with probabilistic uncertainties," in *Preprints of the 19th World Congress, The International Federation of Automatic Control, Cape Town, South Africa*, 2014.
- [15] K.-K. K. Kim, D. M. Raimondo, and R. D. Braatz, "Optimal input design for fault detection and diagnosis: Model-based prediction and statistical distance measures," in *2013 European Control Conference (ECC), Zürich, Switzerland*, 2013.
- [16] C. Combastel, "Merging kalman filtering and zonotopic state bounding for robust fault detection under noisy environment," *IFAC-PapersOnLine*, vol. 48, no. 21, pp. 289–295, 2015.
- [17] B. O. S. Teixeira, L. A. B. Torres, L. A. Aguirre, and D. S. Bernstein, "On unscented kalman filtering with state intercal constraints," *Journal of Process Control*, 2010.
- [18] G. E. P. Box and W. J. Hill, "Discrimination among mechanistic models," *Technometrics*, vol. 9, pp. 57–71, 1967.
- [19] R. F. Stengel, *Optimal Control and Estimation*. Dover Publications, 1994.
- [20] R. Schenkendorf and M. Mangold, "Online model selection approach based on unscented kalman filtering," *Journal of Process Control*, vol. 23, pp. 44–47, 2013.
- [21] X. Z. Zhang, M. M. Polycarpou, and T. Parisini, "A robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 4, 2002.
- [22] L. T. Biegler, *Nonlinear Programming*. Society for Industrial and Applied Mathematics, 2010.