

University of Groningen

## Moment estimation for chemically reacting systems by extended Kalman filtering

Ruess, J; Miliias-Argeitis, A; Summers, S; Lygeros, J

*Published in:*  
Journal of Chemical Physics

*DOI:*  
[10.1063/1.3654135](https://doi.org/10.1063/1.3654135)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2011

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Ruess, J., Miliias-Argeitis, A., Summers, S., & Lygeros, J. (2011). Moment estimation for chemically reacting systems by extended Kalman filtering. *Journal of Chemical Physics*, 135(16), [165102]. <https://doi.org/10.1063/1.3654135>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Moment estimation for chemically reacting systems by extended Kalman filtering

J. Ruess, A. Milias-Argeitis, S. Summers, and J. Lygeros

Citation: *The Journal of Chemical Physics* **135**, 165102 (2011); doi: 10.1063/1.3654135

View online: <https://doi.org/10.1063/1.3654135>

View Table of Contents: <http://aip.scitation.org/toc/jcp/135/16>

Published by the [American Institute of Physics](#)

---

## Articles you may be interested in

[The finite state projection algorithm for the solution of the chemical master equation](#)

*The Journal of Chemical Physics* **124**, 044104 (2006); 10.1063/1.2145882

[Minimal moment equations for stochastic models of biochemical reaction networks with partially finite state space](#)

*The Journal of Chemical Physics* **143**, 244103 (2015); 10.1063/1.4937937

[Multivariate moment closure techniques for stochastic kinetic models](#)

*The Journal of Chemical Physics* **143**, 094107 (2015); 10.1063/1.4929837

[A study of the accuracy of moment-closure approximations for stochastic chemical kinetics](#)

*The Journal of Chemical Physics* **136**, 154105 (2012); 10.1063/1.3702848

[The chemical Langevin equation](#)

*The Journal of Chemical Physics* **113**, 297 (2000); 10.1063/1.481811

[Sensitivity estimation for stochastic models of biochemical reaction networks in the presence of extrinsic variability](#)

*The Journal of Chemical Physics* **146**, 124122 (2017); 10.1063/1.4978940

---

PHYSICS TODAY

WHITEPAPERS

### ADVANCED LIGHT CURE ADHESIVES

Take a closer look at what these environmentally friendly adhesive systems can do

READ NOW

PRESENTED BY  
 **MASTERBOND**  
ADHESIVES | SEALANTS | COATINGS

# Moment estimation for chemically reacting systems by extended Kalman filtering

J. Ruess,<sup>1,2,a)</sup> A. Miliias-Argeitis,<sup>1</sup> S. Summers,<sup>1</sup> and J. Lygeros<sup>1</sup>

<sup>1</sup>*Institut für Automatik, ETH Zürich, Zürich 8092, Switzerland*

<sup>2</sup>*Systems Biology Ph.D. Program, Life Science Zurich Graduate School, ETH Zürich, Zürich 8092, Switzerland*

(Received 6 June 2011; accepted 3 October 2011; published online 26 October 2011)

In stochastic models of chemically reacting systems that contain bimolecular reactions, the dynamics of the moments of order up to  $n$  of the species populations do not form a closed system, in the sense that their time-derivatives depend on moments of order  $n + 1$ . To close the dynamics, the moments of order  $n + 1$  are generally approximated by nonlinear functions of the lower order moments. If the molecule counts of some of the species have a high probability of becoming zero, such approximations may lead to imprecise results and stochastic simulation is the only viable alternative for system analysis. Stochastic simulation can produce exact realizations of chemically reacting systems, but tends to become computationally expensive, especially for stiff systems that involve reactions at different time scales. Further, in some systems, important stochastic events can be very rare and many simulations are necessary to obtain accurate estimates. The computational cost of stochastic simulation can then be prohibitively large. In this paper, we propose a novel method for estimating the moments of chemically reacting systems. The method is based on closing the moment dynamics by replacing the moments of order  $n + 1$  by estimates calculated from a small number of stochastic simulation runs. The resulting stochastic system is then used in an extended Kalman filter, where estimates of the moments of order up to  $n$ , obtained from the same simulation, serve as outputs of the system. While the initial motivation for the method was improving over the performance of stochastic simulation and moment closure methods, we also demonstrate that it can be used in an experimental setting to estimate moments of species that cannot be measured directly from time course measurements of the moments of other species. © 2011 American Institute of Physics. [doi:10.1063/1.3654135]

## I. INTRODUCTION

In chemically reacting systems, the time evolution of the species is often modeled by a set of ordinary differential equations that are derived from the reactions. The state of the system is then continuous and can be thought of as the concentrations of the species. This describes the system on a macroscopic scale and is well suited if molecule counts are large. However, in biology some key molecules come in very low copy numbers. Some transcripts, for example, can be present in less than one copy per cell on average.<sup>1</sup> A stochastic description of the system that takes into account the discrete nature and inherent randomness of individual molecular interactions may be more appropriate in such cases.<sup>2-4</sup>

If the chemically reacting system is well mixed and has fixed volume and temperature, it can be modeled as a Markov process with discrete state space, where the state corresponds to the number of molecules of each species.<sup>5</sup> The evolution of the probability distribution of the system is then described by the chemical master equation (CME). Analytical solution of the CME is generally impossible and approximation methods have to be used. One such method is the finite state projection algorithm,<sup>6</sup> which uses a finite truncation of the state space and gives explicit accuracy guarantees over a finite horizon. However, the number of states one needs to retain in the

truncation to attain a certain accuracy may grow very rapidly, making the method impractical for systems with more than a few species. Another method tries to keep the truncation small by adjusting it according to the direction in which the probability mass moves.<sup>7</sup> Still the truncations needed to obtain accurate approximations may have to be chosen large. Generally, stochastic simulation techniques, e.g., Gillespie's stochastic simulations algorithm (SSA),<sup>8</sup> have to be used. To obtain reasonable accuracy with this approach, many simulations have to be done and the computational time required can be prohibitively large.

Often one is only interested in the lower order moments of the probability distribution. A system for the time evolution of these moments can be derived from the CME. The state of this system is then again continuous and corresponds to the different moments and cross-moments of the number of molecules of each species. If only unimolecular reactions are present in the reacting system, the differential equations for the moments up to any order form a closed linear dynamical system whose solution can be explicitly computed. If, however, bimolecular reactions are also present, the time evolution of the moments of order  $n$  depends on moments of order up to  $n + 1$  and the moment system cannot be closed.<sup>9</sup>

One way to obtain a closed system of equations in the presence of bimolecular reactions is to approximate the exact open system for the first  $n$  moments by a closed system, where the moments of order  $n + 1$  are replaced by nonlinear

<sup>a)</sup>Electronic mail: ruess@control.ee.ethz.ch.

functions of the moments of order up to  $n$ .<sup>10</sup> Such so-called moment closure approximations can be derived in many different ways. Most moment closure methods are based on the assumption that the distribution belongs to a certain class; for example, the underlying distribution is assumed to be normal and the third central moments are replaced by zero in second order low dispersion moment closure.<sup>11</sup> Alternatively, the derivative matching method develops an approximate closed system, such that the derivatives of the approximate system match those of the exact systems up to some small error for a basis of deterministic initial conditions.<sup>12,13</sup> This method is effective for skewed distributions, if populations do not become zero with high probability. If, however, the probability of having zero molecules of one or more species is substantial, the method may fail.<sup>13</sup>

In this paper, we propose an approach that uses SSA and takes into account the additional information provided by the moment equations. To obtain an estimation of the evolution of the moments of order up to  $n$ , we replace the moments of order  $n + 1$  by estimates, obtained from a low number of SSA runs. We then use the resulting noisy, but closed system in an extended Kalman filter where estimates of the moments of order up to  $n$ , obtained from the same SSA runs, serve as measured system outputs. This approach provides an intuitive way of including all the available information without making any assumptions on the distribution. We provide simulation results to demonstrate that the proposed method can outperform both stand-alone SSA (in the sense that less SSA trajectories are required to obtain the same accuracy) and moment closure methods. Like moment closure and stochastic simulation, our method depends on knowledge of the full model, including kinetic parameters. Parameter identification in this context is the topic of on-going work.<sup>14–16</sup> See also related literature on a hybrid moment computation algorithm that switches between moment closure and stochastic simulation.<sup>17</sup>

Building on these insights, we subsequently extend our approach further, to a situation, where experimental measurements of only some of the species are available (e.g., from flow cytometry). We use a similar filtering approach to infer the moments of the unobserved species. If only unimolecular reactions are present in the reacting system, the moment equations close and the resulting moment system is exact. Our method can then be used to infer the moments of the unobserved species if they are observable from the measured moments. Clearly, a necessary condition for observability is that changes in the number of molecules of the unobserved species affect the observed species in some way. Even though sufficient observability conditions, based on the structure of the reactions, remain to be explored, for a given system, it is straightforward to check observability using standard control theory methods and to determine which species have to be measured, such that the moments of the unmeasured species are observable. If bimolecular reactions are also present, estimates of the moments needed to close the moment equations will only be available, if all the moments of order  $n$  are independent of the higher moments that contain unobserved species. In most systems, this is not the case and we have to rely on the moment closure methods to close the equations. Observability then depends on the nonlinear moment system

and on the closure method used, but can still be checked in a given situation. Our numerical results demonstrate that it is usually possible to infer the unobserved moments, up to some small remaining bias that stems from imprecise moment closure approximations.

The paper is organized as follows. In Sec. II, we introduce the concept of Kalman filtering. In Sec. III, we present the chemical master equation and the corresponding moment system and derive our method. Section IV is devoted to the reconstruction of unobserved moments from moment measurements of only some of the species. In Sec. V we provide simulation results for several examples to demonstrate the performance of our approach.

## II. KALMAN FILTERING

Consider a linear dynamical system with state vector  $x(t) \in \mathbb{R}^n$  that is perturbed by white process noise  $w(t) \in \mathbb{R}^n$

$$\dot{x}(t) = A(t) \cdot x(t) + w(t), \quad w(t) \sim N(0, Q(t)),$$

where  $A(t) \in \mathbb{R}^{n \times n}$  and  $Q(t) \in \mathbb{R}^{n \times n}$  is the covariance matrix of  $w(t)$ .

Further assume that at time points  $t_k, k = 1, \dots, l$  noisy measurements are available, according to the observation model

$$y_k = H_k \cdot x(t_k) + v_k, \quad v_k \sim N(0, R_k),$$

where  $y_k \in \mathbb{R}^m$ ,  $H_k \in \mathbb{R}^{m \times n}$  and  $R_k \in \mathbb{R}^{m \times m}$  is the covariance matrix of  $v_k \in \mathbb{R}^m$ .

The Kalman filter is a recursive algorithm that constructs estimates of the state  $x(t)$  of the system. First an *a priori* state estimate for the next measurement time point is computed using the system equations, and then the final *a posteriori* state estimate for that time point is obtained by computing a weighted average of the prediction and the measurements. The weights are chosen according to the uncertainties of the *a priori* estimate and the measurements, that results from the noise terms  $w(t)$  and  $v_k$ , respectively. If the process noise is large, compared to the measurement noise, the weighting should be such that a lot of emphasis is put on the measurements, whereas a small process noise, compared to the measurement noise, should lead to the filter putting more weight on the predictions of the model. The optimal trade-off is achieved by choosing the Kalman gain as the weighting matrix. It can be shown, that with the use of the Kalman gain, the resulting estimates are optimal in the sense that they attain the minimal mean square error (MSE) among all possible estimates.<sup>18</sup> More precisely, the Kalman filtering equations can be written as follows:

*Initialization:*

Given some initial state estimate  $x_{0|0}$  and its covariance  $P_{0|0}$ , set  $t = 0, k = 0$ .

*Prediction:*

Obtain *a priori* estimate  $x_{k+1|k}$  and its covariance by integrating the system

$$\begin{aligned} \dot{x}(t) &= A(t) \cdot x(t), \\ \dot{P}(t) &= A(t) \cdot P(t) + P(t) \cdot A(t)^T + Q(t) \end{aligned}$$

over the interval  $[t_k, t_{k+1}]$ , where  $x(t_k) = x_{k|k}$  and  $P(t_k) = P_{k|k}$ .

*Update:*

Compute the Kalman gain by

$$K_{k+1} = P_{k+1|k} \cdot H_{k+1}^T (H_{k+1} \cdot P_{k+1|k} \cdot H_{k+1}^T + R_{k+1})^{-1}$$

and the *a posteriori* estimate  $x_{k+1|k+1}$  and its covariance by

$$x_{k+1|k+1} = x_{k+1|k} + K_{k+1}(y_{k+1} - H_{k+1} \cdot x_{k+1|k}),$$

$$P_{k+1|k+1} = (I - K_{k+1} \cdot H_{k+1})P_{k+1|k}.$$

Note that the Kalman filter is built on the assumption of Gaussian noise terms. If this assumption is violated, the covariances of the estimates, computed by the Kalman filter, are only approximations of the true covariances and MSE-optimality is lost. In applications, the specific characteristics of the process noise are usually unknown and the matrix  $Q(t)$  is treated as a design parameter to balance predictions and measurements.<sup>19</sup>

Extended Kalman filtering is a version of standard Kalman filtering for nonlinear systems of the form

$$\dot{x}(t) = f(x(t)) + w(t), \quad w(t) \sim N(0, Q(t)).$$

The *a priori* estimates  $x_{k+1|k}$  can still be obtained by integrating the system equations as in the case of standard Kalman filtering. Their covariances  $P_{k+1|k}$ , however, cannot be computed analytically anymore. In extended Kalman filtering, an approximation of  $P_{k+1|k}$  is computed by linearizing  $f$  around the current estimate  $F(t) = \frac{\partial f}{\partial x}|_{x(t)}$  and integrating the linearized system

$$\dot{P}(t) = F(t) \cdot P(t) + P(t) \cdot F(t)^T + Q(t).$$

If the observation model is linear, as will be the case throughout this paper, the update equations remain the same as in the case of standard Kalman filtering.

Due to the linearization involved, the extended Kalman filter is not the minimum MSE estimator. Still, extended Kalman filtering is a common practice in nonlinear state estimation and has successfully been applied in many different situations.<sup>18,19</sup>

### III. CONSTRUCTION OF THE FILTER

Consider a chemically reacting system of  $m$  distinct species that is well-mixed and has fixed volume and temperature. Denote by  $p(x; t)$  the probability that the system is in state  $x$  at time  $t$ , where  $x = (x_1, \dots, x_m) \in \mathbb{N}^m$  and  $x_j$  is the number of molecules of species  $X_j, j = 1, \dots, m$ . Suppose that the system can undergo  $K$  different reactions and let  $a_i(x), i = 1, \dots, K$  be the propensity functions and  $\nu_i, i = 1, \dots, K$  the stoichiometric transition vectors of the  $K$  reactions. The time evolution of the probability distribution is then given by the chemical master equation<sup>20</sup>

$$\dot{p}(x; t) = -p(x; t) \sum_{i=1}^K a_i(x) + \sum_{i=1}^K p(x - \nu_i; t) a_i(x - \nu_i). \quad (1)$$

Computing the solution of the chemical master equation requires solving a system of differential equations with one

equation for each possible state of the system. In most chemically reacting systems, the state space is either very large or countably infinite and computing the solution analytically is rarely possible. Recently, the finite state projection method that approximates the full system using a finite-dimensional subsystem has been developed.<sup>6</sup> The accuracy of this approximation depends on the choice of the subsystem. In many cases, however, this subsystem has to be chosen very large in order to obtain reasonable accuracy and consequently the computational cost can be prohibitively large. Thus, SSA is in general the only method available to compute approximations of the solution of the chemical master equation.

An alternative to approximating the solution of the CME is to concentrate on the moments of the probability distribution. The time evolution of the vector  $\mu_\infty$ , containing all the central moments of  $x$ , can be derived from the chemical master equation<sup>21</sup> and is given by an infinite system of nonlinear differential equations

$$\dot{\mu}_\infty = f_\infty(\mu_\infty). \quad (2)$$

Because this equation is infinite dimensional, it cannot be solved analytically and we consider a truncation of order  $n$ , i.e., we extract from Eq. (2) the differential equations for the moments of order up to  $n$  and obtain the finite system

$$\dot{\mu} = f(\mu) + g(\bar{\mu}) \quad (3)$$

with some functions  $f$  and  $g$ , where  $\mu$  is a vector containing the central moments of order up to  $n$  and  $\bar{\mu}$  contains the central moments of order  $n + 1$ .

It is easy to show that the uncentered moments can also be computed in a similar way; indeed in this case, the resulting differential equations turn out to be linear. Even though this would simplify the approach introduced below (the extended Kalman filter can be replaced by a standard Kalman filter), it turns out that dealing with the seemingly more complex case of central moments tends to be more robust numerically. The intuition behind this empirical observation will be discussed later in this section, after the filtering method proposed in the paper has been introduced.

In moment closure methods,  $\bar{\mu}$  is approximated by some nonlinear functions of  $\mu$ . For example, in low dispersion moment closure,  $\bar{\mu}$  is replaced by zero. For second order truncations, this is equivalent to zero-cumulant moment closure, where closed systems are obtained by setting the cumulants of some order to zero. Another moment closure method is derivative matching,<sup>13</sup> where  $\bar{\mu}$  is approximated by a nonlinear function of  $\mu$ , such that the derivatives of the approximate system match those of the true system, up to some small error, for a basis of deterministic initial conditions. The drawback of the moment closure methods is that approximation errors, even if they are small, can accumulate over time. For example, replacing the third central moments by zero, when the distribution is positively skewed over the whole time course, often leads to divergence of the method.

In this work, we propose to close the system by replacing  $\bar{\mu}$  by estimates, obtained from the SSA. If the estimates used are unbiased and consistent, the resulting stochastic system is an unbiased approximation of the true moment system and the approximation error tends to zero, as the number of SSA

trajectories tends to infinity. However, if low numbers of SSA trajectories are used, the resulting statistical estimates will be inaccurate. This can have deleterious effects, if the system is sensitive to small errors in  $\bar{\mu}$ . To overcome this problem, we use the resulting stochastic system in an extended Kalman filter, where estimates of  $\mu$ , obtained from the same SSA runs, are interpreted as noisy measurements of the moment system.

Let  $M$  be the number of moments of order up to  $n$  of the  $m$  species. Given initial conditions for these moments  $\mu_0 \in \mathbb{R}^M$ , our goal is to estimate their time evolution over a horizon  $[0, T]$  for some  $T > 0$ . We will approximate  $\mu : [0, T] \rightarrow \mathbb{R}^M$  with a sequence of samples  $\{\mu(t_k)\}_{k=0}^J$  where  $0 = t_0 < t_1 < \dots < t_J = T$  is an increasing sequence of sampling times. Equally spaced sampling times are the obvious choice, though this choice is by no means essential, a fact that may be useful for filtering experimental data (Sec. IV). We will use a small number,  $N$ , of SSA runs that will be filtered through the dynamics [Eq. (3)] using an extended Kalman filter.

Our algorithm consists of four steps. In the first two steps, the estimates needed for our approach are computed and the uncertainty associated with model and measurements is chosen. In the third step, the state at the next time point is predicted by integrating the system. This prediction is then updated in the fourth step by computing a weighted average between the predictions and the measurements. Steps 3 and 4 are then repeated until the final time of interest is reached. More precisely, our algorithm works as follows.

#### Step 1: Input

Assume that we are given an initial distribution,  $p_0$ , of the state,  $x$ . Compute its moments up to order  $n$ ,  $\mu_0 \in \mathbb{R}^M$ . Fix a sequence of sampling times  $0 = t_0 < t_1 < \dots < t_J = T$ , choose a number  $N \in \mathbb{N}$  of SSA runs, and choose the model uncertainty covariance matrix  $Q : [0, T] \rightarrow \mathbb{R}^{M \times M}$ .

#### Step 2: Initialization

Execute  $N$  SSA runs, where the initial molecule counts are drawn from  $p_0$ , and use them to construct estimates  $\hat{\mu}(t_k)$ ,  $k = 1, \dots, J$  of the moments of order up to  $n$  and estimates  $\hat{\mu}(t_k)$ ,  $k = 0, \dots, J$  of the moments of order  $n + 1$ . Estimate the variances of the obtained estimates and use the results to choose the observation uncertainty covariance matrix  $R : \{t_k\}_{k=1}^J \rightarrow \mathbb{R}^{M \times M}$ . Set  $\mu_{0|0} = \mu_0$ ,  $P_{0|0} = 0 \in \mathbb{R}^{M \times M}$  and  $k = 0$ .

#### Step 3: Prediction

Calculate the *a priori* state estimate  $\mu_{k+1|k}$  and the *a priori* estimate covariance  $P_{k+1|k}$  by integrating the system

$$\begin{aligned}\dot{\bar{\mu}}(t) &= f(\bar{\mu}(t)) + g(\hat{\mu}(t_k)), \\ \dot{P}(t) &= A(t) \cdot P(t) + P(t) \cdot A(t)^T + Q(t)\end{aligned}$$

over the interval  $[t_k, t_{k+1}]$ , where  $P(t_k) = P_{k|k}$ ,  $\bar{\mu}(t_k) = \mu_{k|k}$ , and  $A(t) = \frac{\partial f}{\partial \mu}|_{\bar{\mu}(t)}$  is the linearization of  $f(\mu)$  around the current state estimate.

#### Step 4: Update

Calculate the *a posteriori* state estimate  $\mu_{k+1|k+1}$  and the *a posteriori* estimate covariance  $P_{k+1|k+1}$  as

$$\begin{aligned}K_{k+1} &= P_{k+1|k}(P_{k+1|k} + R(t_{k+1}))^{-1}, \\ \mu_{k+1|k+1} &= \mu_{k+1|k} + K_{k+1}(\hat{\mu}(t_{k+1}) - \mu_{k+1|k}), \\ P_{k+1|k+1} &= (I - K_{k+1})P_{k+1|k},\end{aligned}$$

where  $I$  denotes the identity matrix.

If  $k < J - 1$ , set  $k = k + 1$  and go back to Step 3.

#### Step 5: Output

Sequence of estimates  $\mu_{k|k}$ ,  $k = 0, \dots, J$ .

Comments on the algorithm:

*a. Estimating the initial condition.* In Step 1, the computation of  $\mu_0$  may be difficult for certain distributions  $p_0$ . If this is the case, one can use estimates  $\hat{\mu}_0$ , computed from the SSA runs, as initial condition. The initial uncertainty  $P_{0|0}$  should then be chosen as the covariance matrix of  $\hat{\mu}_0$ .

*b. Estimating the moments between sampling times.* The prediction in Step 3 could be replaced by multiple smaller steps over smaller time intervals, before the next update step is taken. The algorithm would then also return moment estimates between the sampling times. Since at this stage the sampling times can be arbitrarily chosen, this is not necessary. However, in an experimental setup, like the one considered in Sec. IV, where the sampling times correspond to the time points at which measurements are taken, this fact can be useful.

*c. Choosing the covariance matrices.* In Steps 1 and 2, the covariance matrices  $Q(t)$  and  $R(t_k)$  that describe the uncertainties associated with the model and the observations have to be chosen. The model uncertainty is propagated over each time interval  $[t_k, t_{k+1}]$  by linearizing the system function  $f(\mu)$  around the current state estimate and  $Q(t)$  should represent the linearization error, as well as the noise that is introduced to the model by using  $\hat{\mu}(t_k)$  to close the equations. As this depends on the specific structure of the system, it is *a priori* unclear how to choose  $Q(t)$ . However, in many applications, extended Kalman filtering turns out to be robust with respect to the choice of the model uncertainty (see supplementary material<sup>22</sup> for a related discussion). In our applications, choosing  $Q(t)$  as a constant function of time was sufficient.  $R(t_k)$  represents the uncertainty about the observations  $\hat{\mu}(t_k)$  and is adapted to the number of SSA trajectories. Assume, for example, that we estimate the mean of species  $X$  at some time point  $t_k$  by  $\hat{\mu}_x = \frac{1}{N} \sum_{i=1}^N x_i$ , where  $x_i$ ,  $i = 1, \dots, N$  are the molecule counts of species  $X$  at time  $t_k$  of the  $N$  SSA trajectories. Then, since the SSA trajectories are independent, the variance of  $\hat{\mu}_x$  equals the variance of species  $X$  at time  $t_k$  divided by  $N$  and in the matrix  $R(t_k)$  we, therefore, use the empirical estimate of the variance of  $X$  in the diagonal element corresponding to  $\hat{\mu}_x$ . Further, we assume that correlations of the elements of  $\hat{\mu}(t_k)$  are small and set the off-diagonal entries to zero.

*d. Comparison to SSA and moment closure.* Compared to SSA, our approach requires integrating the closed moment

equations, propagating the covariance matrix and computing the weighted averages in addition to computing the SSA trajectories. However, solving ordinary differential equations usually requires less computational time than stochastic simulation, especially for stiff systems that involve reactions at different time scales. Therefore, the additional computational effort is negligible for many systems. Compared to moment closure, our method is of course more computationally intensive. Our numerical results, however, demonstrate that the additional steps make the method applicable to systems where pure moment closure methods fail.

*e. Justification of the use of the central moment equations.*

If we derive the equations of the uncentered moments  $\mu_u$  from the master equation, then instead of Eq. (3), we obtain a linear system

$$\dot{\mu}_u = A\mu_u + B\bar{\mu}_u, \quad (4)$$

where  $\bar{\mu}_u$  contains the uncentered moments of order  $n + 1$ . To close this system, estimates for  $\bar{\mu}_u$  would be required. Compared to estimates of the central moments, these estimates usually have a larger absolute variance and a stronger influence on the system. To illustrate this, consider the estimation of the third moment of a random variable that consists of a deterministic part  $d$  and a zero mean extrinsic noise term  $v$ . The third uncentered moment can then be written as  $\mathbb{E}[(d + v)^3] = d^3 + d \cdot \mathbb{E}[v^2] + \mathbb{E}[v^3]$  and the variance of any estimate of this moment, based on realizations of the random variable, will depend on the size of the deterministic part  $d$  through the second term in the sum. Contrary to that, the third central moment  $\mathbb{E}[(d + v - \mathbb{E}[d + v])^3] = \mathbb{E}[v^3]$  depends only on the third moment of the noise term. Hence, the absolute variance of estimates of the third uncentered moment scales with the mean of the process, whereas the absolute variance of estimates of the third central moment is independent of the mean. Therefore, using the central moment equations and estimates of the central moments naturally helps to keep the influence of the noise on the system low.

#### IV. ESTIMATING THE MOMENTS FROM MEASUREMENTS OF ONLY SOME OF THE SPECIES

In most realistic chemically reacting systems, it is impossible or very difficult to measure all species, and initial conditions are unknown. If only unimolecular reactions are present in the reacting system, the moment equations close and the problem of estimating the unobserved moments from the measured species reduces to inferring the unmeasured states of a linear system with unknown initial conditions. We can then use our method to infer the moments of the unobserved species, if they are observable from the moments of the measured species. If bimolecular reactions that lead to non-closed moment dynamics are also present and all moments of order  $n$  are independent of the moments of order  $n + 1$  that contain unobserved species, we can close the moment equations in the same way as in Sec. III. In most situations, this is not the case and we have to rely on moment closure approximations to obtain closed dynamics. Our numerical re-

sults suggest that even in this case it is possible to estimate the moments of all species, albeit with some bias.

Starting with the system of equations [Eq. (3)], we approximate the moments in  $\bar{\mu}$  by a function of the  $M$  lower order moments and obtain an approximate closed system

$$\dot{v} = f(v) + \phi(v), \quad (5)$$

where  $v \in \mathbb{R}^M$  is a vector containing the approximate moments of order up to  $n$ .

Consider that measurements of the distribution (or at least the moments) of  $l$  of the  $m$  species are available at arbitrary times  $0 \leq t_1 < t_2 < \dots < t_J = T$  and denote by  $L$  the number of moments of order up to  $n$  of the measured species. This can be written as

$$y(t_k) = H \cdot \mu(t_k) + w(t_k), \quad k = 1, \dots, J, \quad (6)$$

where  $y(t_k) \in \mathbb{R}^L$  is a vector containing  $L$  measured moments,  $\mu(t_k) \in \mathbb{R}^M$  is the vector containing the true  $M$  moments,  $w(t_k) \in \mathbb{R}^L$  is the measurement noise, and  $H \in \{0, 1\}^{L \times M}$ . We then run prediction steps of the filter in the time intervals between the measurements and update steps at the time points, where measurements are available. More precisely our algorithm works as follows.

*Step 1: Input*

If the moment dynamics are open, choose a moment closure method to close the moment system. Fix the model uncertainty covariance matrix  $Q : [0, T] \rightarrow \mathbb{R}^{M \times M}$  and the observation uncertainty covariance matrix  $R : \{t_k\}_{k=1}^J \rightarrow \mathbb{R}^{L \times L}$ .

*Step 2: Initialization*

Guess the initial condition  $\mu_{0|0} \in \mathbb{R}^M$ , set the initial uncertainty  $P_{0|0} \in \mathbb{R}^{M \times M}$  to some positive definite matrix that reflects the uncertainty about the initial guess and set  $k = 0$ .

*Step 3: Prediction*

Calculate the *a priori* state estimate  $\mu_{k+1|k}$  and the *a priori* estimate covariance  $P_{k+1|k}$  by integrating the approximate system

$$\dot{\tilde{v}}(t) = f(\tilde{v}(t)) + \phi(\tilde{v}(t)), \quad (7)$$

$$\dot{P}(t) = A(t) \cdot P(t) + P(t) \cdot A(t)^T + Q(t) \quad (8)$$

over the interval  $[t_k, t_{k+1}]$ , where  $P(t_k) = P_{k|k}$ ,  $\tilde{v}(t_k) = \mu_{k|k}$ , and  $A(t) = \frac{\partial(f+\phi)}{\partial v}|_{\tilde{v}(t)}$  is the linearization of  $(f + \phi)(v)$  around the current state estimate.

*Step 4: Update*

Calculate the *a posteriori* state estimate  $\mu_{k+1|k+1}$  and the *a posteriori* estimate covariance  $P_{k+1|k+1}$  as

$$\begin{aligned} K_{k+1} &= P_{k+1|k} \cdot H^T (H \cdot P_{k+1|k} \cdot H^T + R(t_{k+1}))^{-1}, \\ \mu_{k+1|k+1} &= \mu_{k+1|k} + K_{k+1}(y(t_{k+1}) - H \cdot \mu_{k+1|k}), \\ P_{k+1|k+1} &= (I - K_{k+1} \cdot H)P_{k+1|k}, \end{aligned}$$

where  $I$  denotes the identity matrix.

If  $k < J - 1$ , set  $k = k + 1$  and go back to Step 3.

*Step 5: Output*

Sequence of *a posteriori* estimates  $\mu_{k|k}$ ,  $k = 0, \dots, J$ .

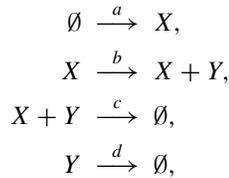
Here, the model uncertainty covariance matrix  $Q(t)$  should represent the performance of the moment closure method. If the moment equations close and no moment closure method is needed, it can be set to zero. The observation uncertainty covariance matrix  $R(t_k)$  should reflect the covariance of the measurements as before.

**V. EXAMPLES**

In this section, we illustrate our method by considering some examples. For simplicity, we only consider truncations of order 2 and replace the third order central moments by their estimated counterparts, obtained from the SSA.

*Example 1:*

Consider the following system:



where  $a = 2 \times 10^{-3}$ ,  $b = 2 \times 10^{-1}$ ,  $c = 2 \times 10^{-4}$ , and  $d = 2 \times 10^{-3}$ .

$X$  can be thought of as a transcription factor for some gene and  $Y$  as the gene product. When  $X$  is present,  $Y$  is produced. Saturation of the promoter of the gene is neglected, since  $X$  comes in very low copy numbers.  $Y$  can bind  $X$  and, hence, inhibit its own production and is slowly degraded or converted to some other species that has no influence on the reactions of  $X$  and  $Y$ . Further, we assume that initially there are no molecules in the system. If the molecules of species  $X$  arrive in rare events, the distribution of the system is skewed and the probability that no molecules of  $X$  are present is large.

From the master equation, we derived the moment equations for the central moments of order up to two. Denote by  $\mu_x, \mu_y, V_x, V_y, V_{xy}, V_{x2y}$ , and  $V_{xy2}$  the mean of  $X$ , mean of  $Y$ , variance of  $X$ , variance of  $Y$ , covariance of  $X$  and  $Y$ ,  $E[(X - E[X])^2(Y - E[Y])]$  and  $E[(X - E[X])(Y - E[Y])^2]$ , respectively. The moment dynamics are then given by

$$\begin{aligned} \dot{\mu}_x &= a - cV_{xy} - c\mu_x\mu_y, \\ \dot{\mu}_y &= b\mu_x - cV_{xy} - c\mu_x\mu_y - d\mu_y, \\ \dot{V}_x &= a - 2cV_{xy}\mu_x - 2c\mu_yV_x + cV_{xy} + c\mu_x\mu_y - 2cV_{x2y}, \\ \dot{V}_y &= -2cV_{xy}\mu_y - 2c\mu_xV_y + (c + 2b)V_{xy} + c\mu_x\mu_y \\ &\quad + b\mu_x - 2dV_y + d\mu_y - 2cV_{xy2}, \\ \dot{V}_{xy} &= -cV_{xy}\mu_y - c\mu_xV_y - cV_{xy}\mu_x - c\mu_yV_x \\ &\quad + (c - d)V_{xy} + c\mu_x\mu_y + bV_x - cV_{x2y} - cV_{xy2}. \end{aligned}$$

To close the equations, we replaced  $V_{x2y}$  and  $V_{xy2}$  by estimates obtained from stochastic simulation. We chose the sam-

pling times  $t_0 = 0, t_1 = 1, \dots, t_J = J$ , where  $J = 500$  and the model uncertainty covariance matrix as

$$Q(t) = 10^{-9} \times \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 100 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad t \in [0, 500].$$

This choice of  $Q(t)$  was motivated as follows: the time derivatives of the means of  $X$  and  $Y$  do not depend on third order moments and the uncertainty associated with their equations only results from linearization error. This is small because no more than bilinear terms are present in the equations and the sampling times are close to each other. The uncertainty associated with the variance of  $X$  was chosen the largest because the variance of  $X$  is the smallest of the second moments and presumably suffers mostly from low signal-to-noise ratios. A comparison of different choices of  $Q(t)$  is given in the supplementary material.<sup>22</sup> The measurement uncertainty covariance matrix  $R(t_k)$  was estimated as described in Sec. III.

Figure 1 gives a typical result for the means of  $X$  and  $Y$  computed from 200 SSA trajectories (green) and the filtering approach of Sec. III using the same 200 SSA trajectories (blue), compared to the solution of the approximate system, obtained from second order derivative matching (cyan). In derivative matching, the closure function involves divisions by the lower order moments that are initially zero. Therefore, we set the initial condition to 0.01 for all moments to compute the solution. The red line shows estimates computed using 100 000 SSA trajectories, which can be considered as the true values of the corresponding quantities. To demonstrate the benefits of including the update steps of the extended Kalman filter, we also computed estimates of the means of  $X$  and  $Y$  by replacing  $\bar{\mu}$  by the estimates  $\hat{\mu}$  and integrating the resulting system without performing update steps (purple). It can be seen that these estimates are still relatively close to the truth. However, as the estimates are never corrected, errors accumulate over time and the system may be driven far away from its true value, especially in systems where the third order moments have a stronger influence on the first and second order moments. Results for the variances of  $X$  and  $Y$  can be found in the supplementary material.<sup>22</sup>

To quantify the performance of our approach, we computed the  $L_2$ -distances from the “true” 100 000 SSA moments normalized by the integral of the true solution of the results of the filter and of SSA on its own, averaged over 500 runs, using  $i = 20, 40, \dots, 400$  SSA trajectories. In Figure 2, the number of SSA trajectories is plotted against the logarithm of the  $L_2$ -error of filter (blue) and SSA on its own (green). For all different numbers of SSA trajectories and all moments, the filter performed better. For example, the error of the estimate of the mean of  $X$  computed from 400 SSA trajectories was about three times bigger than the error of the filter using only 100 SSA trajectories.

We now turn our attention to inferring mean and variance of  $X$  from observations of  $Y$  only using the algorithm

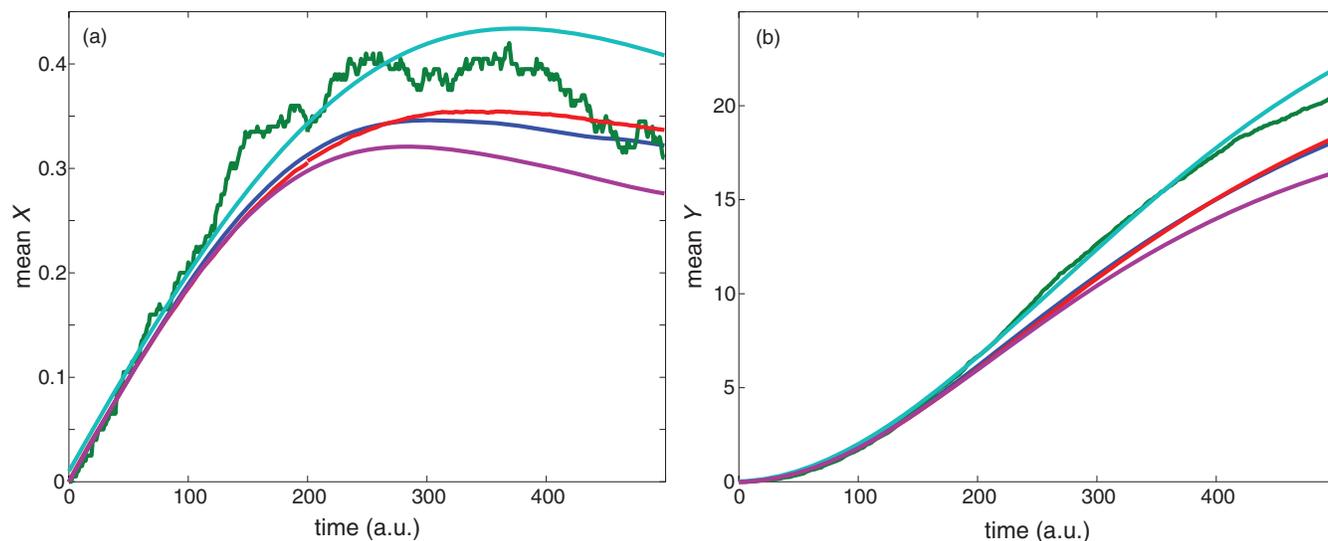


FIG. 1. Example 1. Estimated time evolution of the means of  $X$  (a) and  $Y$  (b) by filter (blue), SSA (green) and second order derivative matching (cyan), compared to the estimation obtained from 100 000 SSA trajectories (red). The purple line gives estimates of the means of  $X$  and  $Y$ , computed by leaving out the update step of the filter.

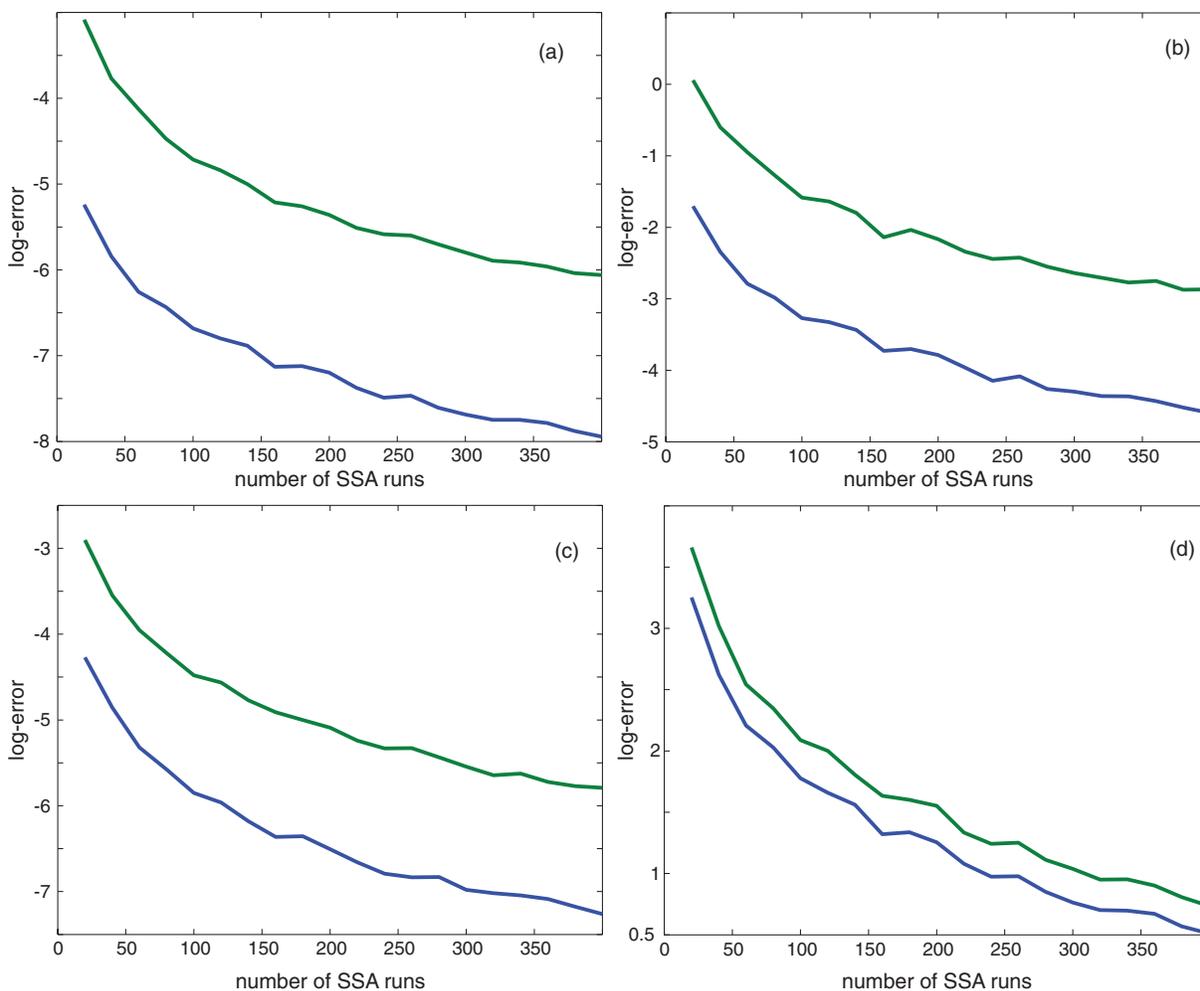


FIG. 2. Example 1. Logarithm of the  $L_2$ -errors of the estimates of the means of  $X$  (a) and  $Y$  (b) and the variances of  $X$  (c) and  $Y$  (d) averaged over 500 runs of filter (blue) and SSA (green).

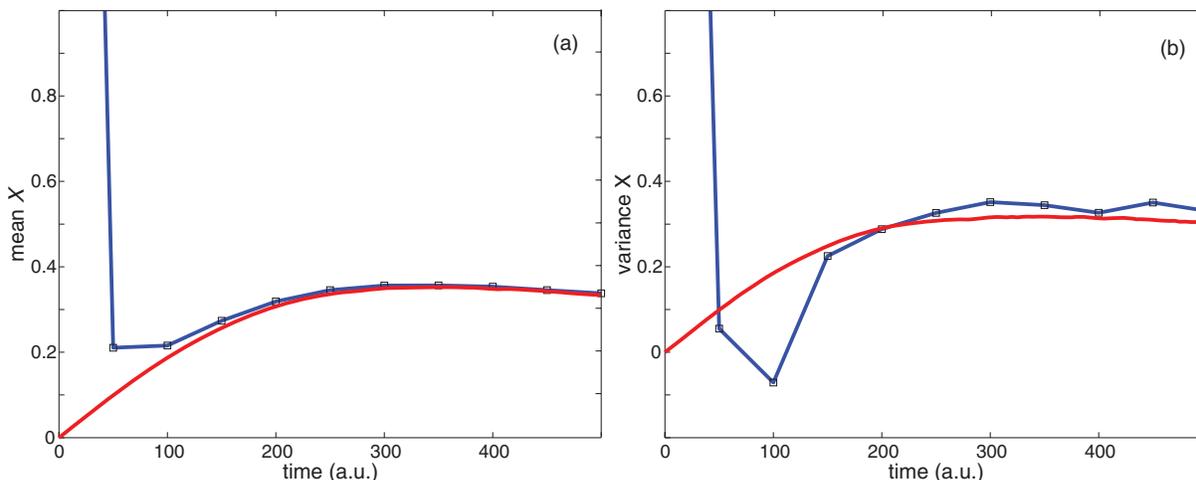


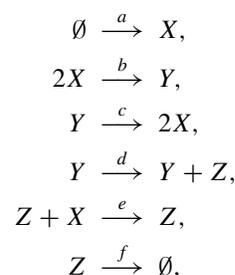
FIG. 3. Example 1. Mean (a) and variance (b) of  $X$  estimated from observations of  $Y$  (blue), compared to estimates computed using 100 000 SSA trajectories, that were started at the true initial conditions (red). Initial conditions for the filtering approach were randomly chosen far from their true values (out of the scale of the plots).

in Sec. IV. We assume that we have observations of all the moments of  $Y$  available every 50 time steps, e.g., from flow cytometry. If the time units are seconds (which corresponds to a half-life of  $Y$  of approximately 6 minutes), this means measurements have to be taken every 50 s. To simulate this situation, we ran 10 000 SSA trajectories and used them to compute estimates of mean and variance of  $Y$ . Higher order moments of  $Y$  do not appear in any equation and are not required. We randomly picked initial conditions and closed the moment equations by using zero cumulant moment closure. The model uncertainty covariance matrix  $Q(t)$  was chosen as above. Mean and variance of  $X$  are both recovered very fast, as can be seen in Figure 3. The mean is close to its true value after only two measurement updates. The variance required three update steps to get close to its true value. Note that SSA and moment closure methods require knowledge about the initial conditions and are not applicable here.

#### Example 2:

In this example, we chose a more difficult situation, where the influence of the third order moments on the lower

order moments is stronger. Consider the following system:



where  $a = 2 \times 10^{-3}$ ,  $b = 2 \times 10^{-4}$ ,  $c = 2 \times 10^{-4}$ ,  $d = 2 \times 10^{-1}$ ,  $e = 2 \times 10^{-4}$ , and  $f = 2 \times 10^{-3}$ .

$X$  is produced in rare events and can form dimers  $Y$ . The dimers induce production of  $Z$  and can also dissociate.  $Z$  inhibits its own production by catalyzing the degradation of  $X$  and is degraded or converted to some other species that has no influence on the reactions of  $X$ ,  $Y$ , and  $Z$ . We further assume that initially there are no molecules in the system. In this system, five third order moments affect the dynamics of the second order moments and their influence is relatively strong. Further, since  $X$  is rarely produced and degraded fast, if  $Z$  is

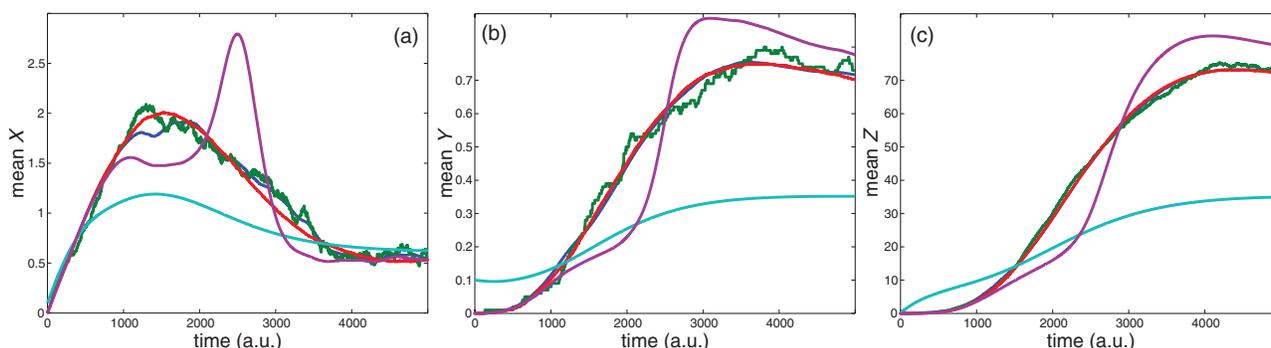


FIG. 4. Example 2. Estimated time evolution of the means of  $X$  (a),  $Y$  (b), and  $Z$  (c) by filter (blue), SSA (green), and second order derivative matching (cyan), compared to the estimation obtained from 10 000 SSA trajectories (red). The purple line gives estimates of the means of  $X$ ,  $Y$ , and  $Z$ , computed by leaving out the update step of the filter.

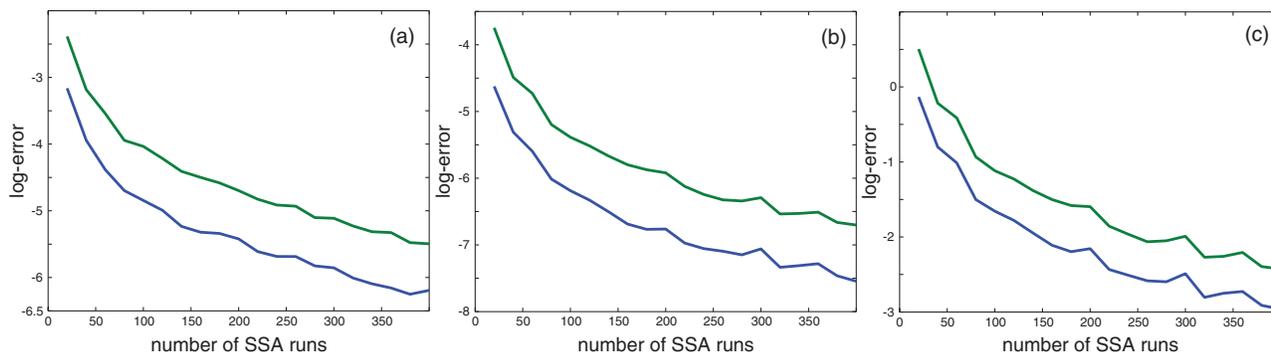


FIG. 5. Example 2. Logarithm of the  $L_2$ -errors of the estimates of the means of  $X$  (a),  $Y$  (b), and  $Z$  (c) averaged over 200 runs of filter (blue) and SSA (green).

present in large amounts, the means of both  $X$  and its dimer  $Y$  are very low. The equations for the first and second order moments and all design parameters of the filter that are not given here, for this and all the following examples, can be found in the supplementary material.<sup>22</sup>

Figure 4 gives a typical result for the means of  $X$ ,  $Y$ , and  $Z$  computed from 100 SSA trajectories (green) and the filtering approach using the same 100 SSA trajectories (blue), compared to the estimates calculated using 10 000 SSA trajectories (red), which can be considered as the true means. Second order derivative matching (cyan) fails to capture the dynamics of this system. Again, we also computed estimates, obtained by leaving out the update step of the filter (purple). Because the influence of the third order moments is relatively strong in this example, the estimates are far worse than the estimates of the filter. Further, differences in the SSA realizations that are used to close the equations may lead to very different estimates, whereas the filter estimates are much more robust (see supplementary material<sup>22</sup>).

Again we computed  $L_2$ -errors of the results of filter and SSA on its own, averaged over 200 runs, using  $i = 20, 40, \dots, 400$  SSA trajectories. In Figure 5, the number of SSA trajectories is plotted against the logarithm of the  $L_2$ -error of the estimates of the means of the three species obtained by the filter (blue) and by SSA on its own (green). For all different num-

bers of SSA trajectories, the filter gives better estimates for the means. The estimates of the second order moments, however, could not be improved much. This is because the noise coming from the estimated third order moments is affecting the second order moments strongly, and the small signal-to-noise ratios make it impossible to gain much information from the equations.

To simulate an experimental setting, we assumed that only species  $Z$  can be measured every 100 time steps and used our approach to infer the moments of  $X$  and  $Y$ . To estimate the moments of  $Z$ , we used 10 000 SSA trajectories. Again, we randomly picked initial conditions and closed the moment equations using zero cumulant moment closure, i.e., we set the third central moments to zero. The true distribution, however, is skewed and the influence of the true third central moments cannot be separated from the influence of the unobserved lower order moments. Therefore, we can only infer the unobserved moments up to some small bias that stems from the imprecise moment closure approximation. The results for the means of  $X$  and  $Y$  are plotted in Figure 6. Because  $Z$  is only changed in unimolecular reactions with propensities that only depend on  $Y$  and  $Z$  itself, the mean of  $Y$  is recovered very fast and without bias. The mean of  $X$  takes longer to get close to its true value and the steady state mean is underestimated.

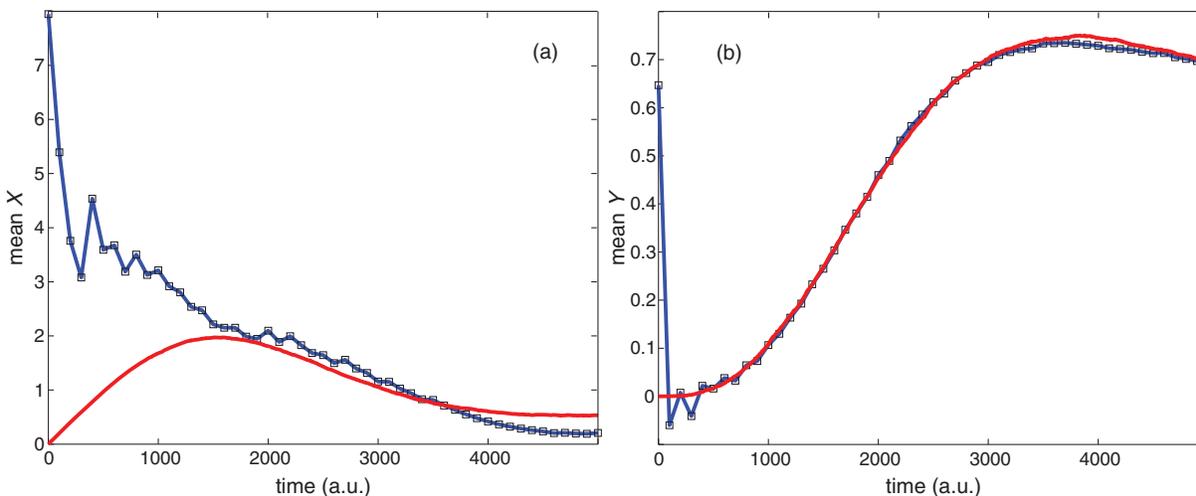


FIG. 6. Example 2. Means of  $X$  (a) and  $Y$  (b) estimated from observations of  $Z$  (blue), compared to estimates obtained from 10 000 SSA trajectories, that were started at the true initial conditions (red).

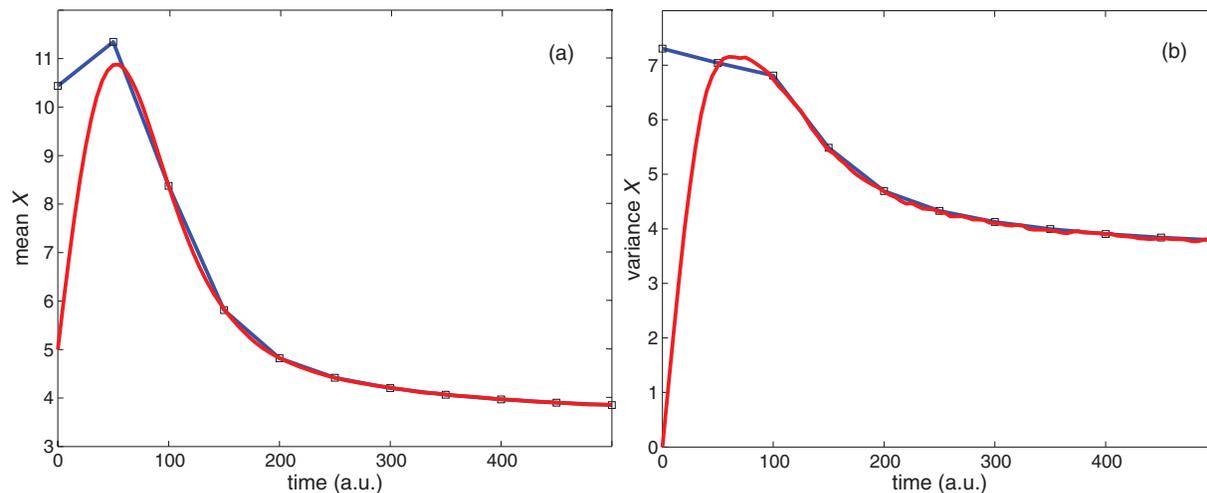


FIG. 7. Example 3. Mean (a) and variance (b) of  $X$  estimated from observations of  $Y$  (blue), compared to estimates obtained from 100 000 SSA trajectories, that were started at the true initial conditions (red).

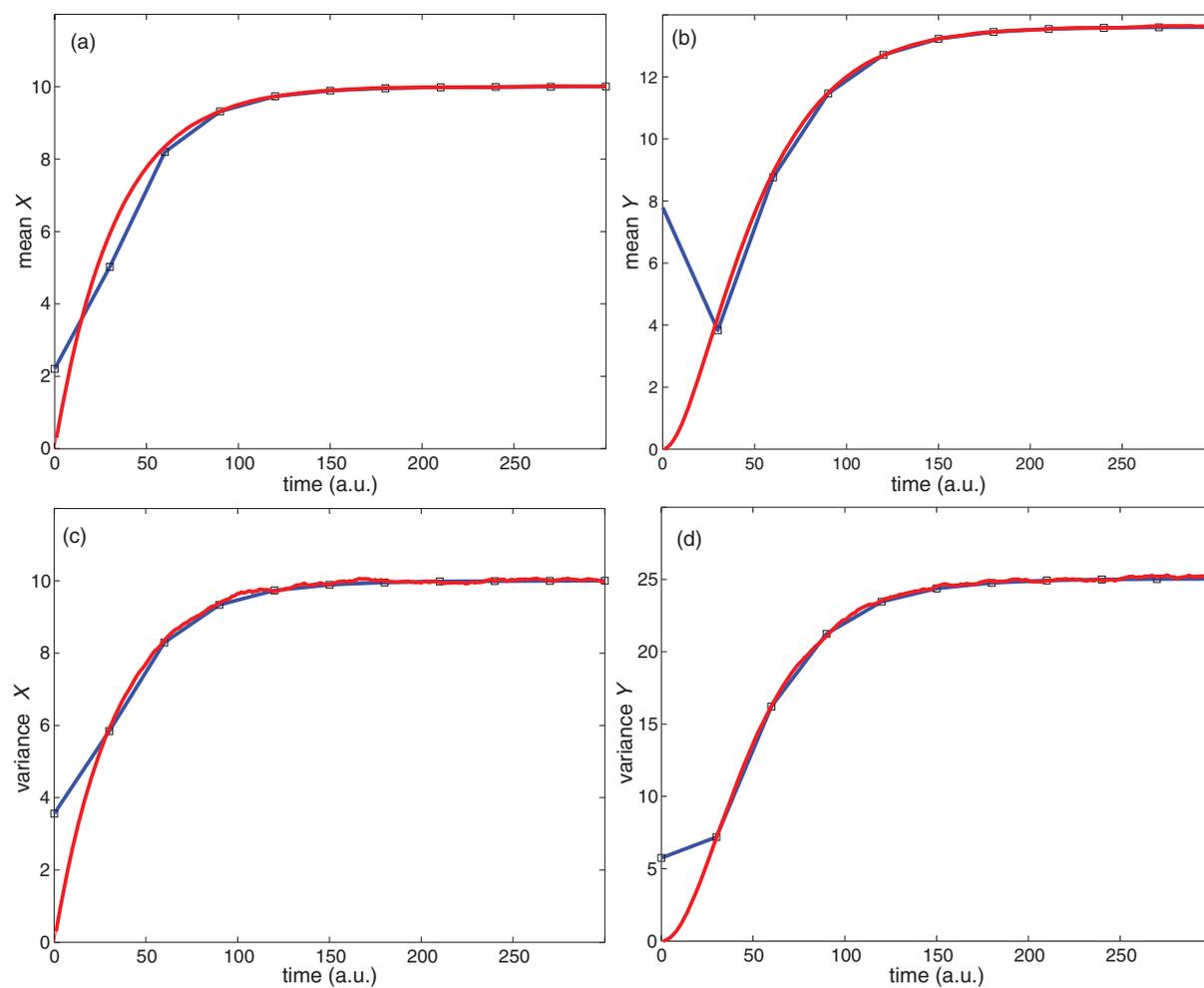


FIG. 8. Example 4. Means [(a) and (b)] and variances [(c) and (d)] of  $X$  and  $Y$  estimated from observations of  $Z$  (blue), compared to the means obtained from 100 000 SSA trajectories, that were started at the true initial conditions (red).

*Example 3:*

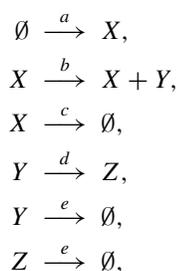
As third example, we consider the same reactions as in Example 1, but change rates and initial conditions such that the probability of having zero molecules is small for both species and the distribution is less skewed. In this example, the moment closure methods lead to good approximations of the true moment dynamics and our method of Sec. IV can be used to infer the moments of unobserved species very precisely and without significant bias, even though moment closure is required to obtain closed dynamics.

We consider the same reactions as in Example 1 having initially five molecules of both species and that  $X$  is produced at rate  $a = 2 \times 10^{-1}$ . The rates of the other reactions are assumed to be the same as in Example 1. The sampling times and the model uncertainty covariance matrix  $Q(t)$  were chosen as in Example 1. Again, we derived truncations of order 2 and ran the filter of Sec. III using 50 SSA trajectories. The estimates of all the moments computed by filter and zero cumulant moment closure were indistinguishable from the true means, hence both methods perform very well.

As in Example 1 we assumed that only species  $Y$  can be measured experimentally every 50 time steps. To infer the unobserved moments from the moments of  $Y$ , we closed the equations using zero cumulant moment closure. Again we ran 10 000 SSA trajectories and used them to compute estimates of mean and variance of  $Y$ . The results are plotted in Figure 7. The filter recovers the unobserved moments very fast and without bias because the moment closure approximation used is accurate.

*Example 4:*

As a final example, we consider a model of gene expression that is motivated by real experiments.



where  $a = 0.3, b = 0.066, c = 0.03, d = 0.0419$ , and  $e = 0.0066$ .

At time zero, a gene is activated and production of the mRNA  $X$  is induced.  $X$  is then translated to yield protein  $Y$  which matures into fluorescent protein  $Z$ .  $X$  is degraded at rate  $c$ .  $Y$  and  $Z$  are also degraded and have the same degradation rate  $e$  because they are different functional forms of the same protein.

In this system, all reactions are unimolecular and the moment equations close. If initial conditions are known, the dynamics of the moments can be computed by solving the closed moment system. To simulate an experimental setting, we assumed that initial conditions are unknown and that only the distribution of the fluorescent protein  $Z$  is measured by flow cytometry every 30 time steps. If the time units are seconds (which corresponds to a half-life of the protein of approx-

imately 2 min), this means measurements have to be taken every 30 s. We used 10 000 SSA trajectories to compute estimates of mean and variance of  $Z$  at those time points, randomly picked initial conditions, and used our approach to estimate the moments of  $X$  and  $Y$ . The model uncertainty covariance matrix  $Q(t)$  was chosen as the zero matrix. In Figure 8, the resulting estimates of the means and variances of  $X$  and  $Y$  (blue) are compared to estimates computed from 100 000 SSA trajectories (red). All moments are recovered very fast, more or less after the first measurement becomes available. Since the moment equations close, the system used in the filter is exact, and once the estimates have reached their true values, no further measurements would be required.

**VI. CONCLUSIONS AND FUTURE WORK**

We presented an algorithm for the estimation of the moment dynamics of chemically reacting systems that uses both stochastic simulation and the differential equations given by the chemical master equation. Our numerical results suggest that the proposed method can outperform the use of SSA on its own, provides a fast way of estimating the moment dynamics accurately and is especially useful in systems, where molecule counts have a high probability of becoming zero, situations where standard moment closure methods tend to fail, and no good alternative to stochastic simulation is available. We also showed how the dynamics of the moments of interest can be estimated if only some of the species are observed in experiments. Although in this setup we relied on moment closure approximations for systems that contain bimolecular reactions, our approach provided good results, even in situations where moment closure methods lead to inaccurate approximations.

In the examples, we only considered truncations of order two. For any order of truncation  $n$ , only the moments of order  $n$  depend directly on the moments of order  $n + 1$  that are replaced by estimates. Therefore, the influence of the noisy estimates on the lower order moments decreases as the order of truncation increases. This suggests that increasing the order of truncation could improve the estimates of the lower order moments. On the other hand, for higher order truncations estimates of higher order moments would be required which have a larger variance. This could possibly reduce the accuracy of the estimates of the lower order moments. Which effect prevails might depend on the system in question. For example, in the system of Example 1, a first order truncation is already sufficient to obtain good estimates for the means of the species (results not shown). A second order truncation then adds good estimates for the variances but does not improve the estimates of the means significantly.

Other classes of systems where moment closure methods are often inaccurate are systems with multi-stability or oscillations. Studying performance of our method for such systems is a direction of current research.

Further we are planning to use the moment equations and filtering techniques to do parameter identification in stochastic systems and filtering combined with finite-state space truncations to estimate the time evolution of the whole probability distribution of the system. Another direction of possible

future research is the derivation of observability conditions based on the structure of the system.

## ACKNOWLEDGMENTS

The authors would like to thank Professor Robert Bitmead for helpful discussions.

- <sup>1</sup>M. Carter, A. Sharov, V. VanBuren, D. Dudekula, C. Carmack, C. Nelson, and M. Ko, *Genome Biol.* **6**, R61 (2005).
- <sup>2</sup>H. McAdams and A. Arkin, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 814 (1997).
- <sup>3</sup>J. Hasty, J. Pradines, M. Dolnik, and J. Collins, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2075 (2000).
- <sup>4</sup>M. Samoilov and A. Arkin, *Nat. Biotechnol.* **24**, 1235 (2006).
- <sup>5</sup>D. Gillespie, *Physica A* **188**, 404 (1992).
- <sup>6</sup>B. Munsky and M. Khammash, *J. Chem. Phys.* **124**, 044104 (2006).
- <sup>7</sup>V. Wolf, R. Goel, and T. Henzinger, *BMC Systems Biology* **4**, 42 (2010).
- <sup>8</sup>D. Gillespie, *J. Comput. Phys.* **22**, 403 (1976).
- <sup>9</sup>It is generally accepted that chemical reactions occur in a series of elementary reactions and reactions containing more than two molecules can be expressed as a combination of mono- and bimolecular reactions. Therefore, we will consider networks that contain at most bimolecular reactions.
- <sup>10</sup>J. Hespanha, "Moment closure for biochemical networks," in *Proceedings of the IEEE 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP), St. Julians, Malta, 2008* (IEEE, New York, 2008), pp. 142–147.
- <sup>11</sup>P. Whittle, *J. R. Stat. Soc. Ser. B (Methodol.)* **19**, 268 (1957).
- <sup>12</sup>J. Hespanha, *IEEE Proc.: Control Theory Appl.* **153**, 520 (2007).
- <sup>13</sup>A. Singh and J. Hespanha, *IEEE Trans. Autom. Control* **56**, 414 (2011).
- <sup>14</sup>B. Munsky, B. Trinh, and M. Khammash, *Mol. Syst. Biol.* **5**, 318 (2009).
- <sup>15</sup>B. Munsky and M. Khammash, *IET Syst. Biol.* **4**, 356 (2010).
- <sup>16</sup>G. Lillacci and M. Khammash, *PLOS Comput. Biol.* **6**, e1000696 (2010).
- <sup>17</sup>Y. Zhao, J. Kim, and J. Hespanha, "Hybrid moment computation algorithm for biochemical reaction networks," in *Proceedings of the 49th IEEE Conference on Decision and Control (CDC), Atlanta, GA, USA, 2010* (IEEE, New York, 2010), pp. 1693–1698.
- <sup>18</sup>A. Jazwinski, *Stochastic Processes and Filtering Theory* (Academic, New York, 1970), Vol. 63.
- <sup>19</sup>B. La Scala, R. Bitmead, and M. James, *Math. Control, Signals, Syst.* **8**, 1 (1995).
- <sup>20</sup>N. van Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 2007).
- <sup>21</sup>C. Lee, K. Kim, and P. Kim, *J. Chem. Phys.* **130**, 134107 (2009).
- <sup>22</sup>See supplementary material at <http://dx.doi.org/10.1063/1.3654135> for additional figures and discussions on the examples.