ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/263942618

How To NOT Make the Extended Kalman Filter Fail

Article *in* Industrial & Engineering Chemistry Research · February 2013

DOI: 10.1021/ie300415d

CITATIONS
15

READS
653

2 authors:

René Schneider
École Polytechnique Fédérale de Lausanne
13 PUBLICATIONS **32** CITATIONS
SEE PROFILE

Christos Georgakis
Tufts University
176 PUBLICATIONS **3,813** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

State and Parameter Estimation for Large-Scale and Distributed Systems View project

Real-Time Optimization of Uncertain Interconnected Systems View project

All content following this page was uploaded by René Schneider on 16 February 2017.

The user has requested enhancement of the downloaded file. All in-text references underlined in blue are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

How to NOT Make the Extended Kalman Filter Fail

René Schneider[†] and Christos Georgakis[*,‡]

AVT Process Systems Engineering, RWTH Aachen University, Aachen, Germany, and Chemical and Biological Engineering, Tufts University, Medford, USA

E-mail: christos.georgakis@tufts.edu

Abstract

In an effort to assess the performance of newer estimation algorithms, many prior publications have presented comparative studies where the Extended Kalman Filter (EKF) failed. This is because the EKF's design parameters were sometimes chosen arbitrarily and with little consideration of their role and impact on filter performance. This paper demonstrates in a tutorial way that EKF failure can often be avoided by a systematic design of its parameters, i.e. its covariance matrices. Particular focus is on the systematic selection of proper initial state and process noise covariance matrices. As in practice, model parameters are first identified from plant measurements. The covariance of these model parameters is subsequently required to calculate the proposed time-varying process noise covariance matrix for the EKF. Finally, the proposed EKF design is evaluated on a popular reactor example process from the literature and converges in all simulation runs. From this extensive set of simulations and by comparison with several out-of-the-box versions of the constrained Unscented Kalman Filter (UKF), we conclude that there exists a systematic way to design the EKF for a very satisfactory performance. Therefore, the EKF's design parameters need not be tuned *ad hoc* by trial and error.

[*]To whom correspondence should be addressed
[†]AVT Process Systems Engineering, RWTH Aachen University, Aachen, Germany
[‡]Chemical and Biological Engineering, Tufts University, Medford, USA

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

Introduction

Many modern model-based control strategies require an estimate of all the dependent variables of the model, called the state variables, that characterize the state of the process at all times. This helps our understanding of the inner workings of the process and greatly facilitates the design of effective multivariable and nonlinear control strategies. Such estimation utilizes the available output measurements, $\boldsymbol{y}(t)$, the process model and the measured inputs, $\boldsymbol{u}(t)$, of the process. For the class of linear systems the optimal state estimation problem has been resolved by the introduction of the famous Kalman Filter.[1] For nonlinear processes, the most common approximate solution to the optimal state estimation problem is the Extended Kalman Filter (EKF).[2] Because the EKF is not the optimal nonlinear estimator a number of alternative filters have been proposed, e.g. Unscented (UKF) or Ensemble Kalman Filters (EnKF).[3] Also, the substantial interest over the last two decades on constrained controllers, in particular the Model Predictive Controller (MPC), has motivated the postulation of a constrained state estimator, the Moving Horizon Estimator (MHE).[4] These additional approaches are of interest, especially if they substantially improve on the performance of the EKF. Unfortunately, comparisons aimed at showing the advantages of such alternative algorithms are not always fair. While novel algorithms are typically carefully tuned, it appears as if proper EKF design is frequently disregarded: It often seems that the EKF's design matrices are chosen more or less arbitrarily. Conversely, the need to select these design matrices – $\boldsymbol{P}_0$, $\boldsymbol{Q}$ and $\boldsymbol{R}$ – in a systematic manner rather than *ad hoc* by trial and error has been discussed more than ten years ago by Valappil and Georgakis.[5] It is our feeling that this important point cannot be stressed enough and that its sensible consideration will result in both more fair and meaningful future comparisons of estimation algorithms. Consequently, the present paper serves as a tutorial one and does not present any new design methodologies beyond those of Valappil and Georgakis.[5] However, it supports this design methodology by presenting the results of several simulation sets, each of which consists of 1,000 to 100,000 simulation runs.

Because of the tutorial nature of this paper, we will focus on a single example process that has been examined by Haseltine and Rawlings[6] and was referred to by many other authors. It is a

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

reaction system described by three state variables, representing the concentrations of three species involved in two reversible gas-phase reactions in an isothermal, well-mixed batch reactor with constant volume. In the paper cited, the reactor benchmark problem was used to demonstrate the superiority of a modern MHE scheme with constraints on the states over the rather traditional unconstrained EKF. While the constrained MHE algorithm is an interesting and promising approach, the comparison was unfair on two counts which we will detail below. Unfortunately, the poor performance of the EKF with this particular configuration was often reiterated as a generic result by subsequent authors.[7–9] Consequently, a widespread wrong perception of what an EKF can or cannot do needs to be corrected so that it does not end up being accepted as a fact.

The aim of this contribution is to point out that there exist systematic ways to design the EKF and its design parameters need not be chosen by trial and error. To achieve this, we examine in detail the above mentioned reactor benchmark problem and show that the EKF can fulfill the estimation task quite successfully, provided it is designed properly. To that end, the design methodology proposed by Valappil and Georgakis[5] is applied. Finally, we compare these results with the performance of several constrained and unconstrained out-of-the-box UKF algorithms as presented by Kolås et al.[10]

Extended Kalman Filter (EKF)

Consider a general nonlinear time-invariant system in continuous time that generates measurements at discrete time steps $t_k = k\Delta t$:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}\left(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{p}\right) + \boldsymbol{w}(t), \tag{1}$$

$$\boldsymbol{y}(t_k) = \boldsymbol{h}\left(\boldsymbol{x}(t_k)\right) + \boldsymbol{v}(t_k), \tag{2}$$

where $\boldsymbol{x} \in \mathbb{R}^{n_x}$, $\boldsymbol{u} \in \mathbb{R}^{n_u}$, $\boldsymbol{p} \in \mathbb{R}^{n_p}$, $\boldsymbol{w} \in \mathbb{R}^{n_w}$, $\boldsymbol{v} \in \mathbb{R}^{n_v}$ and $\boldsymbol{y} \in \mathbb{R}^{n_y}$ are the states, the inputs, the time-invariant parameters, the process and measurement noise variables and the measurements of the system, respectively. Both process and measurement noises are assumed to be uncorrelated

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

zero-mean Gaussian random processes, i.e.

$$\begin{aligned} E[\boldsymbol{w}(t)] = E[\boldsymbol{v}(t_k)] &= \mathbf{0}, \\ E[\boldsymbol{w}(t)\boldsymbol{w}^T(\tau)] &= \boldsymbol{Q}(t)\delta(t-\tau), \\ E[\boldsymbol{v}(t_k)\boldsymbol{v}^T(t_k)] &= \boldsymbol{R}(t_k), \\ E[\boldsymbol{w}(\tau)\boldsymbol{v}^T(t_k)] &= \mathbf{0}. \end{aligned}$$

For this class of systems with continuous time evolution of states but discrete measurements, the so-called *hybrid* EKF can estimate the states from the noisy measurements through the following set of equations that are taken from Simon.[11] First, the *a posteriori* state estimates $\hat{\boldsymbol{x}}(t_{k-1}^+)$ and covariance matrices $\boldsymbol{P}(t_{k-1}^+)$ are propagated one sampling period forward in time, i.e. from time instant t_{k-1}^+ to t_k^-, yielding the *a priori* state estimate and estimated output

$$\begin{aligned} \hat{\boldsymbol{x}}(t_k^-) &= \hat{\boldsymbol{x}}(t_{k-1}^+) + \int_{t_{k-1}^+}^{t_k^-} \boldsymbol{f}(\hat{\boldsymbol{x}}(t), \boldsymbol{u}(t), \boldsymbol{p})\, d\tau, \\ \hat{\boldsymbol{y}}(t_k^-) &= \boldsymbol{h}(\hat{\boldsymbol{x}}(t_k^-)), \end{aligned}$$

and the state covariance matrix

$$\boldsymbol{P}(t_k^-) = \boldsymbol{P}(t_{k-1}^+) + \int_{t_{k-1}^+}^{t_k^-} \left(\boldsymbol{A}(\tau)\boldsymbol{P}(\tau) + \boldsymbol{P}(\tau)\boldsymbol{A}^T(\tau) + \boldsymbol{Q}(\tau)\right) d\tau, \tag{3}$$

immediately before the next measurement arrives at time t_k. In Eq. (3), the linearization matrix $\boldsymbol{A}$ is defined as

$$\boldsymbol{A}(t) = \left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right)_{\hat{\boldsymbol{x}}(t), \boldsymbol{u}(t), \boldsymbol{p}}.$$

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

We further define the matrix $\boldsymbol{C}$ point-wise at all measurement occurences t_k as

$$\boldsymbol{C}(t_k) = \left(\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\right)_{\hat{\boldsymbol{x}}(t_k^-)}.$$

The Kalman filter gain $\boldsymbol{K}$ is then computed as

$$\boldsymbol{K}(t_k) = \boldsymbol{P}(t_k^-)\boldsymbol{C}(t_k)^T \left(\boldsymbol{C}(t_k)\boldsymbol{P}(t_k^-)\boldsymbol{C}(t_k)^T + \boldsymbol{R}(t_k)\right)^{-1}$$

and subsequently used to update the *a priori* state estimates and the covariance matrix to the *a posteriori* values as soon as a new measurement $\boldsymbol{y}(t_k)$ becomes available:

$$\begin{aligned}
\hat{\boldsymbol{x}}(t_k^+) &= \hat{\boldsymbol{x}}(t_k^-) + \boldsymbol{K}(t_k)\left(\boldsymbol{y}(t_k) - \boldsymbol{h}(\hat{\boldsymbol{x}}(t_k^-))\right), \\
\boldsymbol{P}(t_k^+) &= \left(\boldsymbol{I} - \boldsymbol{K}(t_k)\boldsymbol{C}(t_k)\right)\boldsymbol{P}(t_k^-)\left(\boldsymbol{I} - \boldsymbol{K}(t_k)\boldsymbol{C}(t_k)\right)\boldsymbol{P}(t_k^-)^T + \boldsymbol{K}(t_k)\boldsymbol{R}(t_k)\boldsymbol{K}^T(t_k).
\end{aligned}$$

Note that we use the so-called Joseph stabilized version of the covariance measurement update equation which guarantees that $\boldsymbol{P}(t_k^+)$ will always be symmetric and positive definite, as long as $\boldsymbol{P}(t_k^-)$ is symmetric positive definite. In this sense it is more stable and numerically robust than the standard update equation $\boldsymbol{P}(t_k^+) = (\boldsymbol{I} - \boldsymbol{K}(t_k)\boldsymbol{C}(t_k))\boldsymbol{P}(t_k^-)$.[11]

Defining the Design Parameters of the EKF

Whenever one wants to apply the EKF to a particular problem at hand, one must select the measurement noise covariance $\boldsymbol{R}$, the initial state estimate $\hat{\boldsymbol{x}}(t_0^+)$ and the corresponding covariance $\boldsymbol{P}_0$ and the process noise covariance $\boldsymbol{Q}$. In this section, we briefly elaborate on the influence of each of these variables on the behaviour of the EKF.

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

Measurement Noise Covariance

For an appropriate choice of the measurement noise covariance matrix $\boldsymbol{R}$, multiple methods have been developed and are discussed in a recent review.[12] For instance, a few authors try to derive this information directly from process data.[13] However, most scientific publications assume that the measurement noise covariance is time-invariant and readily available from the measurement device's manufacturer.

Suppose that σ_i is the known time-invariant nominal standard deviation of the i-th measurement signal, then a simple way to specify the measurement noise covariance matrix $\boldsymbol{R}$ would be

$$\boldsymbol{R} = k_R \cdot diag(\sigma_i^2), \tag{4}$$

where k_R is typically equal to or larger than one. This approach yields a diagonal and time-invariant measurement noise covariance matrix $\boldsymbol{R}$. If one expects a larger measurement uncertainty than σ_i, one can increase the factor k_R beyond one, thus making the filter more robust to measurement errors. However, if $\boldsymbol{R}$ or k_R is chosen too large, the filter underweights the effect of the measurements, thus reducing the filter feedback. In the limit case of vanishing Kalman gain matrix $\boldsymbol{K}$, caused by the large values of $\boldsymbol{R}$ or k_R, the filter transforms into an open-loop predictor. In the presence of plant-model mismatch or if the plant is unstable, this may lead to an offset or even to the divergence of the filter estimates from the true states.

For the remainder of this text, however, it will be assumed that the measurement noise covariance is an exactly known, time-invariant diagonal matrix.

Initial State Estimate

To initialize the filter, both a suitable initial state guess $\hat{\boldsymbol{x}}_0 \equiv \hat{\boldsymbol{x}}(t_0^+)$ is needed as well as information about the uncertainty of this estimate in the form of an initial state estimate covariance matrix $\boldsymbol{P}_0 \equiv \boldsymbol{P}(t_0^+)$. In contrast to the measurement noise covariance, information about the uncertainty of $\hat{\boldsymbol{x}}_0$ cannot be measured experimentally. Therefore, the filter designer needs to estimate the degree

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

of uncertainty of his particular choice for $\hat{\boldsymbol{x}}_0$. In general, the filter will not be fatally affected if $\hat{\boldsymbol{x}}_0$ is not close to $\boldsymbol{x}_0$ but convergence to the correct estimate may be slow. However, if $\boldsymbol{P}_0$ is chosen too small or optimistic while $\hat{\boldsymbol{x}}_0$ and $\boldsymbol{x}_0$ differ considerably, the Kalman filter gain $\boldsymbol{K}$ becomes small and the estimator relies on the model predictions more than it should. Thus, subsequent measurements do not have the impact on the estimator that they need to have. The filter might then learn the wrong state too well and diverge.[14] The importance of choosing a consistent pair of $\hat{\boldsymbol{x}}_0$ and $\boldsymbol{P}_0$ is also emphasized by other authors.[15]

To that end, some authors, e.g. Jazwinski[14] or Valappil and Georgakis,[5] propose to initialize $\boldsymbol{P}_0$ as follows:

$$\boldsymbol{P}_0 = diag\left((\hat{\boldsymbol{x}}_0 - \boldsymbol{x}_0)^T(\hat{\boldsymbol{x}}_0 - \boldsymbol{x}_0)\right). \tag{5}$$

Unfortunately, the true state $\boldsymbol{x}_0$ is rarely known exactly in practice. However, one often knows reasonably accurate upper ($\boldsymbol{x}_u$) and lower ($\boldsymbol{x}_l$) bounds on the initial state which can be used to approximate $\hat{\boldsymbol{x}}_0$, i.e. $\hat{\boldsymbol{x}}_0 = 0.5(\boldsymbol{x}_u + \boldsymbol{x}_l)$, and the $\hat{\boldsymbol{x}}_0 - \boldsymbol{x}_0$ values, i.e. $\hat{\boldsymbol{x}}_0 - \boldsymbol{x}_0 = 0.5(\boldsymbol{x}_u - \boldsymbol{x}_l)$. By multiplying each diagonal element of $\boldsymbol{P}_0$ by a factor $k_{P_0,i}$, one can further adjust one's confidence into a particular initial state guess $x_{i,0}$. As will be demonstrated later, initializing the EKF as recommended by Eq. (5) leads to good estimation results.

Process Noise Covariance

To many authors, the process noise covariance matrix $\boldsymbol{Q}$ is the most difficult design matrix of the EKF. To select a suitable process noise covariance matrix, several approaches have been suggested in the scientific literature, see e.g. Odelson et al.[13] and the references therein. In the statistical derivation of the EKF where the plant model is perfectly known, $\boldsymbol{Q}$ describes the covariance of random process noise. While random process variations are indeed existent, they play a dominant role only in the very rare case that we have a perfectly accurate model.

However, in the majority of real applications the postulated model will contain parametric and structural errors. Both the identification of parameters and the determination of a suitable model

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

structure are tough research problems on their own. Although approaches to find the best model structure can be found in the literature, e.g. by discriminating among a set of candidate models[16] or by incremental refinement,[17] this particular problem is beyond the scope of this paper. For the remainder of this text we will rather assume that the EKF model structure is consistent with the measurement-generating plant.

Even when the model structure is known, the values of its parameters may be unknown *a priori*. To identify these unknown model parameters, experiments have to be performed and data need to be collected. As a result, such a parameter estimation task provides two pieces of information: the best estimate of the fitted, often time-invariant parameters, $\hat{\boldsymbol{p}}$, but also the covariance of such an estimation, $\boldsymbol{C}_{\hat{p}}$.[18] Unfortunately, the latter is frequently ignored when it comes to the EKF design. Thus, the process noise covariance matrix is typically chosen manually through trial-and-error procedures, yielding a constant and diagonal matrix $\boldsymbol{Q}$. For us, however, the covariance matrix $\boldsymbol{C}_{\hat{p}}$ associated with the identification of the parameters is a valuable measure of the random variability naturally present in the process. As such, it should be exploited in the design of the process noise covariance $\boldsymbol{Q}$. To this end, Valappil and Georgakis[5] propose to compute a time-varying $\boldsymbol{Q}$ matrix on-line, based on knowledge about $\boldsymbol{C}_{\hat{p}}$. We briefly recall the proposed design methodology below.

Suppose that the process can be represented by the stochastic differential equation

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{p}), \tag{6}$$

in which $\boldsymbol{p}$ is the stochastic variable, while $\boldsymbol{f}$ and $\boldsymbol{u}$ are deterministic. As a result, $\boldsymbol{x}$ becomes a stochastic variable as well. In the preceding parameter estimation step, the stochastic properties of the parameters $\boldsymbol{p}$ were assumed to satisfy the following two properties:

$$E[\boldsymbol{p}(t)] = \hat{\boldsymbol{p}}, \tag{7}$$

$$E[\boldsymbol{p}(t)\boldsymbol{p}^T(\tau)] = \boldsymbol{C}_{\hat{p}}\delta(t-\tau). \tag{8}$$

The EKF represents the stochastic part as an additive component by approximating Eq. (6) above

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

with the following one:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \hat{\boldsymbol{p}}) + \boldsymbol{w}(t). \tag{9}$$

Here, $\hat{\boldsymbol{p}}$ are the nominal values of the estimated parameters and $\boldsymbol{w}(t)$ is a zero mean white noise stochastic variable with covariance $\boldsymbol{Q}(t)$. To calculate the covariance $\boldsymbol{Q}$ from the parameter covariance $\boldsymbol{C}_{\hat{p}}$ of the estimated parameters we need to set $\boldsymbol{p} = \hat{\boldsymbol{p}} + \delta\boldsymbol{p}$, in Eq. (6) and expand in a Taylor series with respect to the $\boldsymbol{p}$ variables. Here $\delta\boldsymbol{p}$ represent the stochastic part of the estimated parameters, with zero mean and covariance $\boldsymbol{C}_{\hat{p}}$. Retaining only the linear terms in the Taylor series expansion, we have

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \hat{\boldsymbol{p}}) + \boldsymbol{J}_{\hat{p}}(t)\delta\boldsymbol{p} \tag{10}$$

with

$$\boldsymbol{J}_{\hat{p}}(t) = \left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{p}}\right)_{\boldsymbol{x}(t), \boldsymbol{u}(t), \hat{\boldsymbol{p}}}. \tag{11}$$

From the above equation one can calculate $\boldsymbol{Q}$ as follows:

$$\boldsymbol{Q}(t) = k_Q \boldsymbol{J}_{\hat{p}}(t) \boldsymbol{C}_{\hat{p}} \boldsymbol{J}_{\hat{p}}^T(t). \tag{12}$$

Here, the factor k_Q equals 1 unless one feels that the plant-model mismatch is larger than calculated above, in which case k_Q can be given a value larger than 1. One can observe that even if $\boldsymbol{C}_{\hat{p}}$ is time-invariant, as is the case in this paper, $\boldsymbol{Q}(t)$ is time-varying because $\boldsymbol{J}_{\hat{p}}(t)$ is time-varying. Furthermore, $\boldsymbol{J}_{\hat{p}}(t)$ depends on the estimated process trajectory and so will the values of $\boldsymbol{Q}(t)$.

In the following sections, we will study the influence of the aforementioned design methodology on the performance of the EKF.

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

Benchmark Problem: Batch Reactor

Problem Configuration

As a benchmark problem, we consider the popular batch reactor example which was originally published by Haseltine and Rawlings[6] and later used in a number of other publications for estimator performance comparison.[10,15,19] The model describes the following reversible gas-phase reactions, that take place in an isothermal, well-mixed batch reactor with constant volume:

$$\mathrm{A} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{B} + \mathrm{C}, \tag{13}$$

$$2\,\mathrm{B} \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \mathrm{C}. \tag{14}$$

The true reaction rate constants are given as $\boldsymbol{k} = [k_1,\ k_2,\ k_3,\ k_4]^T = [0.5,\ 0.05,\ 0.2,\ 0.01]^T$.

By using a simple power model for the kinetics, the following nonlinear, continuous-time ordinary differential equation model can be derived for the states $\boldsymbol{x} = [c_\mathrm{A},\ c_\mathrm{B},\ c_\mathrm{C}]^T$ from first principles:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{p}) = \begin{bmatrix} -k_1 c_\mathrm{A} + k_2 c_\mathrm{B} c_\mathrm{C} \\ k_1 c_\mathrm{A} - k_2 c_\mathrm{B} c_\mathrm{C} - 2k_3 c_\mathrm{B}^2 + 2k_4 c_\mathrm{C} \\ k_1 c_\mathrm{A} - k_2 c_\mathrm{B} c_\mathrm{C} + k_3 c_\mathrm{B}^2 - k_4 c_\mathrm{C} \end{bmatrix}, \tag{15}$$

where c_A, c_B, c_C are the concentrations of species A, B and C, respectively. Assuming that the ideal gas law holds, the sum of the species' partial pressures gives the total reactor pressure, the only available measurement y of the process. It is measured at discrete time instants t_k and a sampling rate of $\Delta t = t_{k+1} - t_k = 0.25$:

$$y(t_k) = h(\boldsymbol{x}(t_k)) + v(t_k) = \begin{bmatrix} \tilde{R}T & \tilde{R}T & \tilde{R}T \end{bmatrix} \boldsymbol{x}(t_k) + v(t_k). \tag{16}$$

$\tilde{R}$, T are the ideal gas constant and the reactor temperature, respectively. In agreement with the original paper,[6] $\tilde{R}T$ is set to 32.84. Furthermore, the measurement signal y is subject to $v(t_k)$,

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

an additional zero mean Gaussian measurement noise with known time-invariant variance $R = 0.0625 (= 0.25^2)$.

In the original paper,[6] a *discrete-time* zero-mean Gaussian process noise $\tilde{\boldsymbol{w}}(t_k)$ with covariance $\tilde{\boldsymbol{Q}}(t_k) = \mathrm{diag}(0.001^2,\ 0.001^2,\ 0.001^2)$ was added to the discretized version of the process model Eq. (15) for the simulation of the plant measurements. With regard to Eq. (3), this implies a continuous-time process noise with a covariance of $\boldsymbol{Q}(t) \approx \tilde{\boldsymbol{Q}}(t_k)/\Delta t$ for small sampling times Δt.[11] As this process noise level is very small, it is neglected here when simulating plant measurements. Instead, only the deterministic model stated in Eq. (15) plus some additional measurement noise as in Eq. (16) will be used.

Design and Influence of Initial State Covariance

Haseltine and Rawlings simulated the plant from the initial conditions $\boldsymbol{x}_0 = [0.5,\ 0.05,\ 0]^T$ while they initialized the EKF with $\hat{\boldsymbol{x}}_0 = [0,\ 0,\ 4]^T$.[6] For the particular reaction system given by Eq. (13) and Eq. (14), this provides the EKF with misleading information, because it makes the EKF assume that the reactor was initially filled with products while in fact it is entirely filled with reactants. To make things worse, a too small initial state covariance estimate, $\boldsymbol{P}_0 = \mathrm{diag}(0.5^2,\ 0.5^2,\ 0.5^2)$, is provided to the EKF. Such a small covariance triggers a high confidence into this misleading initial state estimate. Under these conditions, it is not surprising that the performance of the unconstrained filter is poor and that the EKF does not converge to the true states in an acceptable period of time.

Thus, the first reason why the comparison in the original paper was unfair is that it compared an unconstrained with a constrained estimator. This could be remedied by imposing constraints also on the EKF in which case unphysical state estimates and filter failure can be avoided even for misleading prior information.[20] In this paper, however, we argue that an even simpler way than imposing constraints is a meaningful filter initialization.

Beyond utilizing constraints, the MHE had a second unfair advantage upon the EKF in the original benchmark problem: The MHE was allowed to consider not only one measurement at a time but at least eleven. In contrast to the filtered estimates of the EKF, this allows the MHE

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

to obtain smoothed estimates on each horizon.[21] Allowing the MHE to process more than one measurement at a time is the second reason why the original comparison was unfair.

It should be noted, however, that even if the horizon length of the MHE was set to one, the MHE can be expected to outperform the EKF. This is because of the local optimization performed by the constrained MHE. This means that at each time step, the MHE optimization problem is typically iterated until convergence whereas the standard EKF stops after the first iteration. For detailed discussions of these aspects, the reader is referred to Robertson et al.[22] or the dissertation of Rao.[23]

In this remainder of this section, we will show that the EKF can indeed converge, even in the presence of poor initial guesses – *as long as the initial covariance* $\boldsymbol{P}_0$ *is specified accordingly*. This was also recognized by a number of other authors who used that benchmark example. For instance, Vachhani et al.[15] emphasize the role of a consistent choice of $\boldsymbol{P}_0$ and consequently use alternate initial state information for their simulations. Similarly, Prakash et al.[19] silently adapt their initial state estimate $\hat{\boldsymbol{x}}_0$ to be more consistent with the original initial covariance $\boldsymbol{P}_0$.

For the batch reactor example, we study the effect of different choices of diagonal matrices $\boldsymbol{P}_0$ on the estimation performance on the time interval $t = [0,\ 30]$ with $N = 121$ output samples. For each choice of $\boldsymbol{P}_0$ we run the EKF $K = 1{,}000$ times on noisy process measurements. To simulate these noisy measurements, the clean output samples of the true process are corrupted by a series of random measurement noise whose realization is different for each EKF run. For performance comparison, we compute the so-called mean squared error (MSE), adapted from Ungarala:[24]

$$MSE = \frac{1}{MKNn_x} \sum_{m=1}^{M} \sum_{j=1}^{K} \sum_{k=0}^{N-1} \left(\hat{\boldsymbol{x}}^j(t_k^+) - \boldsymbol{x}^j(t_k)\right)^T \left(\hat{\boldsymbol{x}}^j(t_k^+) - \boldsymbol{x}^j(t_k)\right), \tag{17}$$

where the superscript j denotes the values of the j-th EKF run, estimated from the measurements that contain the j-th series of random measurement noise realizations. Let further $M = 1$ for the time being.

We now compare two different choices of $\boldsymbol{P}_0$, namely the one used by Haseltine and Rawlings,[6]

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

i.e. $\boldsymbol{P}_0 = \mathrm{diag}(0.5^2,\ 0.5^2,\ 0.5^2)$, and the one according to Eq. (5), i.e. $\boldsymbol{P}_0 = \mathrm{diag}(0.5^2,\ 0.05^2,\ 4^2)$. The other filter covariance matrices are kept constant, i.e. $R = 0.25^2$, $\boldsymbol{Q} = \mathrm{diag}(0.001^2,\ 0.001^2,\ 0.001^2)/0.25$, and the true kinetic constants are used in the EKF model. Three typical EKF runs are visualized in Figure 1, Figure 2 and Figure 3. It can be seen that with the design of Haseltine and Rawlings some simulations do not converge to the true states in the simulated time frame. In order to quantify the percentage of converging simulations, we consider an EKF run as converged if the absolute estimation error of *all* concentrations $|\hat{c}_i - c_i|$ at time $t = 30$ is less than 0.02. This convergence fraction is presented in Table 1 along with average, minimum and maximum values and the standard deviation of the MSEs. Note that in this case there is a natural lower bound on the MSE for any distinct pair of true initial state $\boldsymbol{x}_0$ and initial state estimate $\hat{\boldsymbol{x}}_0$. This bound can be computed by assuming zero estimation error for $k > 0$ in Eq. (17). For the particular choice of $\boldsymbol{x}_0$ and $\hat{\boldsymbol{x}}_0$ stated above, the smallest possible MSE value equates to 0.0448. In this sense, the EKF with properly designed $\boldsymbol{P}_0$ matrix performs very well. More details on the MSE distribution can be found in the relative frequency plot in Figure 4.

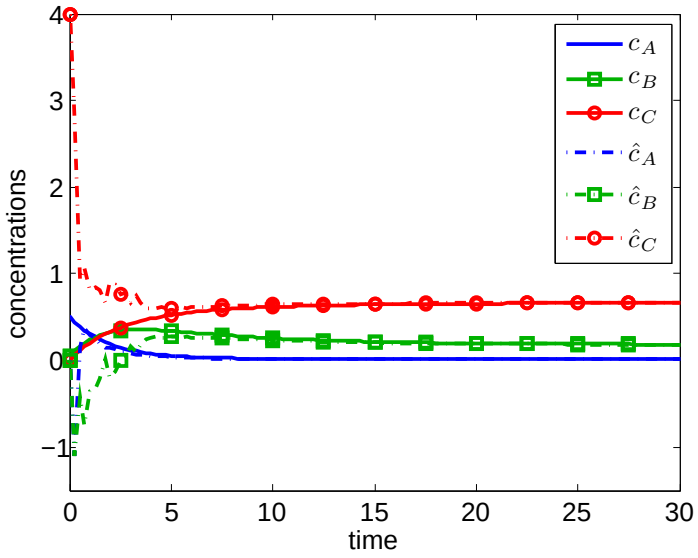

Figure 1: Example of an EKF run that *converges* with the $\boldsymbol{P}_0$ matrix used by Haseltine and Rawlings.[6]

As violation of physical constraints, such as as non-negativity of concentrations, is a major concern in the literature, we additionally introduce a measure for constraint violation (MCV). We

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

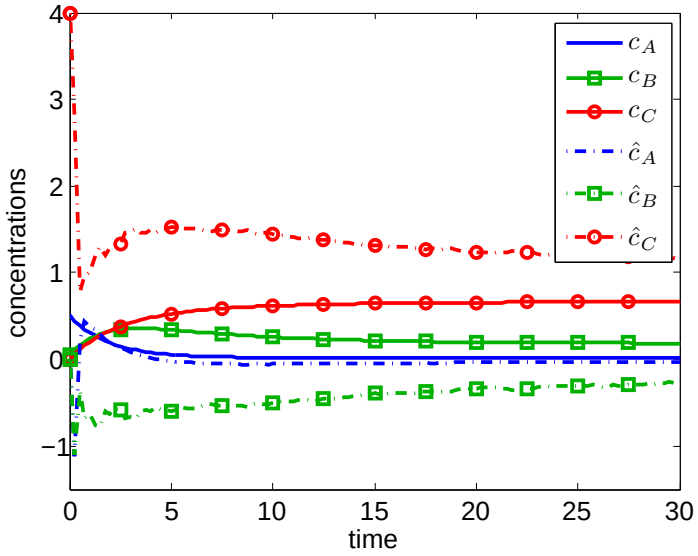

Figure 2: Characteristic example of an EKF run that *diverges* with the $\boldsymbol{P}_0$ matrix used by Haseltine and Rawlings.[6]

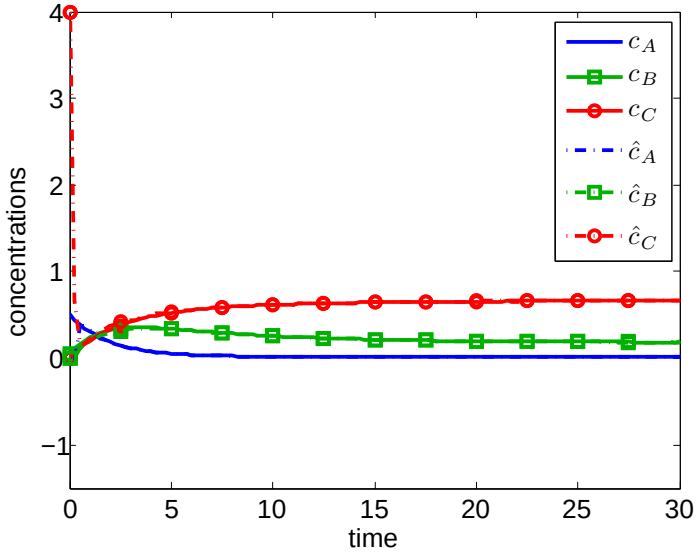

Figure 3: Characteristic example of an EKF simulation run that converges with the $\boldsymbol{P}_0$ matrix chosen according to Eq. (5), as proposed by Valappil and Georgakis.[5]

Table 1: Number of converging runs and estimation error statistics. For each choice of the diagonal $\boldsymbol{P}_0$ matrix, 1,000 EKF runs with different measurement noise realizations were performed.

diag($\boldsymbol{P}_0$)			conv. runs	avg MSE	min MSE	max MSE	std MSE
0.5^2	0.5^2	0.5^2	205	0.3331	0.0803	0.4331	0.1218
0.5^2	0.05^2	4^2	1,000	0.0468	0.0465	0.0478	0.0002

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

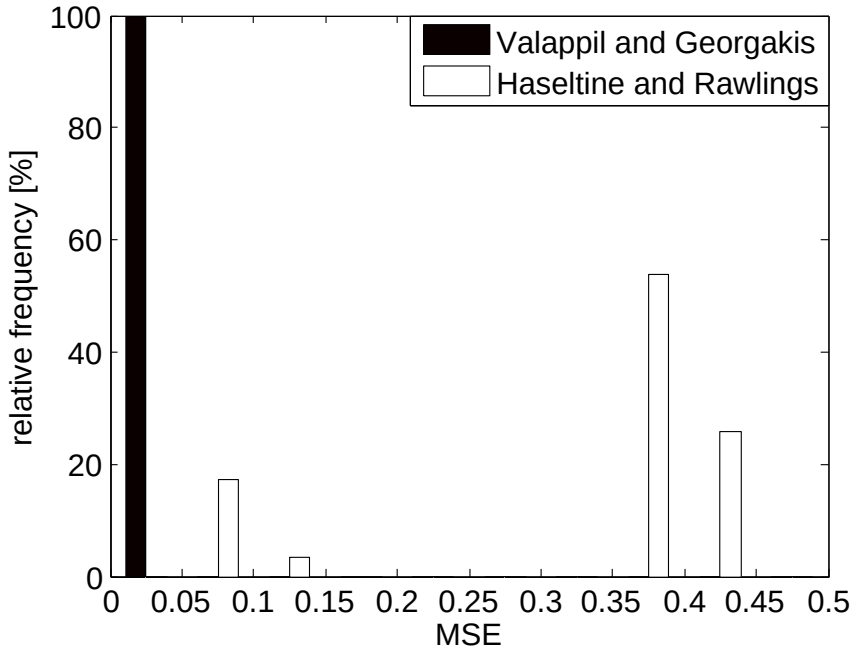

Figure 4: Detailed MSE frequency distribution corresponding to the different $\boldsymbol{P}_0$ choices in Table 1. The first row is shown in white and the second row in black.

Table 2: Measure for constraint violation for different choices of diagonal initial state covariances $\boldsymbol{P}_0$.

diag($\boldsymbol{P}_0$)			avg MCV	min MCV	max MCV	std MCV
0.5^2	0.5^2	0.5^2	97.2520	5	120	44.8276
0.5^2	0.05^2	4^2	1.0260	1	2	0.1592

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

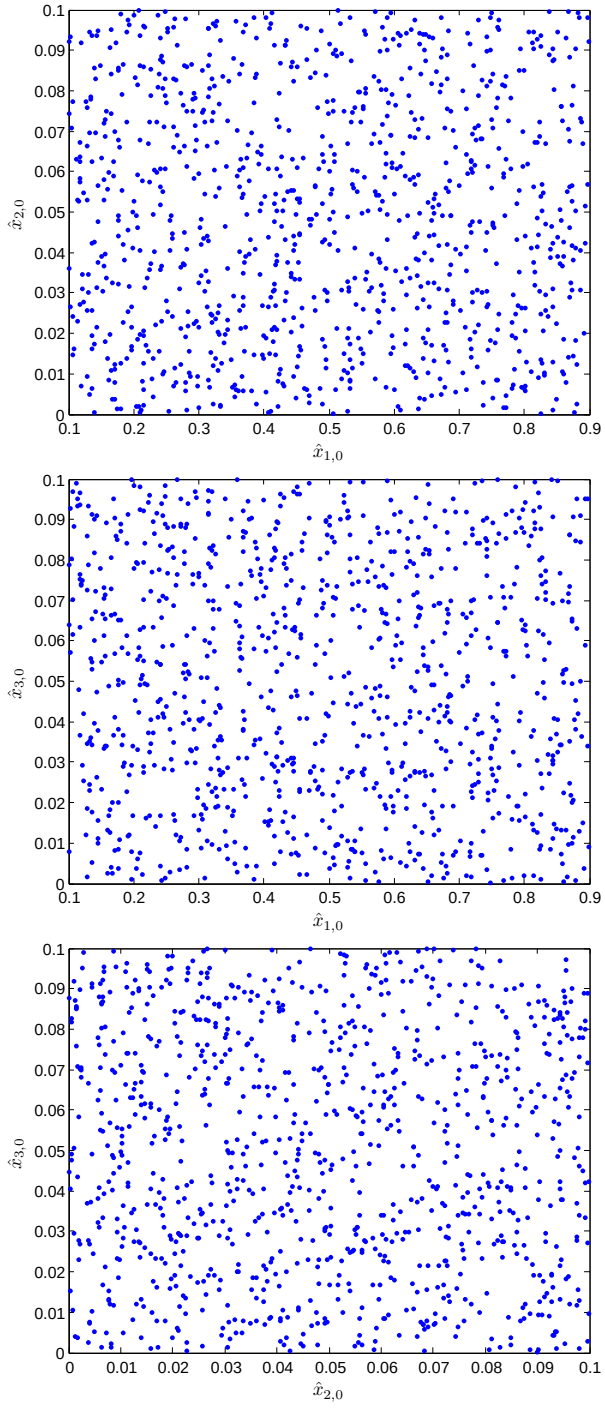

Figure 5: Distribution of randomly generated initial state estimates $\hat{\boldsymbol{x}}_0$.

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

define the MCV to be an integer that counts the number of samples in each run in which at least one of the three estimated concentrations is negative. The average, minimum and maximum values as well as the standard deviation of that measure can be found in Table 2.

These results clearly show that the initialization of $\boldsymbol{P}_0$ proposed in Eq. (5), leads to a considerably improved performance compared to the results of Haseltine and Rawlings[6] and does not necessarily violate constraints, *even though the EKF does not know about them*.

While the large number of simulation runs should already have eliminated most random measurement noise effects, one might still argue that our recommended method of initialization only works well for the particular initial state estimate considered so far. In order to dispel these concerns, we run another series of $M = 1{,}000$ simulations. While the true initial state $\boldsymbol{x}_0$ remains unchanged at $[0.5,\ 0.05,\ 0]^T$, the EKF's initial state estimate $\hat{\boldsymbol{x}}_0$ is now generated randomly in each run. The underlying probability distribution is uniform and defined on the following set:

$$\hat{\boldsymbol{x}}_0 \in [0.1,\ 0.9] \times [0,\ 0.1] \times [0,\ 0.1]. \tag{18}$$

We choose this set to represent a variety of reasonable educated guesses. To that end, the set should contain the true initial state and, if allowed by physical constraints, a desirably symmetric neighborhood of it. The particular realization of these $M = 1{,}000$ different random initial states in the $\mathbb{R}^3$ state space is shown in Figure 5. Subsequently, the corresponding initial state covariance $\boldsymbol{P}_0$ is constructed for each initial state according to Eq. (5). Each EKF initialization data pair, consisting of a particular random initial state estimate $\hat{\boldsymbol{x}}_0$ and corresponding initial state covariance $\boldsymbol{P}_0$, is then tested on $K = 100$ random sequences of measurement noise, added to the $N = 121$ clean output samples of the plant. This leads to a total of $100{,}000$ EKF simulations, ***all*** of which converged in the sense of the absolute estimation error threshold presented above.

The results are tabulated in Table 3 and detailed in Figure 6. The relation between the estimation performance using the suggested design procedure and the initial state estimation error $|\hat{\boldsymbol{x}}_0 - \boldsymbol{x}_0|$ is visualized in Figure 7. As expected, the average estimation error increases with in-

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

creasing error in the initial state. Clearly, the proposed method for choosing a suitable initial state covariance, Eq. (5), leads generally to favorable performance, not just for the particular initial state estimate used in Haseltine and Rawlings.[6]

Table 3: MSE statistics for random initial state estimates $\hat{\boldsymbol{x}}_0$.

Average mean	Global minimum	Global maximum	Average standard deviation
2.1483×10^{-4}	1.9930×10^{-6}	9.9957×10^{-4}	1.8339×10^{-5}

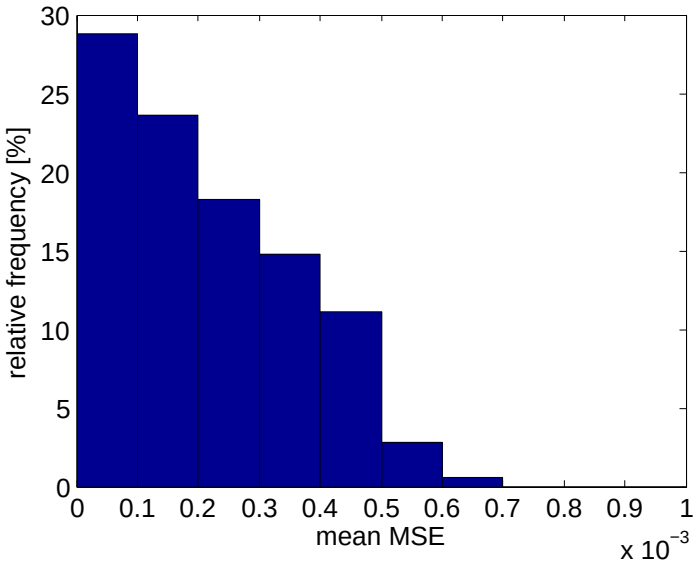

Figure 6: Relative frequency of mean MSE values for all random choices of $\hat{\boldsymbol{x}}_0$.

Design and Influence of Process Noise Covariance

In the simulations presented so far, no model-plant mismatch was present. Such a mismatch, however, is frequently encountered in practice and well known to cause biased or diverged EKF estimates.[14] In the following, we present the typical workflow of designing an EKF in the presence of *parametric* model-plant mismatch caused by unknown plant parameters. By following our proposed design methodology, good results are obtained.

In the very beginning, plant model parameters are usually identified from open-loop operation data. The identified parameters will differ from the true plant parameters. Since only the identified

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

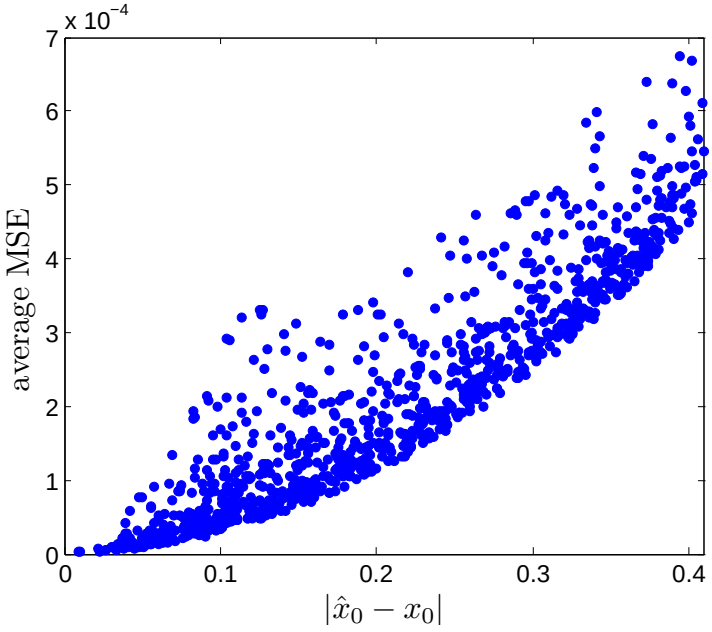

Figure 7: Average MSE as a function of the absolute initial estimation error $|\hat{\boldsymbol{x}}_0 - \boldsymbol{x}_0|$.

parameters are available for the design of the EKF, parametric model-plant mismatch will influence the EKF estimation quality considerably.

In detail, the procedure for the parameter identification step is as follows: Similar to the simulations performed in the previous section, a particular realization of zero-mean Gaussian measurement noise with measurement variance $R = 0.25^2$ is initially added to the true system outputs. From these simulated noisy process measurements the four kinetic parameters k_1, k_2, k_3 and k_4 and their covariance matrix $\boldsymbol{C}_{\hat{p}}$, are estimated using PSE's gPROMS software. To that end, gPROMS is provided the true initial state of the process, the structurally correct plant model and the true value of the measurement noise covariance R. We then solve the maximum likelihood parameter estimation problem with the following objective function Φ,[25] adapted to our particular reactor example:

$$\min_{\hat{\boldsymbol{k}}} \Phi = \frac{N}{2}\ln(2\pi) + \frac{1}{2}\sum_{k=0}^{N-1}\left(\ln(R) + \frac{(p_{meas}(t_k) - p(\hat{\boldsymbol{k}}, t_k))^2}{R}\right),$$

where N is the number of noisy measurement samples, $p_{meas}(t_k)$ is the noisy measured total reactor pressure at time t_k and $p(\hat{\boldsymbol{k}}, t_k)$ is the predicted reactor pressure at time t_k using the model Eq. (15) and Eq. (16) with the parameter estimate $\hat{\boldsymbol{p}}$. We also compute an approximation to the covariance

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

of the identified kinetic constants as follows:[25]

$$\boldsymbol{C}_{\hat{p}} = \left(\frac{\partial^2}{\partial \hat{\boldsymbol{k}}}\Phi\right)^{-1} \left(\frac{\partial^2}{\partial p_{meas}\partial \hat{\boldsymbol{k}}}\Phi\right) R \left(\frac{\partial^2}{\partial p_{meas}\partial \hat{\boldsymbol{k}}}\Phi\right)^{T} \left(\frac{\partial^2}{\partial \hat{\boldsymbol{k}}}\Phi\right)^{-1}.$$

The estimated parameters are $\hat{\boldsymbol{k}} = [0.4938800,\ 0.0313430,\ 0.2122300,\ 0.0099926]^T$ whereas the parameter covariance matrix was estimated as

$$\boldsymbol{C}_{\hat{p}} = \begin{bmatrix} 3.70\times10^{-6} & 9.50\times10^{-6} & -5.83\times10^{-6} & 2.36\times10^{-8} \\ 9.50\times10^{-6} & 3.37\times10^{-4} & -2.55\times10^{-4} & -2.68\times10^{-6} \\ -5.83\times10^{-6} & -2.55\times10^{-4} & 1.97\times10^{-4} & 2.31\times10^{-6} \\ 2.36\times10^{-8} & -2.68\times10^{-6} & 2.31\times10^{-6} & 4.79\times10^{-8} \end{bmatrix}.$$

With the exception of k_2, all identified parameters are very close to the true plant parameters. Since the true plant parameters are typically not known in practice, the identified parameters are validated in a different manner: We perform an open-loop simulation of the system using the identified parameters and compare the simulated, noise-free process measurements with those used to identify the parameters in the first place. In our case, the parameters identified by gPROMS lead to the open-loop trajectories presented in Figure 8. Note that the true internal states are not available in practice but shown here only for comparison. However, the plant's original noisy measurements are available and are compared to the open-loop predictions in Figure 9. Clearly, the prediction based on the estimated parameters matches the plant measurements very well.

The next step is to use the identified parameters and their covariance to design an appropriate EKF. Starting from the true initial state, i.e. $\hat{\boldsymbol{x}}_0 = \boldsymbol{x}_0$, and with a small initial state covariance matrix $\boldsymbol{P}_0 = \mathrm{diag}(0.001^2,\ 0.001^2,\ 0.001^2)$ to express a strong confidence into this choice of $\hat{\boldsymbol{x}}_0$, $K = 1{,}000$ EKF simulation runs are performed ($M = 1$). During each EKF run, the time-varying process noise covariance matrix $\boldsymbol{Q}(t)$ is computed according to Eq. (12). The resulting MSE values are tabulated in Table 4 and their statistics visualized in Figure 10. While the small MSE values indicate good estimation performance, they cannot be compared to those computed in the previous

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

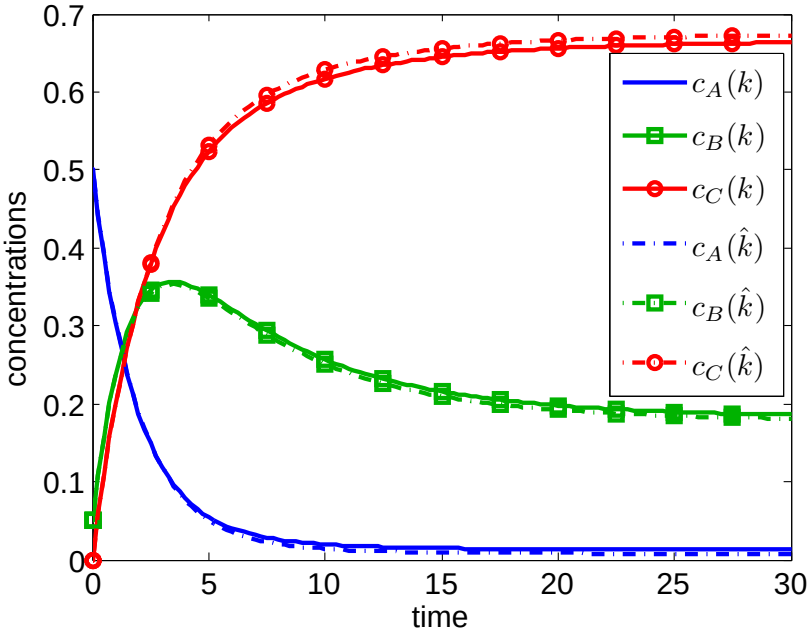

Figure 8: Open loop simulation using the estimated parameters and the true initial state.

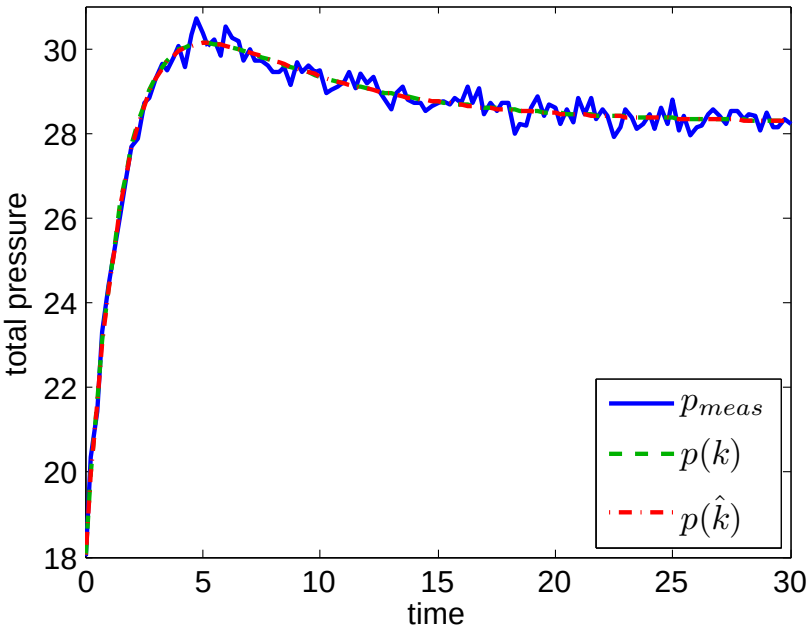

Figure 9: Open loop output prediction corresponding to Figure 8.

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

section. The reason is that in the previous section, the large error of the initial state estimate contributed significantly to the MSE. Exemplary closed-loop EKF trajectories are shown in Figure 11. As expected, only a small steady state estimation error is present due to model-plant parameter mismatch.

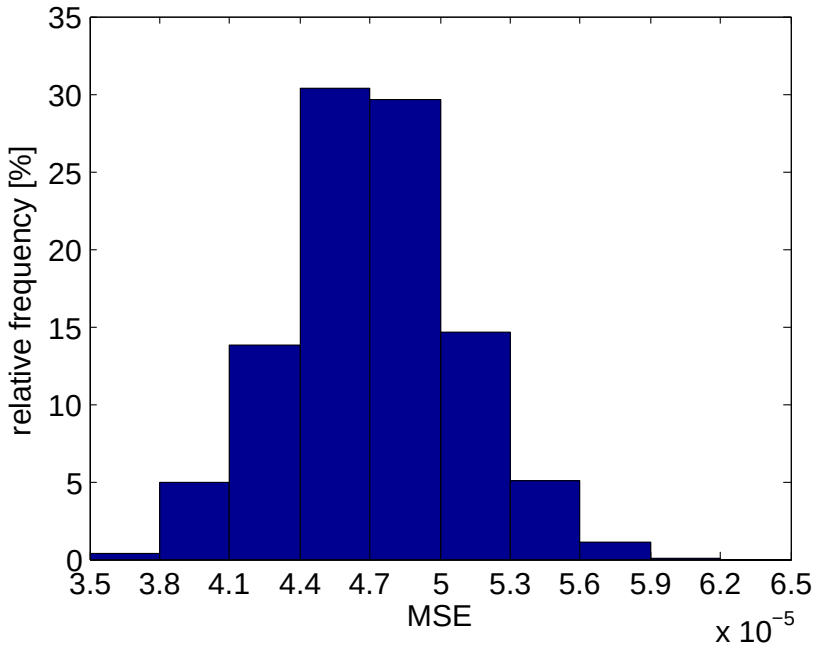

Figure 10: Detailed MSE distribution for EKF runs using the identified model parameters and the corresponding time-varying process noise covariance matrix $\boldsymbol{Q}(t)$.

These simulations emphasize that the proposed method to compute $\boldsymbol{Q}(t)$ on-line using information from a preceding parameter identification step leads to good closed-loop results. These results, obtained by using a time-varying process noise covariance matrix, are now compared with traditional approaches where $\boldsymbol{Q}$ is time-invariant. To that end, we first compute the diagonal elements of two suitable diagonal and time-invariant process noise covariance matrices $\boldsymbol{Q}_{mean}$ and $\boldsymbol{Q}_{max}$ as follows:

$$\begin{aligned} Q_{mean,ii} &= \frac{1}{KN}\sum_{j=1}^{K}\sum_{k=0}^{N-1} Q_{ii}^{j}(t_k), \\ Q_{max,ii} &= \max_{j\in\{1,\ldots,K\},k\in\{0,\ldots,N-1\}} Q_{ii}^{j}(t_k), \end{aligned}$$

where $i \in \{1,\ 2,\ 3\}$ and $\boldsymbol{Q}(t_k)$ are the corresponding time-varying process noise covariance ma-

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

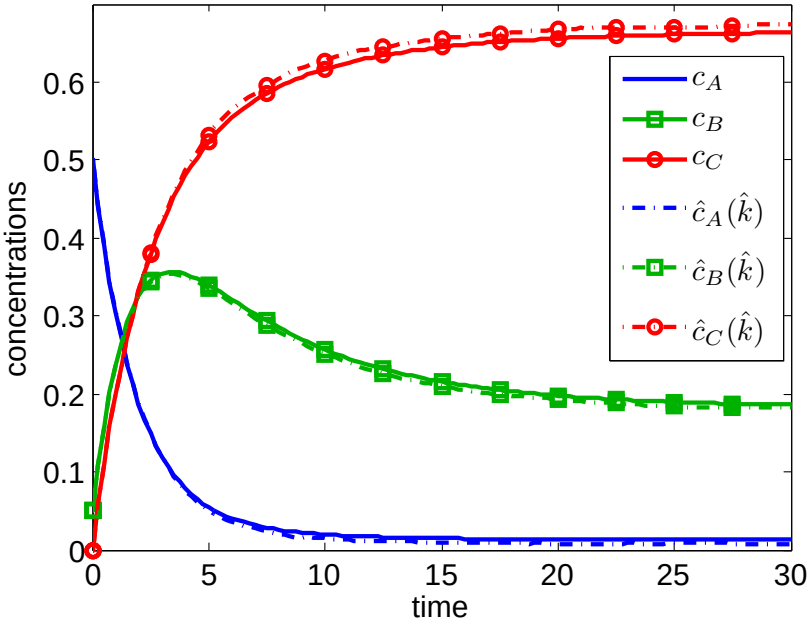

Figure 11: Typical closed loop EKF run using the identified model parameters and the corresponding time-varying process noise covariance matrix $\boldsymbol{Q}(t)$.

trices at each sampling instant t_k from the preceding EKF runs. The resulting diagonal matrices are:

$$\begin{aligned} \boldsymbol{Q}_{mean} &= \text{diag}(6.41 \times 10^{-6},\ 1.77 \times 10^{-6},\ 1.07 \times 10^{-5}), \\ \boldsymbol{Q}_{max} &= \text{diag}(1.09 \times 10^{-5},\ 2.37 \times 10^{-6},\ 2.30 \times 10^{-5}). \end{aligned}$$

Similar to the previous simulations with the time-varying matrix $\boldsymbol{Q}(t)$, another two sets of $K = 1,000$ EKF simulation runs with random measurement noise were performed, now using the time-invariant diagonal matrices $\boldsymbol{Q}_{mean}$ and $\boldsymbol{Q}_{max}$. Again, the true initial state was assumed to be known, i.e. $\hat{\boldsymbol{x}}_0 = \boldsymbol{x}_0$. Table 4 shows the average, minimum, maximum and the standard deviation of the resulting MSE values. In all cases, the performance is acceptable and the absolute estimation error of all concentrations at time $t = 30$ is less than 0.02. Overall, all choices of $\boldsymbol{Q}$ lead to satisfactory EKF performance that is similar to the one obtained with the time-varying $\boldsymbol{Q}(t)$ matrix. One may be tempted to conclude that due to implementational simplicity, the constant and diagonal process covariance matrix is the preferred choice. However, note that such a well-performing matrix is

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

typically not directly availabe in practice. Instead, computing such a suitable constant matrix may require extensive simulations with the time-varying matrix $\boldsymbol{Q}(t)$. Furthermore, $\boldsymbol{Q}_{mean}$ and $\boldsymbol{Q}_{max}$ need to be recalculated for different starting points $\hat{\boldsymbol{x}}_0$. These reasons suggest to rather use the time-varying matrix directly in order to obtain good estimation performance.

Table 4: Error statistics for EKF runs with different choices of the process noise covariance matrix $\boldsymbol{Q}$.

	avg MSE	min MSE	max MSE	std MSE
$\boldsymbol{Q}(t)$	4.71×10^{-5}	3.76×10^{-5}	6.04×10^{-5}	3.70×10^{-6}
$\boldsymbol{Q}_{mean}$	4.86×10^{-5}	3.96×10^{-5}	6.11×10^{-5}	3.45×10^{-6}
$\boldsymbol{Q}_{max}$	4.93×10^{-5}	3.95×10^{-5}	6.06×10^{-5}	3.44×10^{-6}

Comparison with the Unscented Kalman Filter

In the present section we shall relate these results obtained with the EKF to those obtained with more modern estimation methods, such as the UKF. To that end, we briefly present some simulation results for the same reactor case study, based on different UKF algorithms that were proposed in Kolås et al. [10] and the references therein.

The following results are based on $K = 100$ random measurement noise sequence realizations and the same configuration as described earlier. Again, the true initial state is $\boldsymbol{x}_0 = [0.5,\ 0.05,\ 0]^T$ while the UKF starts at $\hat{\boldsymbol{x}}_0 = [0,\ 0,\ 4]^T$ with the misleading initial state covariance matrix $\boldsymbol{P}_0 = \mathrm{diag}(0.5^2,\ 0.5^2,\ 0.5^2)$. The UKF is further supplied the same structurally and parametrically correct process model as the true process.

The average MSE values for several out-of-the-box UKF algorithms are shown in Table 5 in which the notation for the UKF algorithms, variables and constraints is identical to the one used by Kolås et al. [10]

By comparison of these results with Table 1, we notice that the unconstrained UKF performs only slightly better than the EKF when $\boldsymbol{P}_0$ is chosen too optimistic. This was also noted by Daum [26] who comments that the unconstrained UKF "often provides a significant improvement relative to

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

Table 5: Performance of several UKF algorithms on the reactor benchmark example. See Kolås et al.[10] for details on the different algorithms.

UKF algorithm	α	β	κ	Γ	constr.	avg MSE
Standard, Cholesky decomp.	1	2	0	$\sqrt{3}$	-	0.2924
Standard, symmetric square root	1	2	0	$\sqrt{3}$	-	0.3044
Augmented, Cholesky decomp.	1	2	0	$\sqrt{7}$	-	0.2091
Augmented, symmetric square root	1	2	0	$\sqrt{7}$	-	0.2188
Augmented, Cholesky decomp.	1	2	0	$\sqrt{7}$	1	0.0820
QP, symmetric square root	1	10	0	$\sqrt{3}$	1, 3, 7	0.0486

the EKF but sometimes it does not".[26] However, it should be noted that both the unconstrained UKF and EKF diverge in most runs. For the UKF, this corresponds to the observations of Kolås et al.[10]

On the other hand, the constrained UKF performs notably better, even with poor prior information. However, the EKF performs comparably well when its $\boldsymbol{P}_0$ matrix is properly designed. Thus, considering the efforts required to implement and design a constrained UKF, one might as well specify more reasonable initial state information for use by an existing EKF implementation.

Conclusion

In this work we have shown that systematic approaches to the EKF design can be powerful. When its design matrices are chosen in an appropriate way, the EKF delivers good performance that can keep up with more recent methods, such as the MHE and / or the UKF. This is even true in the case of a poor initial state guess, which should, however, not be confused with an intentionally misleading initialization. We emphasized that it is important to choose a consistent initial state covariance matrix in order to realistically reflect the confidence into the initial state estimate. We further proposed an effective design methodology for the process noise covariance matrix that exploits information about parametric uncertainties in the process model when these are known. Considering the implementational simplicity, we conclude that the EKF remains an attractive state estimation method.

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

Acknowledgement

Financial support by Cybernetica AS through the project "Model identification methods for online monitoring, control and economic optimization" is gratefully acknowledged. The authors also thank Prof. Wolfgang Marquardt for helpful discussions. This project was initiated while Professor Georgakis was at RWTH Aachen University on sabbatical leave from Tufts University. He is grateful to both institutions for making this possible.

References

1. Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. *Trans. ASME, Ser. D* **1960**, *82*, 35–45.

2. Lewis, F. *Optimal Estimation: With an Introduction to Stochastic Control Theory*; John Wiley and Sons, 1986.

3. Mesbah, A.; Huesman, A.; Kramer, H.; van den Hof, P. A Comparison of Nonlinear Observers for Output Feedback Model-Based Control of Seeded Batch Crystallization Processes. *J. Process Control* **2011**, *21*, 652–666.

4. Rao, C. V.; Rawlings, J. B. Constrained Process Monitoring: Moving-Horizon Approach. *AIChE J.* **2002**, *48*, 97–109.

5. Valappil, J.; Georgakis, C. Systematic Estimation of State Noise Statistics for Extended Kalman Filters. *AIChE J.* **2000**, *46*, 292–308.

6. Haseltine, E. L.; Rawlings, J. B. A Critical Evaluation of Extended Kalman Filtering and Moving-Horizon Estimation. *Ind. Eng. Chem. Res.* **2005**, *44*, 2451–2460.

7. Darby, M. L.; Nikolaou, M. A Parametric Programming Approach to Moving-Horizon State Estimation. *Automatica* **2007**, *43*, 885 – 891.

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

8. Zavala, V. M.; Laird, C. D.; Biegler, L. T. A Fast Computational Framework for Large-Scale Moving Horizon Estimation. *Proceedings of the 8th International IFAC Symposium on Dynamics and Control of Process Systems* **2007**, *3*, 21–30.

9. Zavala, V. M.; Laird, C. D.; Biegler, L. T. A Fast Moving Horizon Estimation Algorithm based on Nonlinear Programming Sensitivity. *J. Process Control* **2008**, *18*, 876 – 884.

10. Kolås, S.; Foss, B.; Schei, T. Constrained Nonlinear State Estimation based on the UKF Approach. *Comput. Chem. Eng.* **2009**, *33*, 1386 – 1401.

11. Simon, D. *Optimal State Estimation: Kalman, H [infinity] and Nonlinear Approaches*; John Wiley and Sons, 2006.

12. Garriga, J. L.; Soroush, M. Model Predictive Control Tuning Methods: A Review. *Ind. Eng. Chem. Res.* **2010**, *49*, 3505–3515.

13. Odelson, B. J.; Rajamani, M. R.; Rawlings, J. B. A New Autocovariance Least-Squares Method for Estimating Noise Covariances. *Automatica* **2006**, *42*, 303 – 308.

14. Jazwinski, A. H. *Stochastic Processes and Filtering Theory*; Academic Press, New York, 1970.

15. Vachhani, P.; Narasimhan, S.; Rengaswamy, R. Recursive State Estimation in Nonlinear Processes. *Proceedings of the 2004 American Control Conference* **2004**, *1*, 200–204.

16. Michalik, C.; Stuckert, M.; Marquardt, W. Optimal Experimental Design for Discriminating Numerous Model Candidates - The AWDC Criterion. *Ind. Eng. Chem. Res.* **2009**, *49*, 913–919.

17. Bhatt, N.; Kerimoglu, N.; Amrhein, M.; Marquardt, W.; Bonvin, D. Incremental Model Identification for Reaction Systems - A Comparison of Rate-based and Extent-based Approaches. *Chem. Eng. Sci.* **2012**, *dx.doi.org/10.1016/j.ces.2012.05.040*, In Press.

18. Aster, R.; Borchers, B.; Thurber, C. *Parameter Estimation and Inverse Problems*; Academic Press, 2004.

This document is the Accepted Manuscript version of a Published Work that appeared in final form in Industrial & Engineering Chemistry Research, copyright © American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see pubs.acs.org/doi/abs/10.1021/ie300415d

19. Prakash, J.; Patwardhan, S. C.; Shah, S. L. Constrained Nonlinear State Estimation Using Ensemble Kalman Filters. *Ind. Eng. Chem. Res.* **2010**, *49*, 2242–2253.

20. Ungarala, S.; Dolence, E.; Li, K. Constrained Extended Kalman Filter for Nonlinear State Estimation. *Proceedings of the 8th International IFAC Symposium on Dynamics and Control of Process Systems* **2007**, *2*, 63–68.

21. Gelb, A. *Applied Optimal Estimation*; MIT Press, 1974.

22. Robertson, D. G.; Lee, J. H.; Rawlings, J. B. A Moving Horizon-based Approach for Least-squares Estimation. *AIChE J.* **1996**, *42*, 2209–2224.

23. Rao, C. V. Moving Horizon Strategies for the Constrained Monitoring and Control of Nonlinear Discrete-Time Systems. Ph.D. thesis, University of Wisconsin-Madison, 2000.

24. Ungarala, S. Computing Arrival Cost Parameters in Moving Horizon Estimation Using Sampling Based Filters. *J. Process Control* **2009**, *19*, 1576 – 1588.

25. gPROMS Advanced User Guide. Process Systems Enterprise Ltd.: Bridge Studios, 107a Hammersmith Bridge Road, London W6 9DA, United Kingdom, 2004.

26. Daum, F. Nonlinear Filters: Beyond the Kalman Filter. *IEEE Aerospace and Electronic Systems Magazine* **2005**, *20*, 57 –69.

View publication stats