BioSystems xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems

Adaptive moment closure for parameter inference of biochemical reaction networks

Christian Schilling [a], Sergiy Bogomolov [b], Thomas A. Henzinger [b], Andreas Podelski [a], Jakob Ruess [b,*]

[a] University of Freiburg, 79110 Freiburg, Germany
[b] IST Austria, 3400 Klosterneuburg, Austria

ARTICLE INFO

Article history:
Received 28 November 2015
Received in revised form 30 June 2016
Accepted 12 July 2016
Available online xxx

Keywords:
Stochastic reaction networks
Continuous-time Markov chains
Parameter inference
Moment closure

ABSTRACT

Continuous-time Markov chain (CTMC) models have become a central tool for understanding the dynamics of complex reaction networks and the importance of stochasticity in the underlying biochemical processes. When such models are employed to answer questions in applications, in order to ensure that the model provides a sufficiently accurate representation of the real system, it is of vital importance that the model parameters are inferred from real measured data. This, however, is often a formidable task and all of the existing methods fail in one case or the other, usually because the underlying CTMC model is high-dimensional and computationally difficult to analyze. The parameter inference methods that tend to scale best in the dimension of the CTMC are based on so-called moment closure approximations. However, there exists a large number of different moment closure approximations and it is typically hard to say a priori which of the approximations is the most suitable for the inference procedure. Here, we propose a moment-based parameter inference method that automatically chooses the most appropriate moment closure method. Accordingly, contrary to existing methods, the user is not required to be experienced in moment closure techniques. In addition to that, our method adaptively changes the approximation during the parameter inference to ensure that always the best approximation is used, even in cases where different approximations are best in different regions of the parameter space.

1. Introduction

With the advancement of measurement technologies for biochemical processes in the last decades, quantitative mathematical modeling of biochemical reaction networks has continuously increased in importance (Bertaux et al., 2014; Neuert et al., 2013; Ruess et al., 2015). Chemical reactions inside cells, where some of the reacting species may be present in very low amounts of molecules, are inherently driven by random fluctuations (Hasty et al., 2000; McAdams and Arkin, 1997; Samoilov and Arkin, 2006; Munsky et al., 2009). Accordingly, an accurate mathematical model should take this stochasticity into account. The most widely used class of stochastic models in this context are continuous-time Markov chains (CTMCs) (Goutsias and Jenkinson, 2013). The advantage of these models is that they are easy to formulate and can be justified based on first principles (Gillespie, 1992). The major drawback is that their analytical or computational analysis can be extremely difficult, especially when more than just a few different chemical species play a role for the reaction network. This is because the chemical master equation (CME), which governs the time evolution of the probability distribution of the CTMC, cannot be solved for anything but the simplest systems and even approximation techniques (Munsky and Khammash, 2006; Wolf et al., 2010) tend to fail when the CTMC is high-dimensional. In such cases, an alternative is to focus only on some low-order moments of the probability distribution. Ordinary differential equations that describe the time evolution of these moments can be derived from the CME (Engblom, 2006), but their solution typically requires some kind of approximation (Ruess et al., 2011; Singh and Hespanha, 2006). These approximations, known as moment closure, are usually based on an assumption about the underlying probability distribution and exist in many different varieties (Hespanha, 2008). Often, for a given system and given model parameters, some of these approximations provide good results whereas others fail

* Corresponding author.
E-mail address: jakob.ruess@ist.ac.at (J. Ruess).

http://dx.doi.org/10.1016/j.biosystems.2016.07.005
0303-2647/

to be sufficiently accurate or fail entirely. Unfortunately, there exists no approach for determining a priori which moment closure technique will provide the best approximation. In general, the only approach that is guaranteed to provide at least statistically exact results is to simply simulate the CTMC using a stochastic simulation algorithm (SSA) (Gillespie, 1976) and to compute Monte Carlo estimates of the system output of interest based on the simulation results. To obtain precise estimates, however, a large number of simulations may be required, leading to a high computational cost. For the forward analysis of a system, i.e. when the model parameters are known, this is not always a serious problem. For the reverse engineering task of identifying the model parameters from measured data, however, the CTMC needs to be analyzed for many different parameter values in order to determine those in best agreement with the measured data. Accordingly, for this task the computational cost of approaches based on stochastic simulation (Lillacci and Khammash, 2013) is often prohibitively large.

In this paper, we propose an approach for parameter inference based on moment closure that is complemented by stochastic simulation. In particular, the parameter inference is performed based on the computationally cheap moment closure approximation, whereas the stochastic simulation is employed whenever new regions in the parameter space are explored, either to ensure that the approximation is still sufficiently accurate, or to propose a new approximation that outperforms the previously used one. With this approach we are able to combine the computational advantages of moment closure with the statistical exactness of SSA and obtain a method that is both scalable and does not require a priori knowledge of the performance of different moment closure techniques. Importantly, the method is completely automated and chooses and adapts the approximation from a precomputed library of moment closure methods. Thus, the user only has to specify the model and supply the data and, contrary to previous approaches (Kügler, 2012; Ruess and Lygeros, 2015; Zechner et al., 2012), no expertise in the analysis of CTMCs is required.

The remaining paper is structured as follows. In Section 2, we introduce biochemical reaction networks, the chemical master equation and moment closure methods. In Section 3, we formulate a maximum-likelihood estimation problem for the model parameters and describe previously published moment-based methods for solving these problems. In Section 4, we propose our automated adaptive parameter inference method. In Section 5, we study the performance of our method for some benchmark reaction networks. Section 6 is devoted to a study of the computational cost of our approach. Finally, in Section 7, we discuss our results and provide some concluding remarks.

2. Stochastic modeling of biochemical reaction networks

Consider a biochemical reaction network consisting of m different chemical species $X^1, \ldots, X^m$ that interact according to K different reactions:

$$\nu'_{1k}X^1 + \cdots + \nu'_{mk}X^m \xrightarrow{\theta_k} \nu''_{1k}X^1 + \cdots + \nu''_{mk}X^m, \quad k = 1, \ldots, K, \tag{1}$$

where the coefficients ν'_{ik} and ν'_{ik} determine how many molecules of the i-th species are consumed and produced in the k-th reaction, respectively. Under the assumption that the reaction network is well-stirred and in thermal equilibrium, it can be described by a continuous-time Markov chain $X(t,\theta) = [X^1(t,\theta)\cdots X^m(t,\theta)]^T$ that takes states $x = [x^1 \cdots x^m]^T \in \mathbb{N}_0^m$ (Gillespie, 1992). The transition probabilities of this CTMC are determined by the reaction parameters $\theta = [\theta_1 \cdots \theta_K]^T \in (\mathbb{R}_0^+)^K$ and the kinetic rate law of the reactions. Here, we restrict our attention to mass action kinetics and elementary chemical reactions (i.e. reactions of order at most 2). These assumptions simplify the computation of moments of the CTMC. It should be noted, however, that they are not strictly necessary for the results of this paper and are mainly imposed because it is very unlikely that, in a three-dimensional space, more than two molecules meet at exactly the same time. Accordingly, any more complicated biochemical reaction can essentially be decomposed into a series of elementary reactions whose reaction rates are governed by the law of mass action. These assumptions lead to transition probabilities of the CTMC that are determined by propensity functions of the form $a_k(x,\theta) = \theta_k h_k(x)$, $k = 1, \ldots, K$, where $h_k(x)$ are at most quadratic polynomials in x. The time evolution of the probability distribution of $X(t,\theta)$ can then be described by the chemical master equation:

$$\dot{p}(x,t) = -p(x,t)\sum_{k=1}^{K} a_k(x,\theta) + \sum_{k=1}^{K} p(x-\nu_k,t)a_k(x-\nu_k,\theta), \tag{2}$$

where $\nu_k = [\nu_{1k}\cdots\nu_{mk}]^T$, $\nu_{ik} = \nu''_{ik} - \nu'_{ik}$, $i = 1, \ldots, m$, and $p(x,t)$ is the probability $P(X(t,\theta) = x)$ that x molecules of the m chemical species are present at time t.

Since $X(t,\theta)$ has a countably infinite state space, computing the probabilities $p(x,t)$ requires solving an infinite system of coupled ordinary differential equations, which is generally not possible. Approximate solutions can be obtained in some cases, for instance by projection to a finite state space (Munsky and Khammash, 2006; Wolf et al., 2010), but we will not discuss these approaches here.

An alternative is to focus only on some low-order moments of the probability distribution. Ordinary differential equations describing their time evolution can be derived from the CME (Engblom, 2006; Ale et al., 2013) and written as

$$\dot{\eta}(t) = A(\theta)\eta(t) + B(\theta)\bar{\eta}(t), \tag{3}$$

where $\eta(t)$ is a vector containing the (uncentered) moments up to some desired order L and $\bar{\eta}(t)$ contains moments of order $L+1$. Eq. (3) shows that the time evolution of $\eta(t)$ depends on moments of higher order; hence $\eta(t)$ cannot be computed without knowledge of $\bar{\eta}(t)$. Accordingly, the open system of equations Eq. (3) is typically replaced by an approximate closed system of equations

$$\dot{\tilde{\eta}}(t) = A(\theta)\tilde{\eta}(t) + B(\theta)f(\tilde{\eta}(t)), \tag{4}$$

where $\tilde{\eta}(t)$ are approximations of $\eta(t)$. The function f is usually chosen according to an assumption on the underlying probability distribution. Typical examples are to assume that the centered moments (or cumulants) of order $L+1$ are zero (Whittle, 1957; Matis et al., 2010), or to choose f according to a log-normal distribution (Singh and Hespanha, 2006). In general, the choice of f is made rather arbitrarily without actual knowledge of the underlying distribution. Furthermore, whether a given closure will provide good approximations depends on the system that is being studied, the model parameters, and the order L at which the moment equations are closed. This makes it practically impossible for someone who is not an expert in the use of these methods to choose an appropriate closure. Despite all this, moment closure methods have been successfully applied for analyzing CTMCs, and specifically also for parameter inference (Ruess and Lygeros, 2015; Zechner et al., 2012; Lück and Wolf, 2016). The choice of the closure method used in these references, however, was based on trial and error and the success of the performed studies accordingly required a portion of luck.

An alternative approach for analyzing biochemical reaction networks is by using a stochastic simulation algorithm (SSA). It is straightforward to generate statistically exact sample paths $x_1(t)$,

Please cite this article in press as: Schilling, C., et al., Adaptive moment closure for parameter inference of biochemical reaction networks. BioSystems (2016), http://dx.doi.org/10.1016/j.biosystems.2016.07.005

$\ldots, x_n(t)$ of $X(t,\theta)$ in this way. From these sample paths, estimators of any system output, for instance some moments or the entire probability distribution at a certain time point, can be constructed. While such an approach is easy to implement and can always be used, it comes with the major drawback that often a large number of sample paths n is required to obtain precise estimates. This can make the use of stochastic simulation for reverse engineering tasks computationally prohibitively expensive.

3. Moment-based parameter inference

In this section, we formulate the parameter inference problem and review previous methods that have been developed to solve it. The goal in this paper is to estimate the reaction rate constants θ from measured data that is of the form $y = \{x_1^j(t_s), \ldots, x_n^j(t_s) \mid j \in J, \quad s = 1, \ldots, S\}$ and corresponds to measuring the number of molecules of each chemical species from index set J in n cells at each measurement time point t_s, $s = 1, \ldots, S$. We assume that all measurements collected at different times and in different cells are statistically independent. This is, for instance, the case for flow cytometry data where the cells are discarded after being measured so that two different measurements can never come from the same cell. When measuring more than one species in the same experiment, i.e. $|J| > 1$, the measurements of the different species that are collected at the same time and in the same cell will, however, be correlated. To simplify the presentation, we will first describe the case $J = \{j\}$ where only a single species X^j is measured.

The task of identifying the model parameters from the data can be posed as a maximum-likelihood estimation problem

$$\theta_{\mathrm{MLE}}(y) = \underset{\theta}{\mathrm{arg\,max}}\, \mathcal{L}(y,\theta), \tag{5}$$

where y is the measured data and $\mathcal{L}(y,\theta) = p(y|\theta)$ is the likelihood of the parameters θ, i.e. the probability (density) of the data given that θ are the model parameters. Analytically computing the likelihood is usually impossible, and accordingly, the optimization problem in Eq. (5) is typically solved by iterative numerical evaluation of $\mathcal{L}(y,\theta)$ for many different values of θ. Unfortunately, evaluating the likelihood for given parameters θ requires solving the CME with these parameters, which, as discussed in the previous section, is often impossible or computationally expensive itself. For this reason, one option is to use sample moments of the data as measurements instead of the entire data (Zechner et al., 2012). For instance, one can compute sample means $\hat{\mu}_1(t_s)$ and sample variances $\hat{\mu}_2(t_s)$, $s = 1, \ldots, S$ from the data y and treat the vector $\hat{\mu} := \left[\hat{\mu}(t_1) \cdots \hat{\mu}(t_S)\right]^T$, where $\hat{\mu}(t_s) := \left[\hat{\mu}_1(t_s) \;\; \hat{\mu}_2(t_s)\right]$, as new data. In earlier publications (Zechner et al., 2012; Ruess et al., 2013), we have shown that the probability density function $p(\hat{\mu}|\theta)$ of $\hat{\mu}$ is given by

$$p(\hat{\mu}|\theta) = \prod_{s=1}^{S} p(\hat{\mu}(t_s)|\theta),$$
$$\text{where} \quad p(\hat{\mu}(t_s)|\theta) = \mathcal{N}(M(t_s), \Sigma(t_s)) \tag{6}$$

and

$$M(t_s) = \begin{bmatrix} \mu_1(t_s) \\ \mu_2(t_s) \end{bmatrix}$$

and

$$\Sigma(t_s) = \frac{1}{n} \begin{bmatrix} \mu_2(t_s) & \mu_3(t_s) \\ \mu_3(t_s) & \mu_4(t_s) - \dfrac{n-3}{n-1}(\mu_2(t_s))^2 \end{bmatrix},$$

where $\mathcal{N}$ stands for the normal distribution, $\mu_1(t_s) = \mu_1(t_s,\theta)$ is the mean and $\mu_i(t_s) = \mu_i(t_s,\theta)$, $i = 2, 3, 4$ are the centered moments of the measured species $X^j(t_s,\theta)$ at time t_s for model parameters θ. Since these moments can be computed from the solution of Eq. (4), we can use this result to approximately compute the likelihood $\mathcal{L}(\hat{\mu},\theta) = p(\hat{\mu}|\theta)$ without having to solve the CME. Accordingly, we can in principle solve the optimization problem in Eq. (5) using $\hat{\mu}$ instead of y to compute the maximum-likelihood estimator $\theta_{\mathrm{MLE}}(\hat{\mu})$.

However, there are two practical problems in evaluating the covariance matrix $\Sigma(t_s) = \Sigma(t_s,\theta)$. The first is that moment closure typically does not preserve the properties of the underlying stochastic process. This means that the relations between the moments $\mu_i(t_s,\theta)$, $i = 1, \ldots, 4$ that ensure that $\Sigma(t_s,\theta)$ is positive definite and a proper covariance matrix might be lost due to the approximation. The second problem is that $\Sigma(t_s,\theta)$ depends on the first four moments even if only the first two sample moments are used as data. Consequently, moment closure methods of order less than four cannot be used to evaluate the likelihood. To circumvent these problems, we implemented a slightly different likelihood in our algorithm. More specifically, we estimate the covariance matrices $\Sigma(t_s)$ from the data by computing empirical estimates of the moments up to order four and plugging them into the above equation. This means that the covariance matrix is fixed throughout the optimization problem and chosen as an estimate of the covariance matrix $\Sigma(t_s,\theta)$ in Eq. (6) evaluated at the true unknown parameters, i.e. as the covariance matrix to which the parameter dependent covariance matrix $\Sigma(t_s,\theta)$ would converge when the parameter search converges to the true parameters. Throughout this paper, we will follow this strategy and denote by μ_{data} the moments up to fourth order of the data, i.e. $\mu_{\mathrm{data}} := [\mu_{\mathrm{data}}(t_1) \cdots \mu_{\mathrm{data}}(t_S)]^T$, where $\mu_{\mathrm{data}}(t_s) := \left[\hat{\mu}_1(t_s) \;\; \hat{\mu}_2(t_s) \;\; \hat{\mu}_3(t_s) \;\; \hat{\mu}_4(t_s)\right]$ contains the first four centered empirical moments of the data set at time t_s. This strategy is appropriate whenever sufficiently many cells are measured so that the moments up to order four can be estimated with reasonable precision. For flow cytometry data, the number of cells measured per time point typically ranges in the order of thousands or even tens of thousands; hence sufficing precision is always guaranteed.

Extension to multi-dimensional measurements. The likelihood in Eq. (6) is only valid for a single measured species. It can, however, also be extended to more general settings where multiple chemical species are measured simultaneously in the same experiment (e.g. by using different fluorescent markers for labeling different proteins in flow cytometry experiments). In the following, we consider a case where two species are measured, i.e. $J = \{i, j\}$ so that X^i and X^j are the measured species. The data is then of the form $y = \left\{ \{x_1^i(t_s), x_1^j(t_s)\}, \ldots, \{x_n^i(t_s), x_n^j(t_s)\}, \quad s = 1, \ldots, S \right\}$ and the vector of sample moments becomes $\hat{\mu} := \left[\hat{\mu}(t_1) \cdots \hat{\mu}(t_S)\right]^T$, where at each time point t_s the sample means and variances of both species and the sample covariance between the species are measured, i.e. $\hat{\mu}(t_s) := \left[\hat{\mu}_1^i(t_s) \;\; \hat{\mu}_1^j(t_s) \;\; \hat{\mu}_2^i(t_s) \;\; \hat{\mu}_2^{ij}(t_s) \;\; \hat{\mu}_2^j(t_s)\right]$.

It has previously been shown (Ruess et al., 2013) that in this case the likelihood of the sample moments is given by

$$p(\hat{\mu}|\theta) = \prod_{s=1}^{S} p(\hat{\mu}(t_s)|\theta), \quad \text{where} \quad p(\hat{\mu}(t_s)|\theta) = \mathcal{N}(M(t_s), \Sigma(t_s)) \tag{7}$$

$$M(t_s) = \left[\mu_1^i(t_s) \;\; \mu_1^j(t_s) \;\; \mu_2^i(t_s) \;\; \mu_2^{ij}(t_s) \;\; \mu_2^j(t_s)\right]^T$$

and

$$\Sigma(t_s) = \frac{1}{n}\begin{bmatrix} \mu_2^i & \mu_2^{ij} & \mu_3^i & \mu_3^{i2j} & \mu_3^{ij2} \\ \cdot & \mu_2^j & \mu_3^{i2j} & \mu_3^{ij2} & \mu_3^j \\ \cdot & \cdot & \mu_4^i - \frac{n-3}{n-1}(\mu_2^i)^2 & \mu_4^{i3j} - \frac{n-3}{n-1}\mu_2^i\mu_2^{ij} & \mu_4^{i2j2} - \frac{n-3}{n-1}\mu_2^i\mu_2^j \\ \cdot & \cdot & \cdot & \mu_4^{i2j2} - \frac{n-2}{n-1}(\mu_2^{ij})^2 + \frac{1}{n-1}\mu_2^i\mu_2^j & \mu_4^{ij3} - \frac{n-3}{n-1}\mu_2^{ij}\mu_2^j \\ \cdot & \cdot & \cdot & \cdot & \mu_4^j - \frac{n-3}{n-1}(\mu_2^j)^2 \end{bmatrix},$$

where for simplicity we omitted the lower left entries of the symmetric matrix $\Sigma(t_s)$ and the dependence of the moments in $\Sigma(t_s)$ on the time point t_s. The entries of $\Sigma(t_s)$ are functions of the centered moments up to order four of the joint distribution of $X^i(t_s)$ and $X^j(t_s)$. The notation has to be understood as follows: lower case indices at μ determine the order of the moment; upper case indices determine the identity of the moment, i.e. μ_3^{i2j} refers to the third order moment $\mathbb{E}[(X^i(t_s) - \mathbb{E}[X^i(t_s)])^2 \cdot (X^j(t_s) - \mathbb{E}[X^j(t_s)])]$ and μ_4^{i2j2} refers to the fourth order moment $\mathbb{E}[(X^i(t_s) - \mathbb{E}[X^i(t_s)])^2 \cdot (X^j(t_s) - \mathbb{E}[X^j(t_s)])^2]$. In our algorithm, all the entries of $\Sigma(t_s)$ are estimated from the measurements and collected in the vector μ_{data}, as already described above for the case of a single measured species.

4. Adaptive approach for parameter inference

The drawback of the approach described in the previous section is that a moment closure method has to be chosen in advance and this closure will be used throughout the entire parameter search. This leads to the problems that, on the one hand, it is a priori very difficult to choose the best closure and, on the other hand, which closure is best may also be different for different parts of the parameter space. The main idea of the method that we propose in the following is to use a small number of simulated trajectories of the system that are generated using a stochastic simulation algorithm (SSA) in order to test different approximations during the parameter space exploration. Specifically, whenever the parameter search leaves a certain area in parameter space, defined as an ϵ-neighborhood around the point at which the last SSA run was carried out, new simulations are performed and all closure methods from a predefined library are evaluated by comparing the different approximations at the current point in parameter space to the simulation results. Importantly, all the approximate moment systems, corresponding to closures of different types and degrees, are precomputed only once, and thus new derivations of the moment equations are not required during the search. To generate these systems we make use of Hespanha's StochDynTools toolbox (Hespanha, 2007).

A flowchart describing our approach is given in Figure 1. We now explain the approach step by step using the pseudocode given in Algorithm 1. The inputs of the algorithm are the CTMC model $X(t,\theta)$, parametrized by the reaction rate constants θ, a set of ODE systems $CL = \{c_1(\theta), \ldots, c_q(\theta)\}$ corresponding to different approximations of the moment dynamics obtained through various closures of different types and degrees, the centered moments up to the fourth order μ_{data} of a measured data set Y, and a maximal number of iterations $i_{\max}$ that determines for how many steps in parameter space the search is performed. The algorithm returns the maximum likelihood estimator θ_{MLE}. The core idea of our approach works independently of the actual parameter search technique used in the background. Thus, it can be applied in conjunction with any standard optimization scheme used to minimize some distance between model output and data (for instance simple gradient descent). Accordingly, we focus on the adaptive update of the closure method while abstracting from the actual details of the parameter search for a fixed approximation by the function NextParameter (line 18). It takes the current values of the parameters θ_i and the chosen approximate ODE system $c_{\text{best}}(\theta_i)$ and moves the search to the new parameters θ_{i+1} according to some criteria. In our implementation, we instantiate it with a Markov chain Monte Carlo method and a Metropolis-Hastings sampler, based on the likelihood in Eq. (6) (Zechner et al., 2012). Additionally, this function also takes care of updating the value of the maximum likelihood estimator θ_{MLE} based on the likelihood of the new parameters θ_{i+1}. The remaining pseudocode describes how and when the used closure method is adjusted. We first check whether the current parameter values θ_i are still within the ϵ-neighborhood $N_\epsilon(\theta_{\text{ref}})$, where θ_{ref} are the parameters at which the previous

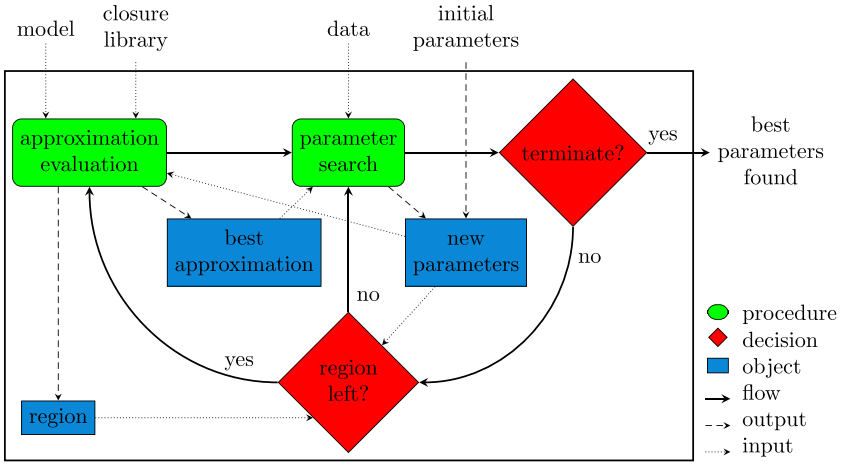

Fig. 1. Flowchart describing the high-level workflow of our method.

simulation was performed (line 5). In our implementation, we choose a neighborhood in the form of a hyperrectangle of relative size $N_\epsilon(\theta_{\text{ref}}) = \{\theta \mid |\theta - \theta_{\text{ref}}|_k \leq \epsilon \cdot |\theta_{\text{ref}}|_k, k = 1, \ldots, K\}$. If $\theta_i \in N_\epsilon\left(\theta_{\text{ref}}\right)$, we directly proceed with the standard inference method in line 18, relying on the ODE system $c_{\text{best}}(\theta_i)$ from the most recent evaluation. Otherwise, stochastic simulation is employed with the current parameter values θ_i to compute estimates of the moments $\mu_{\text{SSA}}(\theta_i)$ using the function COMPUTE$_{\text{SSA}}$ (line 6), for which we utilize a standard implementation of Gillespie's SSA in our implementation. These estimates are then compared to the approximations $\mu_{\text{ODE}}(\theta_i)$ obtained with all the different closure methods using the function COMPUTE$_{\text{ODE}}$ which numerically computes the solution of the system of ODEs $c(\theta_i) \in CL$ (lines 8–15). The best approximate system $c_{\text{best}}(\theta_i)$ is chosen as the one that minimizes some distance DIST between estimation and approximation. In general, this distance could be determined in many different ways. In our implementation, we choose DIST as the likelihood of the estimated moments for the measured species (Eq. (6)), i.e. we measure the performance of the approximations by evaluating how precise the approximated moments of the system output (not of the entire state) are. Finally, we update the reference point θ_{ref} to θ_i (line 16) and the search continues in the standard way until the next ϵ-neighborhood is left.

Algorithm 1. Adaptive moment-based parameter inference algorithm

Input: CTMC $X(t, \theta)$, where $\theta \in (\mathbb{R}_0^+)^K$, data μ_{data}, maximum number of iterations $i_{\max}$, and set of approximate moment systems $CL = \{c_1(\theta), \ldots, c_q(\theta)\}$ obtained using different closure methods
Output: Maximum likelihood estimator θ_{MLE}

1: θ_1 := random initial parameter values
2: $\theta_{\text{MLE}} := \theta_1$
3: $\theta_{\text{ref}} := +\infty$
4: **for** $i := 1$ **to** $i_{\max}$ **do**
5: **if** $\theta_i \notin N_\epsilon\left(\theta_{\text{ref}}\right)$ **then**
6: $\mu_{\text{SSA}}(\theta_i) := \text{COMPUTE}_{\text{SSA}}(X(t, \theta_i))$
7: $d_{\text{best}} := +\infty$
8: **for all** $c(\theta_i) \in CL$ **do**
9: $\mu_{\text{ODE}}(\theta_i) := \text{COMPUTE}_{\text{ODE}}(c(\theta_i))$
10: $d := \text{DIST}(\mu_{\text{SSA}}(\theta_i), \mu_{\text{ODE}}(\theta_i))$
11: **if** $d < d_{\text{best}}$ **then**
12: $d_{\text{best}} := d$
13: $c_{\text{best}}(\theta_i) := c(\theta_i)$
14: **end if**
15: **end for**
16: $\theta_{\text{ref}} := \theta_i$
17: **end if**
18: $\langle \theta_{i+1}, \theta_{\text{MLE}} \rangle := \text{NEXTPARAMETER}(\theta_i, c_{\text{best}}(\theta_i), \mu_{\text{data}}, \theta_{\text{MLE}})$
19: **end for**
20: **return** θ_{MLE}

5. Case studies

We applied our inference method to several benchmark stochastic reaction networks. In this section, we report some exemplary results. For all examples, to generate the set of approximate ODE systems CL we used *derivative matching* (dm), *zero cumulants* (zc), *zero variance* (zv) moment closure, each with degree 2, 3, and 4, and *low dispersion* (ld) moment closure with degree 3 and 4 (we refer the reader to the literature (Hespanha, 2008) for details).

Example 1. The first network is a model that has recently been used to describe agricultural pests (Gillespie and Golightly, 2010; Parise et al., 2015) but can also be regarded as a model of gene expression given by the following reactions:

$$\emptyset \xrightarrow{\theta_1} N + C \qquad N \xrightarrow{\theta_2} 2N + C$$

$$N + C \xrightarrow{\theta_3} C \qquad C \xrightarrow{\theta_4} \emptyset.$$

The produced protein is positively regulated by the current amount of protein and negatively regulated (through an increased degradation rate) by past amounts of protein (i.e. species C could be regarded as an abstraction of a slow process that is activated by N and leads to the production of proteases that degrade N). We assume that $N(0) = C(0) = 0$, that the true parameters are given by $\theta_1 = 0.03$, $\theta_2 = 0.012$, $\theta_3 = 0.25 \times 10^{-4}$ and $\theta_4 = 0.003$, and that 5000 cells are measured at the time points $t_1 = 100, \ldots, t_9 = 900$. As settings for our algorithm we use $\varepsilon = 0.2$ and perform 200 simulations whenever the search leaves an ϵ-neighborhood, i.e. in line 6 of Algorithm 1.

An exemplary run of our parameter search for $i_{\max} = 2000$ iterations, started from random initial parameter values, is shown in Figure 2. It can be seen that all the inferred parameters, i.e. the maximum-likelihood estimates $\theta_{\text{MLE}}(\hat{\mu})$, agree with the true parameter values up to negligible errors with basically no uncertainty. The former is a sign that a precise moment closure method exists for this example, whereas the latter stems from the large number of measurements that we assumed to be available. Fig. 3 shows that also the model predictions, computed with the inferred parameters $\theta_{\text{MLE}}(\hat{\mu})$ and the best closure method, agree well both with the data and with SSA estimates of mean and variance obtained with the inferred parameters. We can conclude that the moment closure approximation is very precise and can match the data up to very small errors.

To evaluate on the one hand how important it is to choose a good approximation, and on the other hand whether it is necessary to adaptively change the closure method during the search, we perform the parameter inference with the same data and the same algorithm, but fix an initial closure method and do not allow the search to switch between different approximations (i.e. by choosing $\varepsilon = +\infty$). Table 1(a) compares the error in the inferred parameters obtained from our approach to the error in the results when the closure is fixed. It can be seen that for some of the fixed closure approaches the error in the parameter estimates is very large (specifically for all of the zero variance closures). Other methods provide more precise results. It is interesting to note that the fourth order zero cumulants (zc4) closure is the computationally most expensive one and the parameter search with fixed zc4 closure actually takes three minutes longer than the adaptive search, despite the additional stochastic simulations and evaluations of all closure methods needed here.

To further test our results, we investigate how often the approximation is changed during the run of our algorithm and which closure methods are used most often. Table 1(b), column Ex 1, shows how often the different closure methods are chosen as best. It can be seen that some approximations are never chosen (for instance all of the zero variance closures but also the second and third order zero cumulants closures) whereas derivative matching and low dispersion closures are chosen most often. Overall, high order closures are preferred over low order closures. This is to be expected, since these usually provide more precise results at the cost of an increased computational effort. Also we highlight that the option to switch the approximation is often used (in 15 out of 26 evaluations), and, compared to a pure simulation-based approach, we need to employ stochastic simulation only 26 times (instead of 2000 times).

Example 2. As second case study, we choose a model of two genes that regulate each other's production such that a negative feedback loop is formed (see Fig. 4). The model is given by the following reaction network:

$$G_1^{\text{off}} \xrightarrow{\theta_1} G_1^{\text{on}} \qquad G_2^{\text{off}} + P_1 \xrightarrow{\theta_4} G_2^{\text{on}} + P_1$$

$$G_1^{\text{on}} + P_2 \xrightarrow{\theta_2} G_1^{\text{off}} + P_2 \qquad G_2^{\text{on}} \xrightarrow{\theta_5} G_2^{\text{off}}$$

$$G_1^{\text{on}} \xrightarrow{\theta_3} G_1^{\text{on}} + P_1 \qquad G_2^{\text{on}} \xrightarrow{\theta_6} G_2^{\text{on}} + P_2$$

$$P_1 \xrightarrow{\theta_7} \emptyset \qquad P_2 \xrightarrow{\theta_7} \emptyset$$

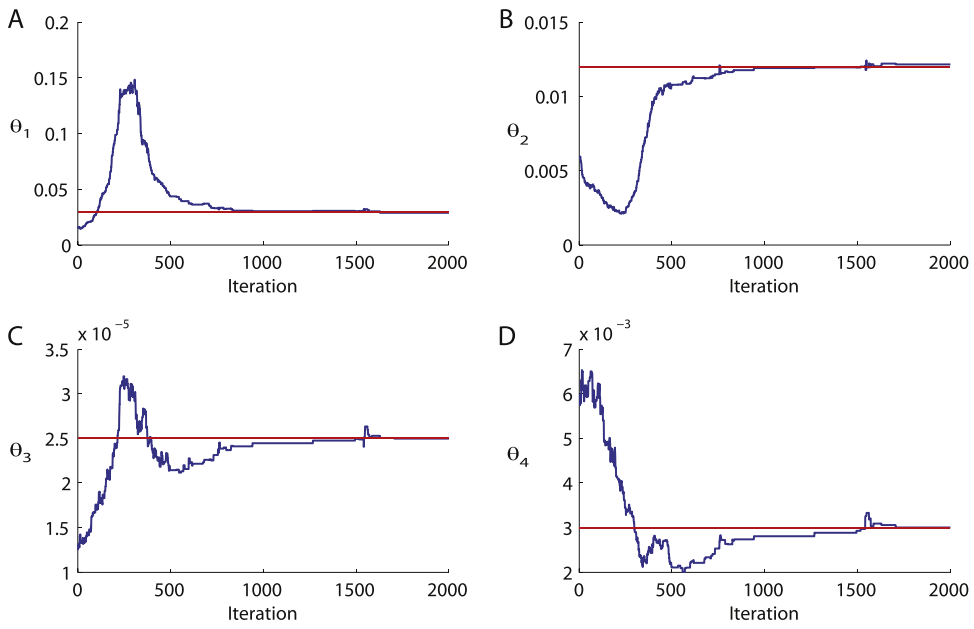

Fig. 2. Parameter search for Example 1. The panels show the values of the parameters in the search as a function of the iteration (blue). It can be seen that after approximately 1500 iterations the search is very close to the true values (red lines) for all parameters and retains these values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

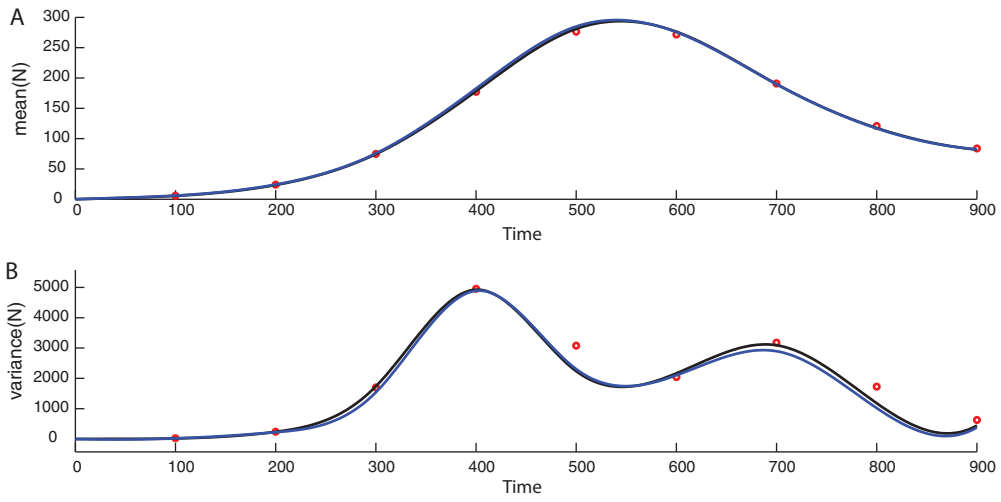

Fig. 3. Model output and data for the inferred parameters of Example 1. (A) The mean computed with the best closure method (black) and the inferred parameters agrees very well both with the data (red) and the results of stochastic simulation with the inferred parameters (blue). (B) Also all the variances agree very well. The color coding is the same as in (A). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

If a gene is active (G_1^{on} and G_2^{on}), the respective protein is produced. Otherwise, the gene is in an inactive state (G_1^{off} and G_2^{off}). We assume that initially no protein molecules are present (i.e. $P_1(0) = P_2(0) = 0$) and that both genes are in the inactive state. Thus,

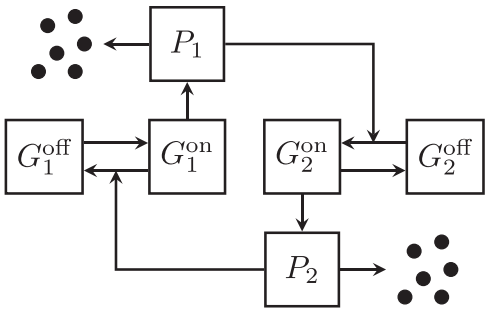

Fig. 4. Network model of the two genes negative feedback loop.

only the first reaction can take place (with rate θ_1), leading to activation of gene G_1 and enabling production of protein P_1 (with rate θ_3). Gene G_2 can only be activated if the protein P_1 is present (rate θ_4), and similarly, gene G_1 can only be deactivated if protein P_2 is present (rate θ_2). Additionally, both proteins degrade with rate θ_7. It has previously been observed that for this example measuring only one of the produced proteins does not provide sufficient information to reliably estimate all the model parameters (Ruess and Lygeros, 2015). To test whether this is confirmed by our method, we applied our algorithm to data containing measurements of only P_1 or only P_2. For both cases the search did not converge to the true parameters and showed that many different parameters lead to similar agreement of the model output with the data, i.e. that the model parameters are practically unidentifiable if only one protein is measured. It follows that both P_1 and P_2 have to be measured and the extended likelihood in Eq. (7) has to be used in our algorithm.

We assume that the true parameters are given by $\theta_1 = 0.1$, $\theta_2 = 0.01$, $\theta_3 = 10$, $\theta_4 = 0.005$, $\theta_5 = 0.1$, $\theta_6 = 10$ and $\theta_7 = 0.1$. We

Table 1
(a) Relative distance between true and inferred parameters. The distances (in percent) are obtained from our adaptive algorithm (adapt) and the different closure methods on their own for the reaction network from Example 1. The smallest distance is marked in bold. (b) Statistics of the used closure methods. Columns correspond to the different reaction networks (Ex stands for example). The top block of rows reports in percent how often each of the closure methods was chosen as best in our adaptive search. The bottom block shows how often the approximation was changed as our search progressed through the parameter space (switch), how often stochastic simulation was performed, i.e. how often ε-neighborhoods were left and all the closure methods were tested (SSA), and the total number of iterations in the search ($i_{\max}$).

Closure	θ_1	θ_2	θ_3	θ_4
(a) Example 1				
Adapt	4.09	1.20	0.08	**0.10**
dm2	**0.13**	0.23	0.52	1.02
zc2	10.37	6.20	6.28	10.73
zv2	379.28	85.45	56.99	92.45
dm3	2.25	**0.03**	0.90	2.72
zc3	0.94	0.47	7.29	13.04
zv3	365.21	91.94	62.47	96.75
ld3	28.80	5.07	0.60	1.55
dm4	2.79	0.55	**0.07**	0.34
zc4	17.52	1.77	6.64	9.76
zv4	364.76	86.07	55.81	88.76
ld4	44.69	10.67	10.97	23.71

Closure	Ex 1	Ex 2	Ex 3	Ex 4	Ex 5
(b) Search statistics					
dm2	13	0	39	0	6
zc2	0	0	0	46	0
zv2	0	0	0	0	0
dm3	13	11	0	0	3
zc3	0	0	17	0	3
zv3	0	0	17	0	0
ld3	18	11	0	0	35
dm4	18	45	5	54	11
zc4	13	0	0	0	21
zv4	0	0	11	0	0
ld4	25	33	11	0	21
Switch	15	8	17	12	33
SSA	26	53	23	48	47
$i_{\max}$	2k	10k	2k	4k	5k

initialize the search with arbitrary parameter values and set the maximal number of iterations i_{max} to 10,000. For the reference data we assume that 5000 cells are measured at the time points $t_1 = 10$, ..., $t_{10} = 100$. We set the ε-parameter to 0.2 and the number of simulations for evaluating the closure methods to 100. The search (see Fig. 5) needs roughly 3000–4000 iterations to converge to a region close to the true parameters. This period can be seen as the time that the used Markov chain Monte Carlo algorithm needs to converge to its stationary distribution. The parameter values at the remaining iterations can then be seen as samples from a posterior distribution (with flat priors), i.e. a histogram over these values shows the shape of the likelihood in parameter space. One-dimensional histograms for each of the parameters are depicted in Fig. 6. It can be seen that the histograms are very narrow, i.e. there is only little uncertainty in the inferred parameter values, similar to Example 1. An ideal result would be if all the histograms were narrow and centered around the true parameter values (shown by the red line). However, errors that are introduced by the moment closure approximations are propagated to the inferred parameters, which leads to the result that not all the histograms in Fig. 6 are perfectly centered around the true values. Fig. 7 shows the sample moments of the measured data and the final model output, i.e. the moments of the joint distribution of P_1 and P_2 computed with the best closure method and the inferred parameters $\theta_{\text{MLE}}(\hat{\mu})$.

As in Example 1, we also test how well the parameters can be inferred when the closure method is not adapted during the search. We find that only four of the considered moment closure methods can be used on their own, as can be seen in Table 2(a). The closure methods ld4 (45%), dm4 (33%), dm3 (11%) and ld3 (6%) are used exclusively in the adaptive search. It is interesting to note that, while method ld4 is very precise around the true parameters, it cannot be used in the other regions visited during the search. This is a situation where the capability of our adaptive algorithm to change the approximation during the search is important to obtain good results. Our algorithm uses the dm3 and dm4 methods in the other regions to guide the search toward the true parameters, and then switches to the ld4 method only when it becomes precise enough.

Further examples. In addition to Examples 1 and 2 we apply our algorithm to three further reaction networks and perform the same comparisons. Specifically, we consider two models from the literature, namely the model of transient gene expression reported in reference (Zechner et al., 2012) (termed here Example 3) and the first case study in reference (Ruess et al., 2011) (termed here Example 4). Additionally, we design a model of a negatively autoregulated gene where the produced protein P induces the degradation of the transcription factor T that is responsible for activating the gene (see Rosenfeld et al., 2002; Becskei and Serrano, 2000 for examples of such a system) (termed here Example 5):

$$\emptyset \xrightarrow{\theta_1} T \qquad T + G^{\text{off}} \xrightarrow{\theta_2} G^{\text{on}} \qquad G^{\text{on}} \xrightarrow{\theta_3} T + G^{\text{off}}$$
$$G^{\text{on}} \xrightarrow{\theta_4} P + G^{\text{on}} \qquad T + P \xrightarrow{\theta_5} P \qquad P \xrightarrow{\theta_6} \emptyset.$$

The results are overall similar to those obtained for Examples 1 and 2 and we only report the results on the identified parameters in Tables 2(b) and 2(c), respectively; additionally, Table 1(b), columns Ex 3–Ex 5, shows how often the different closure methods are used by our adaptive search. It can be seen that in Example 4 the second order zero cumulants and the fourth order derivative matching closure are chosen exclusively, whereas in Example 3, different closures, including the zero variance closures, are used and there is no noticeable preference for higher order closures. In Example 5, mainly the third and fourth order low dispersion closures are used. Overall, we can observe that for different examples different closure methods are employed and it is never the case that a single closure method alone performs best during a whole parameter identification run, which demonstrates the relevance of our adaptive algorithm.

6. Evaluation of the computational cost

It is clear that the difficulty of the parameter inference problem increases with the dimensionality of the parameter space that needs to be searched. In our implementation, this means that for high dimensional parameter spaces, a larger number of iterations i_{max} of the search should be used. How many iterations are necessary depends on the shape of the likelihood and the chosen starting point of the search and cannot be answered in general. To compare different examples, we therefore focus on the computational cost normalized by the number of iterations throughout this section.

The second important question is how the computational cost scales with the size of the reaction network, i.e. with the number of chemical species m that are part of the reaction network. The number of states of the whole CTMC model grows exponentially in m even after truncation of a possibly infinite state space and the chemical master equation typically becomes intractable already for small values of m. The main benefit of moment-based approaches is that the number of moment equations scales only polynomially in m with a degree that grows in the maximal order of moments that are used. The precise number of equations that is needed has recently been investigated (Ruess, 2015), in particular also for reaction networks like our Example 2 where some species

Please cite this article in press as: Schilling, C., et al., Adaptive moment closure for parameter inference of biochemical reaction networks. BioSystems (2016), http://dx.doi.org/10.1016/j.biosystems.2016.07.005

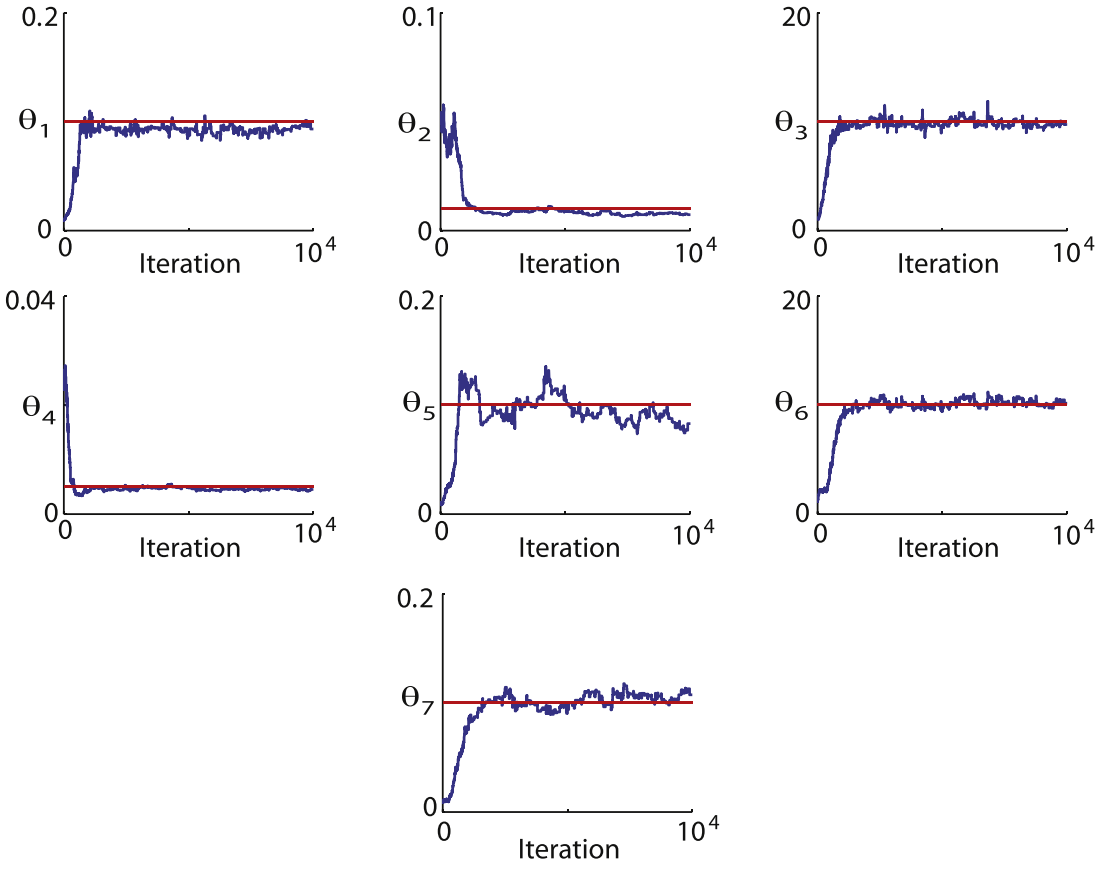

Fig. 5. Parameter search for Example 2. The panels show the values of the parameters in the search as a function of the iteration (blue). It can be seen that after approximately 4000 iterations the search is very close to the true values (red lines) for all parameters and remains close to these values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(e.g. those representing gene states) can only be present in finite numbers of molecules. Here, we focus on numerical results for our case studies and report the run times of our algorithm, as well as the run times of parameter searches where the closure method is not adaptively changed, in Table 3. As would be expected, the adaptive search is usually the slowest because it employs additional stochastic simulations and evaluates the precision of all moment closure methods whenever an ε-neighborhood is left, i.e. several times during the search.

Appropriate parallelization and/or use of a larger value for ε could mitigate this overhead, however. As a conclusion, our approach scales with the run time of the most precise closure method, with additional cost per evaluation of the approximation, and we can control the evaluation frequency by the value of ε. The run times per iteration in Table 3 also show the expected tendency that parameters of small reaction networks can be identified faster. In particular, Example 1 and Example 4 consist only of two species each and require less computation time per iteration of the search than Example 2 and Example 3. The computational cost observed for Example 5, on the other hand, is similar to that of the smaller systems in Example 1 and Example 4, which shows that the computational cost of parameter inference does not necessarily increase with the size of the reaction network.

7. Discussion

Using mathematical models to help in the understanding of complex biological systems is the core idea of systems biology. Up to some years ago, the main bottleneck in the identification of models was the availability of sufficiently precise and abundant data. Recently, measurement technologies have been improving at an amazing pace and nowadays enable us to simultaneously observe the dynamics of many different chemical species at single cell resolution. As these developments continue, we will gain access to data that is sufficiently informative to allow us to infer mathematical models of complex reaction networks from the measurements. However, for stochastic kinetic models that capture the inherent randomness of chemical reactions, this leads to a new bottleneck: the chemical master equation becomes intractable for high-dimensional models and especially the reverse engineering task of identifying model parameters from the measured data quickly becomes computationally infeasible. Parameter inference methods based on moment closure offer a solution to this problem but come with their own drawbacks. The goal of this paper was to address these drawbacks and to provide an automated moment-based inference method that can be used without in-depth knowledge of moment closure.

To this end, we interfaced previously proposed approaches with a stochastic simulation algorithm by continuously checking the quality of the approximations and adaptively adjusting the used closure method to the best one available. Accordingly, our approach is generally applicable whenever a sufficiently accurate approximation in the generated library of moment closure methods exists. Importantly, since the approach can adapt the used closure during the exploration of the parameter space, it is not required that a unique closure method provides good approximations for the entire parameter space.

Naturally, these benefits come with an increased computational cost compared to most standard moment-based inference approaches. This increase can primarily be attributed to the

Please cite this article in press as: Schilling, C., et al., Adaptive moment closure for parameter inference of biochemical reaction networks. BioSystems (2016), http://dx.doi.org/10.1016/j.biosystems.2016.07.005

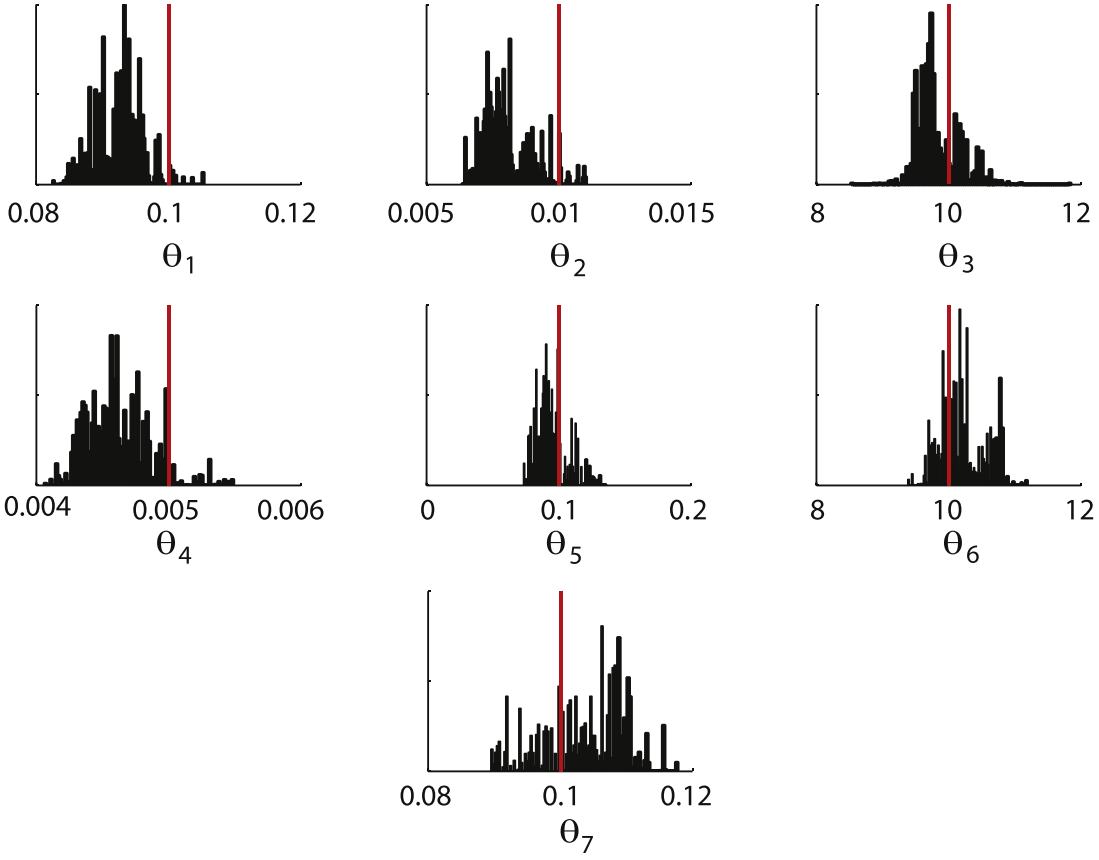

Fig. 6. Histograms for Example 2. The histograms were taken over all parameter values visited by the search in the iterations 4000 to 10,000 (black). The true values of the parameters are given as red lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

additional stochastic simulation and the evaluation of all the closure methods that is performed whenever the parameter search leaves an ε-neighborhood around the point in parameter space where the last simulation was performed. Accordingly, the parameter ε provides a trade-off between computational cost and guarantees that a good approximation is used. For $\varepsilon \to \infty$ our approach becomes a standard moment-based inference method, whereas $\varepsilon \to 0$ essentially leads to a method akin to those based entirely on stochastic simulation. We believe that this flexibility will prove to be valuable and allow us to investigate a large variety of different reaction networks with one unified inference method.

As future work, we plan to extend the generality of our algorithm further. More specifically, we plan to include and test more moment closure methods (e.g. the linear noise approximation) and possibly also to include approximation methods for the full solution of the CME as alternatives to moment closure (Munsky and Khammash, 2006; Wolf et al., 2010; Ammar et al., 2012; Chinesta et al., 2015). Furthermore, one could also envision

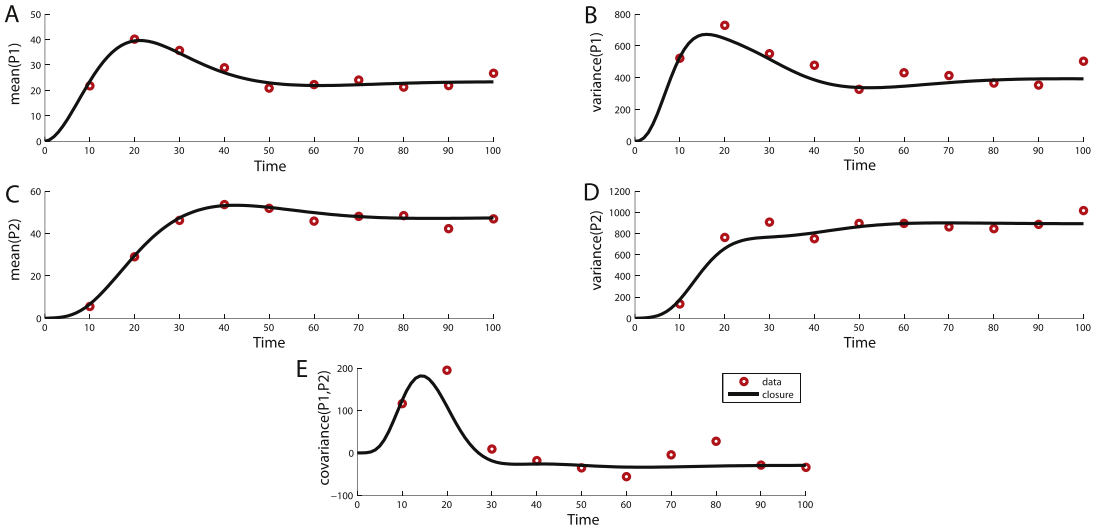

Fig. 7. Model output and data for the inferred parameters of Example 2. The moments of the joint distribution of P_1 and P_2 computed with the best closure method and the inferred parameters (black) are compared to the sample moments of the measured data (red dots). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: Schilling, C., et al., Adaptive moment closure for parameter inference of biochemical reaction networks. BioSystems (2016), http://dx.doi.org/10.1016/j.biosystems.2016.07.005

Table 2
Relative distance between true and inferred parameters. The tables show the results as in Table 1(a) for the respective Examples 2–5. Dashes (–) indicate that the particular parameter search did not succeed due to problems with the closure method.

Closure	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7
(a) Example 2							
Adapt	5.54	15.92	**0.24**	**1.43**	8.02	2.64	6.54
dm2	**1.09**	35.65	2.44	33.62	40.09	5.61	18.58
zc2	–	–	–	–	–	–	–
zv2	–	–	–	–	–	–	–
dm3	6.11	17.01	4.92	10.42	2.40	0.45	**1.43**
zc3	–	–	–	–	–	–	–
zv3	–	–	–	–	–	–	–
ld3	9.80	32.23	3.41	14.73	19.12	8.37	15.11
dm4	4.56	**9.33**	3.26	5.91	**4.19**	**0.05**	2.52
zc4	–	–	–	–	–	–	–
zv4	–	–	–	–	–	–	–
ld4	–	–	–	–	–	–	–

	Example 3				Example 4			
Closure	θ_1	θ_2	θ_3	θ_4	θ_1	θ_2	θ_3	θ_4
(b) Examples 3 &4								
Adapt	**2.97**	**4.07**	2.23	**0.04**	17.85	12.37	65.88	123.59
dm2	14.69	18.86	2.31	0.31	32.99	41.68	93.43	235.62
zc2	16.77	10.27	1.74	0.57	**10.37**	12.12	85.29	213.26
zv2	25.36	18.36	7.77	5.84	–	–	–	–
dm3	–	–	–	–	19.82	18.00	78.79	173.42
zc3	19.94	21.01	1.10	1.33	–	–	–	–
zv3	4.08	15.90	2.23	4.94	–	–	–	–
ld3	24.05	28.20	1.20	1.11	10.28	13.06	84.31	214.49
dm4	–	–	–	–	18.37	**11.72**	**64.09**	**112.11**
zc4	–	–	–	–	–	–	–	–
zv4	33.65	31.75	3.90	3.11	–	–	–	–
ld4	25.79	30.10	**0.55**	1.29	–	–	–	–

Closure	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
(c) Example 5						
Adapt	**3.98**	**7.83**	43.38	7.24	13.63	7.76
dm2	36.18	69.35	7.59	1.82	**9.00**	1.96
zc2	475.17	84.71	43.38	6.09	93.97	3.53
zv2	–	–	–	–	–	–
dm3	20.29	30.98	16.41	5.49	30.63	5.45
zc3	56.19	200.10	292.49	11.83	231.22	14.71
zv3	–	–	–	–	–	–
ld3	56.98	171.69	14.71	**0.34**	173.90	**0.45**
dm4	–	–	–	–	–	–
zc4	–	–	–	–	–	–
zv4	–	–	–	–	–	–
ld4	9.48	15.82	**4.23**	3.87	50.61	3.72

integrating a minimal requirement for the approximation quality and force the algorithm to search through the parameter space using only stochastic simulation as long as none of the considered approximations fulfills the requirements. We also plan to apply our algorithm to larger and more challenging reaction networks, and to make a complete toolbox for moment-based parameter inference publicly available. In addition to this, in order to speed up our algorithm, we plan to introduce a trade-off between precision and computational cost of the different approximations such that the more expensive high order closure methods are only chosen when the low order closures do not provide acceptable precision.

Table 3
Run times. Columns correspond to the different reaction networks (Ex stands for example). Run time is given in seconds, both total (left) and normalized (right). We ran all experiments on a 2.26 GHz dual-core Linux notebook with 4 GB RAM.

Closure	Ex 1		Ex 2		Ex 3		Ex 4		Ex 5	
Adapt	452	0.22	4193	0.41	1511	0.75	835	0.20	1080	0.21
dm2	184	0.09	1413	0.14	537	0.26	263	0.06	454	0.09
zc2	200	0.10	–	–	379	0.18	351	0.08	744	0.14
zv2	126	0.06	–	–	354	0.17	–	–	–	–
dm3	217	0.10	1971	0.19	–	–	327	0.08	516	0.10
zc3	267	0.13	–	–	557	0.27	–	–	1473	0.29
zv3	141	0.07	–	–	948	0.47	–	–	–	–
ld3	228	0.11	2,049	0.20	901	0.45	314	0.07	503	0.10
dm4	259	0.12	2,990	0.29	–	–	388	0.09	–	–
zc4	636	0.31	–	–	–	–	–	–	–	–
zv4	163	0.08	–	–	1326	0.66	–	–	–	–
ld4	232	0.11	–	–	1492	0.74	–	–	610	0.12

Please cite this article in press as: Schilling, C., et al., Adaptive moment closure for parameter inference of biochemical reaction networks. BioSystems (2016), http://dx.doi.org/10.1016/j.biosystems.2016.07.005

Contributions

All authors designed the research. C.S. and J.R. implemented the algorithm. C.S., S.B. and J.R. analyzed the results. All authors wrote the paper.

Acknowledgments

This work is based on the CMSB 2015 paper "Adaptive moment closure for parameter inference of biochemical reaction networks" (Bogomolov et al., 2015). The work was partly supported by the German Research Foundation (DFG) as part of the Transregional Collaborative Research Center "Automatic Verification and Analysis of Complex Systems" (SFB/TR 14 AVACS[1]), by the European Research Council (ERC) under grant 267989 (QUAREM) and by the Austrian Science Fund (FWF) under grants S11402-N23 (RiSE) and Z211-N23 (Wittgenstein Award). J.R. acknowledges support from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. 291734.

References

Ale, A., Kirk, P., Stumpf, M., 2013. A general moment expansion method for stochastic kinetic models. J. Chem. Phys. 138, 174101, http://dx.doi.org/10.1063/1.4802475.
Ammar, A., Cueto, E., Chinesta, F., 2012. Reduction of the chemical master equation for gene regulatory networks using proper generalized decompositions. Int. J. Numer. Methods Biomed. Eng. 28 (9), 960–973, http://dx.doi.org/10.1002/cnm.2476.
Becskei, A., Serrano, L., 2000. Engineering stability in gene networks by autoregulation. Nature 405 (6786), 590–593, http://dx.doi.org/10.1038/35014651.
Bertaux, F., Stoma, S., Drasdo, D., Batt, G., 2014. Modeling dynamics of cell-to-cell variability in TRAIL-induced apoptosis explains fractional killing and predicts reversible resistance. PLOS Comput. Biol. 10 (10), e1003893, http://dx.doi.org/10.1371/journal.pcbi.1003893.
Bogomolov, S., Henzinger, T.A., Podelski, A., Ruess, J., Schilling, C., 2015. Adaptive moment closure for parameter inference of biochemical reaction networks. In: Proceedings of Computational Methods in Systems Biology – 13th International Conference, CMSB 2015, Nantes, France, September 16–18, 2015, pp. 77–89, http://dx.doi.org/10.1007/978-3-319-23401-4_8.
Chinesta, F., Magnin, M., Roux, O., Ammar, A., Cueto, E., 2015. Kinetic theory modeling and efficient numerical simulation of gene regulatory networks based on qualitative descriptions. Entropy 17 (4), 1896–1915, http://dx.doi.org/10.3390/e17041896.
Engblom, S., 2006. Computing the moments of high dimensional solutions of the master equation. Appl. Math. Comput. 180 (2), 498–515, http://dx.doi.org/10.1016/j.amc.2005.12.032.
Gillespie, C., Golightly, A., 2010. Bayesian inference for generalized stochastic population growth models with application to aphids. J. R. Stat. Soc. Ser. C: Appl. Stat. 59, 341–357, http://dx.doi.org/10.1111/j.1467-9876.2009.00696.x.
Gillespie, D., 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. 22 (4), 403–434, http://dx.doi.org/10.1016/0021-9991(76)90041-3.
Gillespie, D., 1992. A rigorous derivation of the chemical master equation. Physica A 188 (1-3), 404–425, http://dx.doi.org/10.1016/0378-4371(92)90283-v.
Goutsias, J., Jenkinson, G., 2013. Markovian dynamics on complex reaction networks. Phys. Reports 529, 199–264, http://dx.doi.org/10.1016/j.physrep.2013.03.004.
Hasty, J., Pradines, J., Dolnik, M., Collins, J., 2000. Noise-based switches and amplifiers for gene expression. Proc. Natl. Acad. Sci. U. S. A. 97 (5), 2075–2080, http://dx.doi.org/10.1073/pnas.040411297.
Hespanha, J., 2007. StochDynTools – a MATLAB toolbox to compute moment dynamics for stochastic networks of bio-chemical reactions, Available at http://www.ece.ucsb.edu/hespanha/software.
Hespanha, J., 2008. Moment closure for biochemical networks. In: Proceedings of the 3rd International Symposium on Communications, Control and Signal Processing (IEEE), St Julians, Malta, pp. 142–147, http://dx.doi.org/10.1109/ISCCSP.2008.4537208.
Kügler, P., 2012. Moment fitting for parameter inference in repeatedly and partially observed stochastic biological models. PLoS ONE 7 (8), e43001, http://dx.doi.org/10.1371/journal.pone.0043001.
Lück, A., Wolf, V., 2016. Generalized Method of Moments for Estimating Parameters of Stochastic Reaction Networks. arXiv:1605.01213.
Lillacci, G., Khammash, M., 2013. The signal within the noise: efficient inference of stochastic gene regulation models using fluorescence histograms and stochastic simulations. Bioinformatics 29 (18), 2311–2319, http://dx.doi.org/10.1093/bioinformatics/btt380.
Matis, T., Guardiola, I., 2010. Achieving moment closure through cumulant neglect. Math. J. 12, http://dx.doi.org/10.3888/tmj.12-2.
McAdams, H., Arkin, A., 1997. Stochastic mechanisms in gene expression. Proc. Natl. Acad. Sci. U. S. A. 94 (3), 814–819.
Munsky, B., Khammash, M., 2006. The finite state projection algorithm for the solution of the chemical master equation. J. Chem. Phys. 124, 044104, http://dx.doi.org/10.1063/1.2145882.
Munsky, B., Trinh, B., Khammash, M., 2009. Listening to the noise: random fluctuations reveal gene network parameters. Mol. Syst. Biol. 5 (1), 318, http://dx.doi.org/10.1038/msb.2009.75.
Neuert, G., Munsky, B., Tan, R., Teytelman, L., Khammash, M., van Oudenaarden, A., 2013. Systematic identification of signal-activated stochastic gene regulation. Science 339, 584–587, http://dx.doi.org/10.1126/science.1231456.
Parise, F., Lygeros, J., Ruess, J., 2015. Bayesian inference for stochastic individual-based models of ecological systems: an optimal pest control case study. Front. Environ. Sci. 3, 42, http://dx.doi.org/10.3389/fenvs.2015.00042.
Rosenfeld, N., Elowitz, M.B., Alon, U., 2002. Negative autoregulation speeds the response times of transcription networks. J. Mol. Biol. 323 (5), 785–793, http://dx.doi.org/10.1016/S0022-2836(02)00994-4.
Ruess, J., Lygeros, J., 2015. Moment-based methods for parameter inference and experiment design for stochastic biochemical reaction networks. ACM Trans. Model. Comput. Simul. 25 (2), 8, http://dx.doi.org/10.1145/2688906.
Ruess, J., Milias-Argeitis, A., Summers, S., Lygeros, J., 2011. Moment estimation for chemically reacting systems by extended Kalman filtering. J. Chem. Phys. 135, 165102.
Ruess, J., Milias-Argeitis, A., Lygeros, J., 2013. Designing experiments to understand the variability in biochemical reaction networks. J. R. Soc. Interface 10 (88), 20130588, http://dx.doi.org/10.1098/rsif.2013.0588.
Ruess, J., Parise, F., Milias-Argeitis, A., Khammash, M., Lygeros, J., 2015. Iterative experiment design guides the characterization of a light-inducible gene expression circuit. Proc. Natl. Acad. Sci. U. S. A. 112 (26), 8148–8153, http://dx.doi.org/10.1073/pnas.1423947112.
Ruess, J., 2015. Minimal moment equations for stochastic models of biochemical reaction networks with partially finite state space. J. Chem. Phys. 143, 244103, http://dx.doi.org/10.1063/1.4937937.
Samoilov, M., Arkin, A., 2006. Deviant effects in molecular reaction pathways. Nat. Biotechnol. 24 (10), 1235–1240, http://dx.doi.org/10.1038/nbt1253.
Singh, A., Hespanha, J., 2006. Lognormal moment closures for biochemical reactions. In: 45th IEEE Conference on Decision and Control, pp. 2063–2068, http://dx.doi.org/10.1109/CDC.2006.376994.
Whittle, P., 1957. On the use of the normal approximation in the treatment of stochastic processes. J. R. Stat. Soc. Ser. B: Stat. Methodol. 19, 268–281.
Wolf, V., Goel, R., Mateescu, M., Henzinger, T., 2010. Solving the chemical master equation using sliding windows. BMC Syst. Biol. 4, 42, http://dx.doi.org/10.1186/1752-0509-4-42.
Zechner, C., Ruess, J., Krenn, P., Pelet, S., Peter, M., Lygeros, J., Koeppl, H., 2012. Moment-based inference predicts bimodality in transient gene expression. Proc. Natl. Acad. Sci. U. S. A. 109 (21), 8340–8345, http://dx.doi.org/10.1073/pnas.1200161109.

[1] http://www.avacs.org/.