Systems Biology

Efficient Simulation of Intrinsic, Extrinsic and External Noise in Biochemical Systems

Dennis Pischel [1,], Kai Sundmacher [1,2] and Robert J. Flassig [2,*]

[1] Otto-von-Guericke-University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany
[2] Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg, Germany

*To whom correspondence should be addressed.

Abstract

Motivation: Biological cells operate in a noisy regime influenced by intrinsic, extrinsic, and external noise, which leads to large differences of individual cell states. Stochastic effects must be taken into account to characterize biochemical kinetics accurately. Since the exact solution of the chemical master equation, which governs the underlying stochastic process, cannot be derived for most biochemical systems, approximate methods are used to obtain a solution.

Results: In this study a method to efficiently simulate the various sources of noise simultaneously is proposed and benchmarked on several examples. The method relies on the combination of the sigma point approach to describe extrinsic and external variability and the τ-leaping algorithm to account for the stochasticity due to probabilistic reactions. The comparison of our method to extensive Monte Carlo calculations demonstrates an immense computational advantage while losing an acceptable amount of accuracy. Additionally the application to parameter optimization problems in stochastic biochemical reaction networks is shown, which is rarely applied due to its huge computational burden. To give further insight a MATLAB$^{®}$ script is provided including the proposed method applied to a simple toy example of gene expression.

Contact: pischel@mpi-magdeburg.mpg.de, sundmacher@mpi-magdeburg.mpg.de, flassig@mpi-magdeburg.mpg.de

Supplementary Information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Variability and heterogeneity are fundamental properties of biological systems. Cells differ in all kinds of attributes including cell size, protein abundances, and morphology (Spiller *et al.*, 2010), which is caused by various sources of noise. In this study we refer to intrinsic noise as an inherent stochastic biochemical process, extrinsic noise as cell-to-cell variability, and external noise as environmental fluctuations. Intrinsic noise is very dominant for small biochemical reaction systems involving low copy numbers of chemical species (*e.g.*, gene networks), whereas extrinsic and external noise increase with system size (Patnaik, 2006). Since the cellular abundance of numerous chemical species span all scales from just a few (*e.g.*, genes) to several millions (*e.g.*, proteins) all sources of noise contribute to biological heterogeneity. The interplay between intrinsic, extrinsic, and external noise and their effects on system

dynamics is hardly exploited due to experimental and numerical challenges (Spiller *et al.*, 2010; Lencastre Fernandes *et al.*, 2011; Delvigne *et al.*, 2014). Monte Carlo (MC) techniques are the standard approach to tackle stochastic biochemical reaction networks, but they suffer from an immense computational burden especially in optimization problems, where a system has to be simulated numerous times. Approximations have to be used in order to make computations feasible.

Fig. 1A illustrates a cell population that is corrupted by intrinsic, extrinsic, and external noise. In this scenario, a heterogeneous population of cells under realistic conditions, which differ in size, shape, and number of organelles is situated in an inhomogeneous medium containing concentration gradients (gray background). In addition to that, intrinsic noise caused by switching of a single gene between different states contributes to the overall variability. To clarify our understanding of intrinsic, extrinsic, and external noise their impact on a simple decay process $P \rightarrow \emptyset$ describing the degradation of the protein P is shown in
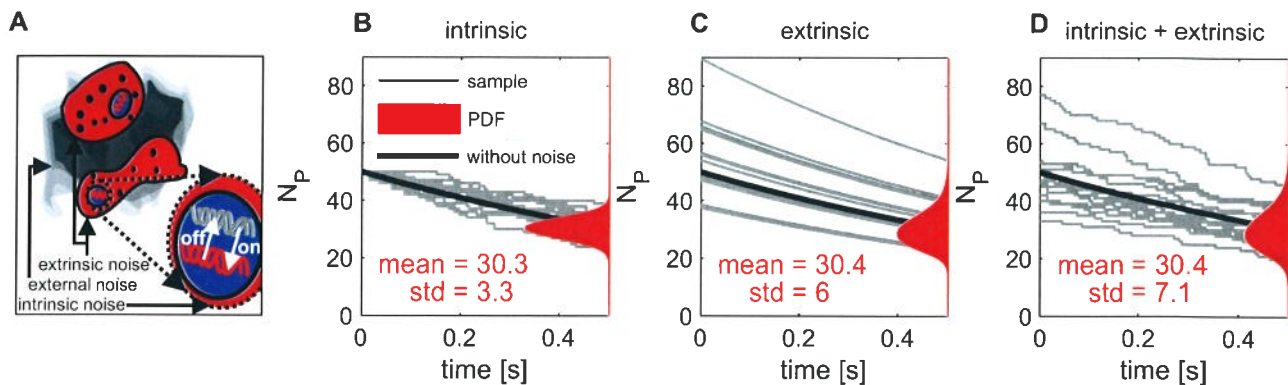
Fig. 1. Noise in biochemical systems. (A) Various sources of noise corrupt biochemical reaction systems. The influence of intrinsic noise (B), extrinsic noise (C), and intrinsic noise combined with extrinsic noise (D) on a decay process results in differences of the probability density function (PDF), the mean, and the standard deviation (std).

Fig. 1B-D. Here intrinsic noise is modeled by the Gillespie algorithm (SSA) (Gillespie, 1977) that treats reactions as stochastic events, whereas extrinsic noise is modeled by distributed initial conditions, which accounts for cell-to-cell variability. Both effects rely on different mechanisms, but result in a probability density function (PDF) characterizing the abundance of the protein (red). Adding both effects yields a further spread of the resulting PDF indicated by an increase of the standard deviation, see Fig. 1D. Note that external noise was left out in the investigation of the decay process, because in the case of a stationary random process its mathematical treatment is identical to extrinsic noise. In case the external noise is a stochastic dynamic process, its mathematical treatment is identical to intrinsic noise. For simplicity we refer from now on only to the terms extrinsic and intrinsic noise, representing a stationary random variable or a stochastic dynamic process, respectively.

In this study we propose an approximate method to model distributed stochastic processes by means of stochastic differential equations. The method relies on the combination of the sigma point (SP) approach (Julier *et al.*, 2000) accounting for extrinsic noise and the τ-leaping algorithm (Cao *et al.*, 2006) capturing probabilistic reactions. The article is organized as follows. In Sec. 2 a brief introduction to modeling of stochastic biochemical systems and a short overview of existing methods is given. In Sec. 3 the idea and advantages of the proposed method are described. In Sec. 4 a detailed benchmark on an examples of gene expression is performed followed by the application of our method on several parameter optimization examples. Sec. 5 summarizes and discusses our results. We provide supplementary material with additional information concerning our methodology and the benchmark systems. The gene expression example is illustrated with a MATLAB® script using the proposed method.

2 Theoretical Background

The simplest and most common approach to model biochemical systems is by means of ordinary differential equations (ODEs)

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}(t) = f(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\theta}) = \boldsymbol{N}\boldsymbol{v}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\theta}). \qquad (1)$$

They describe the temporal evolution of the continuous state vector $\boldsymbol{x}$ by reaction rate equations. In this context $\boldsymbol{N}$ denotes the stoichiometric matrix, $\boldsymbol{v}$ the reaction rates, $\boldsymbol{\theta}$ a set of parameters, and $\boldsymbol{u}$ external deterministic forcing. ODEs in this form neglect stochasticity introduced by intrinsic noise and fail to model the underlying process correctly. A more detailed description can be achieved with the chemical master equation

(CME) (Gillespie *et al.*, 2013)

$$\frac{\mathrm{d}}{\mathrm{d}t}P(\boldsymbol{x}(t), t) = \sum_{k=1}^{m} a_k(\boldsymbol{x}(t) - \boldsymbol{N}_k)P(\boldsymbol{x}(t) - \boldsymbol{N}_k, t) - \dots$$
$$a_k(\boldsymbol{x}(t))P(\boldsymbol{x}(t), t), \qquad (2)$$

taking into account the inherent stochastic nature of biochemical reactions. The CME governs the temporal evolution of the probability P to find the system in the discrete state $\boldsymbol{x}$. a_k denotes the propensity of reaction k and $\boldsymbol{N}_k$ the k's column of the stoichiometric matrix. For simplicity $\boldsymbol{u}$ and $\boldsymbol{\theta}$ are left out in Eq. (2), but they can be easily introduced by interpretation as additional reaction channels (Sanft *et al.*, 2011). From the solution P for all reachable states a PDF ρ describing the stochastic variable $\boldsymbol{x}$ can be reconstructed for every time point. ρ can be interpreted as a vector, whose entries represent the probability of a certain abundance interval of a chemical species. For most biochemical systems it is not possible to find an exact solution of the CME, so approximate methods have to be used. The SSA and its derivatives (Gillespie *et al.*, 2013), such as leaping (Cao *et al.*, 2006; Fu *et al.*, 2013) and time-scale separation approaches (Marchetti *et al.*, 2016) are powerful methods, which rely on a statistical mechanics ansatz treating chemical reactions as discrete molecular collision events. Although these algorithms are very popular they fail to capture variability introduced by extrinsic noise. The finite state projection algorithm (Munsky and Khammash, 2006) obtains a solution by integration of the CME. This algorithm accounts for intrinsic and extrinsic noise, but due to the curse of dimensionality it is not applicable to systems with a large state space. Another approach is the method of moments (Lakatos *et al.*, 2015), which relies on the integration of coupled ODEs for the statistical moments. This method incorporates intrinsic and extrinsic noise, but does not reconstruct the state vector's PDF and numerical instabilities make it difficult to handle (Lee *et al.*, 2009; Azunre *et al.*, 2011). A promising approach consists of solving the CME directly *via* tensor trains, but the derivation of the tensor trains is nontrivial (Kazeev *et al.*, 2014). To overcome these drawbacks we propose a novel method paving the way to further understanding of variability and heterogeneity in biochemical systems.

3 Efficient Modeling of Intrinsic and Extrinsic Noise

A straight forward approach to model intrinsic and extrinsic noise simultaneously is to perform MC sampling for extrinsic noise *e.g.*, distributed initial conditions or parameters and to compute intrinsic noise
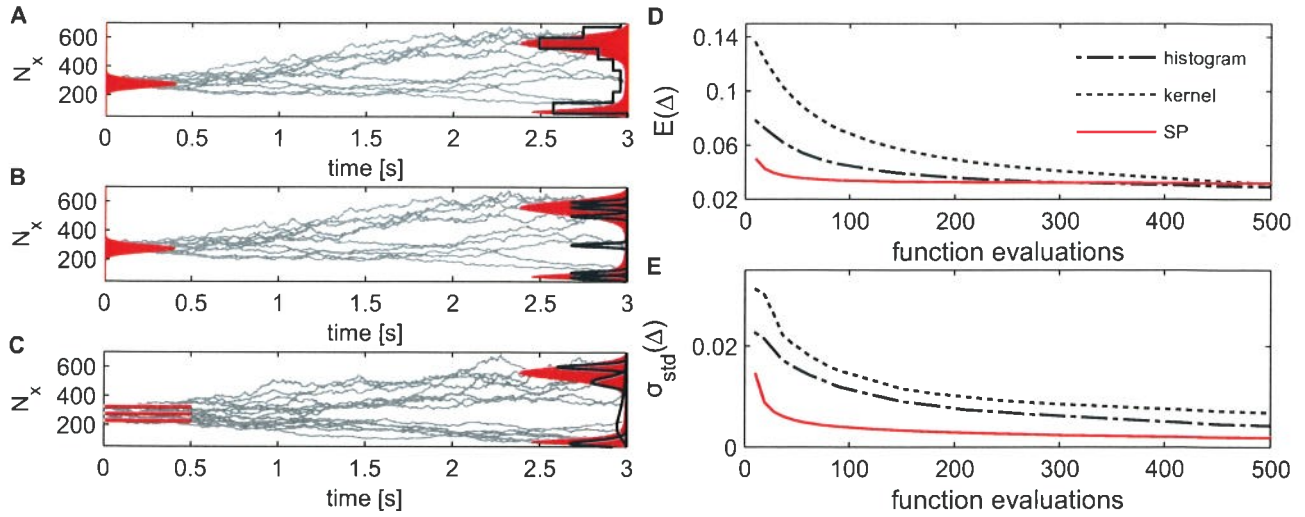
Fig. 2. Approximation of MC methods. Histogram binning (A) and kernel smoothing (B) represent common methods to derive PDFs from MC trajectories (gray) sampled from distributed initial conditions (red, left). The proposed SP approach (C) provides an efficient simulation technique sampling from $2n_\Theta + 1$ starting points (red, left). The red distribution on the right indicates the pseudo exact solution obtained by MC sampling combined with the SSA. The black histogram illustrates the histogram binning in (A) and the black distributions in (B) and (C) are summed up to obtain an approximate solution (not shown) to the exact solution. For better visualization all distributions are scaled. The accuracy (D) and convergence (E) elucidate the computational advantages of the proposed method.

i.e., the temporal evolution of a stochastic process with the SSA. In the limit of infinite function evaluations the obtained histogram is equal to the solution of the CME. Since both methods rely on the generation of random numbers they are extremely time consuming and computationally intense. In order to speed up the algorithm it is possible to lower the resolution of the histograms and use fewer, broader bins or to perform kernel smoothing, which results in the summation of kernel distributions while keeping a high resolution. This is illustrated for the Schlögl model (Schlögl, 1972) in Fig. 2A and B. In this study only Gaussian kernels are considered with the bandwidth denoting the kernel's standard deviation. The idea of the here proposed method is to further accelerate the computation by approximating the PDF of the extrinsic noise and propagating it through time by a stochastic process, see Fig. 2C. We suggest using the SP approach to account for the extrinsic noise. In contrast to MC sampling the SP approach chooses only $2n_\Theta + 1$ sigma points ϕ deterministically and estimates mean $\mathbf{E}$ and covariance $\mathbf{C}$ of a random variable Y given by a nonlinear transformation $\mathbf{h}$ (see supplementary material for further information). In this context Y refers to the abundance of chemical species and $\mathbf{h}$ governs their temporal evolution due to chemical reactions. n_Θ denotes the number of distributed initial conditions Θ, but it can also include kinetic parameters of the right hand side of Eq. (2). The samples are chosen according to

$$\phi_0 = \mathbf{E}(\Theta) \tag{3}$$

$$\phi_i = \mathbf{E}(\Theta) \pm \sqrt{(n_\Theta + \kappa)\mathbf{C}_i}, \tag{4}$$

where $\mathbf{C}_i$ denotes the ith row of the covariance matrix in the original domain. We chose the free parameter κ according to $\kappa = 3 - n_\Theta$ (see supplementary material for more information). The transformed samples

$$y_i = \mathbf{h}(\phi_i, t) \tag{5}$$

are used to compute the mean and covariance

$$\mathbf{E}(Y) = \sum_{i=0}^{2n_\Theta} w_i y_i \tag{6}$$

$$\mathbf{C}(Y) = \sum_{i=0}^{2n_\Theta} w_i (y_i - \mathbf{E}(Y)) \cdot (y_i - \mathbf{E}(Y))^T \tag{7}$$

with the weights

$$w_0 = \frac{\kappa}{n_\Theta + \kappa} \tag{8}$$

$$w_i = \frac{1}{2(n_\Theta + \kappa)}. \tag{9}$$

Assuming normality, log-normality or any other appropriate PDF characterized by mean and variance-covariance, the underlying distribution $\hat{\rho}$ can be estimated from the mean and covariance (Julier *et al.*, 2000).

To incorporate intrinsic noise due to stochastic reactions on top of the extrinsic variability the solution of CME at each point in time can be attributed a corresponding transformation $\mathbf{h}_{CME}$ of a probability function. Alternatively, when approximating the CME by sampling the underlying stochastic process with the SSA, single realizations of the SSA yield realizations of a transformation $\mathbf{h}_{SSA}$ of realizations of $\mathbf{x}$ in $\mathbb{R}^n$. Note that $\mathbf{h}_{CME}$ is a mapping between functional spaces with elements $P_t(\mathbf{x}) \mapsto P_{t+\Delta t}(\mathbf{x})$, whereas $\mathbf{h}_{SSA}$ maps from $\mathbb{R}^{n_r + n_\Theta} \rightarrow \mathbb{R}^{n_r}$, *i.e.*, $(\mathbf{x}_t, \theta) \mapsto \mathbf{x}_{t+\Delta t}$. To overcome the SSA's huge computational load we used an approximate version, which is the τ-leaping algorithm implementation of StochKit2 (Sanft *et al.*, 2011). Note that by combining intrinsic and extrinsic noise we have to deal with a distributed CME. As for real valued stochastic variables, the corresponding ensemble of probability functions or PDFs of the CME may be characterized by statistical moments. To obtain an estimate of the true average PDF ρ the scheme is repeated n times and the resulting distributions are used as

$$\bar{\rho} = \sum_{k=1}^{n} \omega_k \hat{\rho}_k = \frac{1}{n} \sum_{k=1}^{n} \hat{\rho}_k, \tag{10}$$

where we assumed that $\omega_k = 1/n$. A scheme of the algorithm is given in Fig. 3. A corresponding pseudo code can be found in the supplementary material.

The sigma points have been applied frequently to deterministic ODEs *e.g.*, (Flassig and Sundmacher, 2012; Schenkendorf *et al.*, 2009; Toni and Tidor, 2013), whereas (Toni and Tidor, 2013) have proposed a
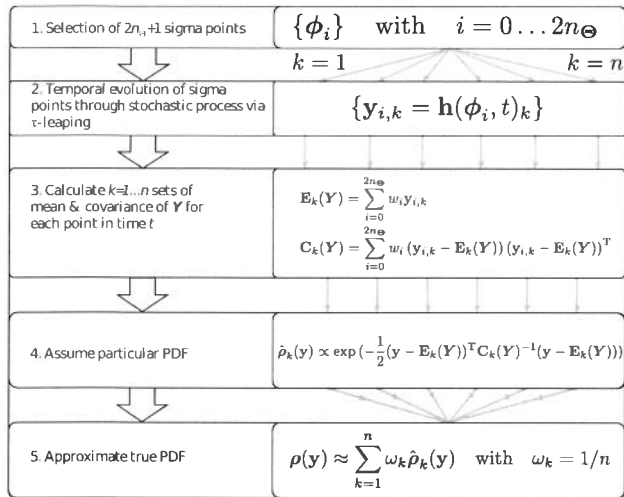
Fig. 3. Workflow of the proposed algorithm. Note that in 4, one may assume any PDF that is characterized by mean and covariance and also in 5 one may use different weights ω_k for each sample k. However the choice is a priori not clear, and we therefore suggest to use a Gaussian PDF and an equal weighting.

combination of sigma points and Ω-expansion to describe extrinsic and intrinsic variability. We are not aware of applications of sigma points to stochastic ODEs, which we apply here for sampling realizations of the CME to ultimately get an approximate solution to a distributed CME. However, the accuracy of deterministic functions given in (Julier *et al.*, 2000) should also apply for each incremental time step τ in the τ-leaping realization, since for each realization we have formally a deterministic mapping, which we can expand into a Taylor series as done in (Julier *et al.*, 2000) for the accuracy analysis. We note, however, that the convergency of $\tilde{\rho} \to \rho$ is in general not guaranteed. This depends on the ensemble of $\mathbf{h}_{SSA}$ and the choice of ω_k and $\hat{\rho}_k$. Even though our results show that the proposed approach works well, a thorough analysis is required at this point which is, however, out of scope of this contribution.
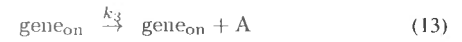
4 Benchmarking

4.1 Comparison to Extensive Monte Carlo Sampling

To compare all methods illustrated in Fig. 2 the Euclidean distance

$$\Delta = \sqrt{(\tilde{\rho} - \rho)^2} \tag{11}$$

is used as similarity measure (Cha, 2007) between the approximate solution and the pseudo exact solution obtained by MC sampling combined with the SSA. Since the final distribution of all methods is a random variable due to the underlying stochastic process, which is used for temporal evolution of the system, the computation was repeated $n_{stat} = 10^3$ times to derive a statistical statement. The mean and standard deviation of the Euclidean distance of the n_{stat} distributions are exploited as measures for accuracy and convergence. In Fig. 2D and E the dependency of accuracy and convergence on the number of function evaluations for the Schlögl model introduced in the previous section is illustrated. For the histogram and kernel method the number of function evaluations is given by the number of MC samples and for the proposed method by $n(2n_{\Theta} + 1)$. As can be seen, our proposed method outperforms the others in accuracy as well as convergence for up to 310 function evaluations, which highlights its computational benefit. For this illustration we used a kernel bandwidth of 10 and a histogram binwidth of 75.

To further demonstrate the computational efficiency of the proposed method an example of gene expression resulting in multimodal probability distributions is investigated. The gene expression system consists of the following reactions

$$\text{gene}_{\text{off}} \overset{k_1}{\underset{k_2}{\rightleftharpoons}} \text{gene}_{\text{on}} \tag{12}$$

$$\text{gene}_{\text{on}} \overset{k_3}{\rightarrow} \text{gene}_{\text{on}} + A \tag{13}$$

$$A + B \overset{k_4}{\rightarrow} \varnothing. \tag{14}$$

A single gene is considered which is able to flip between an active and inactive state. In the active state the gene produces protein A, which is degraded by protein B. This example involves intrinsic noise due to the low abundance of the gene, but also extrinsic noise due to the distributed initial conditions of protein B. In Fig. 4A the proposed method is illustrated for protein A. Several distributions of $\hat{\rho}_k$ (gray, dashed) are averaged to obtain the resulting approximate distribution $\tilde{\rho}$ (gray, solid). The approximate solution mimics the distributed character of the pseudo reference obtained by MC sampling combined with the SSA, which constitutes a bimodal distribution with - regarding any approximation approach - a challenging sharp peak for low abundances. The temporal evolution of the distributions of protein A and B are shown in Fig. 4B and E. The approximation is very similar to the solution obtained by MC sampling combined with the SSA demonstrating the capability to qualitatively model intrinsic and extrinsic noise simultaneously. Furthermore, accuracy and convergence were investigated in comparison to the kernel approach. Since *a priori* the optimal selection of the kernel bandwidth is not clear several bandwidths were tested systematically. In Fig. 4C and F the difference of the mean Euclidean distance of the proposed method and the kernel approach $\mathbf{E}(\Delta_{Kernel}) - \mathbf{E}(\Delta_{SP})$ is shown for the time point $t = 1000$ s. If this term takes positive values (shades of red) the proposed method outperforms the kernel approach and for negative values (white) the kernel approach excels. The same applies to the difference of the standard deviation of the Euclidean distance $\sigma_{std}(\Delta_{Kernel}) - \sigma_{std}(\Delta_{SP})$. For protein A it can be seen that for less than $1.5 \cdot 10^3$ function evaluations the accuracy of the proposed method is always higher than the accuracy of the kernel method. For more function evaluations an optimal bandwidth yields better results at the price of higher computational costs, see Fig. 4C. For protein B the proposed method is superior for all bandwidths and function evaluations, see Fig. 4F. Concerning the convergence it was found that the proposed method outperforms the kernel approach for both proteins, except for kernel bandwidths much larger than the spread of the underlying distribution resulting in very low accuracy, see Fig. 4D and G. With this example including multimodality, low, and high copy numbers of chemical species the strengths of the proposed method has been clearly demonstrated. Note that *a priori* the optimal kernel bandwidth is not known and changes with time meaning that this parameter has to be estimated for the optimal representation of every PDF. The proposed method avoids this computational demanding task making it easy to handle and applicable.

4.2 Application to Parameter Optimization

Having shown that the proposed methods yields convincing results we utilize it for optimization of several biochemical reaction networks. In biology it is not possible to measure all parameters directly that are necessary for computational modeling leading to parameter estimation problems. The exact simultaneous simulation of intrinsic and extrinsic noise is computational extremely intense, wherefore approximate methods are needed. In this subsection we use the proposed method to estimate the unknown rate constants of five different example systems described in Tab. 1. Therefore six reference measurements computed with the SSA
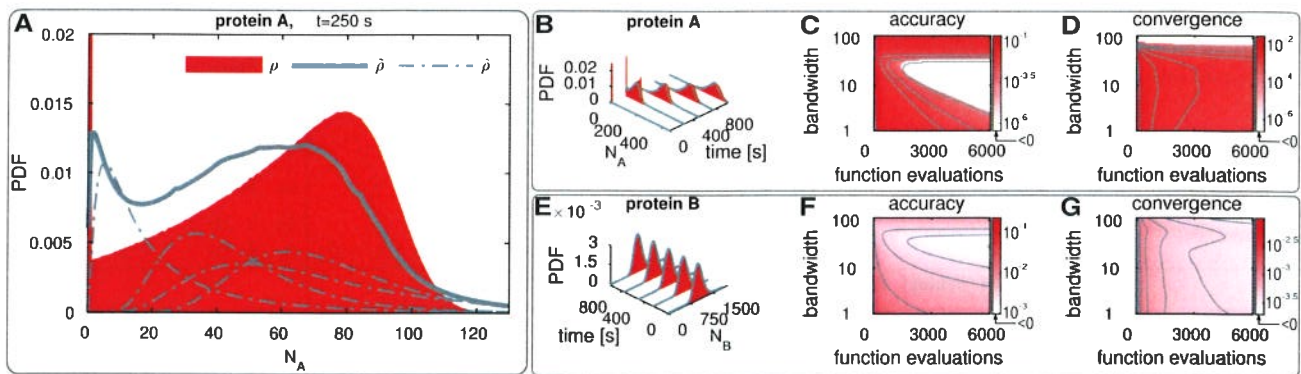
Fig. 4. Comparison of performance. (A) Averaging of $\hat{\rho}_k$ (gray, dashed) yields an approximate solution $\tilde{\rho}$ (gray, solid) of the CME for protein A at the time point 250 s. The pseudo exact solution obtained by MC sampling combined with the SSA is shown in red. For better visualization the $\hat{\rho}_k$ are scaled and hence smaller than the approximate distribution. The corresponding temporal evolution of the probability distributions for protein A and B are illustrated in (B) and (E). In (C) and (F) the difference in accuracy of the proposed method and the kernel approach is shown for protein A and B. Shades of red indicate superior and white the inferior accuracy of the proposed method (all negative values are marked white). The difference in convergence is illustrated in (D) and (G) for protein A and B with a similar color code.

and MC sampling with equal time spacing were used for comparison with the results of our approximate method. For the example of gene expression only protein A and B are utilized for optimization, whereas for the other examples all chemical species were used. The objective function was the sum of the eucledian distance between the pseudo reference and the approximate solution for all chemical species and time points. The optimization was performed with a genetic algorithm, since gradient based algorithms fail for stochastic systems and get easily stuck in local extrema (Poovathingal and Gunawan, 2010). For all example systems 500 generations with a population size of 40 were used. In Tab. 1 the true parameters k_{true} and the optimized ones k_{opti} are provided. For all example systems the optimized parameters are very close to the true parameters indicating that the proposed method yields excellent results and depicts a promising technique for optimization of stochastic biochemical reaction systems. In the supplementary material further comments regarding the benchmarking can be found.

5 Discussion

In this study the problem of efficiently simulating biochemical reaction systems containing intrinsic and extrinsic noise is addressed. We propose a novel algorithm relying on the SP approach and the τ-leaping algorithm, which computes approximate solutions of the distributed CME. The method is benchmarked on several examples illustrating its computational benefit compared to others. Choosing only $2n_\Theta + 1$ SPs deterministically the proposed method converges very fast to an approximate solution. The accuracy of our approximative approach is barely dependent on system size but rather on the complexity of the stochastic mapping itself. This can be seen when the estimated PDFs of the Schlögl are compared to the Virus model (one vs. four states). Therefore our results are very likely to also apply for large reaction systems. The strength of the method, i.e., fast convergency at reduced sample size, also introduces some of its weaknesses. Since only a few samples from the initial distribution are chosen it might be sampled too sparsely. Additionally the choice of PDF kernels is *a priori* not clear. Although the proposed method might not converge to the exact solution of the CME (see saturation behavior in Fig. 2D) it qualitatively describes important characteristics of the underlying distribution (see Fig. 4A, B, E, and also additional illustrations of approximate vs. pseudo exact solution in the supplementary material).

The method may serve as a tool for rapid analysis and optimization of a stochastic system. For precise predictions, we suggest to use our

method as a starting point for refinement by means of numerically demanding, but asymptotically exact methods. Although there is a lot of development making stochastic analysis tools with state of the art algorithms and fast implementations (Drawert et al., 2016; Fan et al., 2016; Somogyi et al., 2015) available for a broad community the optimization of stochastic biochemical systems is hardly performed due to its challenging computational effort (Poovathingal and Gunawan, 2010). With this study we demonstrated that the proposed method is capable of describing biochemical systems containing intrinsic and extrinsic noise and it represents a promising tool suited for optimization and analysis of distributed, stochastic biochemical reaction systems.

Supplementary Material

Further details on the models, our algorithm, and methods can be found in the supplementary material. We provide a MATLAB® script, which can be used to execute the proposed method on the example of gene expression. This script can be easily adapted for the simulation of other models. All models are provided in SBML format to run simulations in other software.

References

Azunre, P., Gómez-Uribe, C., and Verghese, G. (2011). Mass fluctuation kinetics: Analysis and computation of equilibria and local dynamics. *IET Systems Biology*, 5(6), 325–335.

Cao, Y., Gillespie, D. T., and Petzold, L. R. (2006). Efficient step size selection for the tau-leaping simulation method. *Journal of Chemical Physics*, 124(4).

Cha, S.-H. (2007). Comprehensive survey on distance/similarity measures between probability density functions. *International Journal of Mathematical Models and Methods in Applied Sciences*, 1(4), 300–307.

Delvigne, F., Zune, Q., Lara, A. R., Al-Soud, W., and Sørensen, S. J. (2014). Metabolic variability in bioprocessing: Implications of microbial phenotypic heterogeneity. *Trends in Biotechnology*, 32(12), 608–616.

Drawert, B., Hellander, A., Bales, B., Banerjee, D., Bellesia, G., Daigle, Jr., B. J., Douglas, G., Gu, M., Gupta, A., Hellander, S., Horuk, C., Nath, D., Takkar, A., Wu, S., Lötstedt, P., Krinzt, C., and Petzold, L. R. (2016). Stochastic simulation service: Bridging the gap between the computational expert and the biologist. *PLoS Computational Biology*, 12(12).

Table 1. Benchmark model description.

model	initial conditions		k_{true}		k_{opti}
Gene Model:					
$\text{gene}_{\text{off}} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \text{gene}_{\text{on}}$	N_{ON}	0	k_1	10^{-2}	$8.8 \cdot 10^{-3}$
$\text{gene}_{\text{on}} \overset{k_3}{\rightarrow} \text{gene}_{\text{on}} + \text{A}$	N_{OFF}	1	k_2	10^{-3}	$6.6 \cdot 10^{-4}$
$\text{A} + \text{B} \overset{k_4}{\rightarrow} \varnothing$	N_A	0	k_3	$5 \cdot 10^{-1}$	$5.2 \cdot 10^{-1}$
	N_B	$e^{\mathcal{N}}(\mu = 10^3, \sigma = 1.5 \cdot 10^2)$	k_4	$5 \cdot 10^{-7}$	$5.2 \cdot 10^{-7}$
Schlögl Model (Schlögl, 1972):					
$2\text{X} + \text{A} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 3\text{X}$	N_X	$e^{\mathcal{N}}(\mu = 2.75 \cdot 10^2, \sigma = 5)$	k_1	$3 \cdot 10^{-7}$	$3.1 \cdot 10^{-7}$
$\text{B} \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \text{X}$	N_A	10^5 (const.)	k_2	10^{-4}	10^{-4}
	N_B	$2 \cdot 10^5$ (const.)	k_3	10^{-3}	10^{-3}
			k_4	3.5	3.5
Michaelis-Menten-Kinetics (Michaelis and Menten, 1913):					
$\text{E} + \text{S} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \text{ES}$	N_E	$2.5 \cdot 10^2$	k_1	10^{-4}	10^{-4}
$\text{ES} \overset{k_3}{\rightarrow} \text{P}$	N_S	$\mathcal{N}(\mu = 10^4, \sigma = 10^3)$	k_2	$5 \cdot 10^{-3}$	$2.4 \cdot 10^{-3}$
	N_{ES}	0	k_3	10^{-1}	10^{-1}
	N_P	0			
Virus Model (Gupta and Rawlings, 2014):					
$\text{V} \overset{k_1}{\rightarrow} \text{G}$	N_V	$e^{\mathcal{N}}(\mu = 50, \sigma = 5)$	k_1	$1.5 \cdot 10^{-1}$	$1.4 \cdot 10^{-1}$
$\text{G} \overset{k_2}{\rightarrow} \text{G} + \text{M}$	N_G	0	k_2	$2 \cdot 10^{-2}$	$2.1 \cdot 10^{-2}$
$\text{G} \overset{k_3}{\rightarrow} 2\text{G}$	N_M	0	k_3	$5 \cdot 10^{-2}$	$5.1 \cdot 10^{-2}$
$\text{M} \overset{k_4}{\rightarrow} \text{M} + \text{P}$	N_P	0	k_4	1	$9.9 \cdot 10^{-1}$
Yeast Model (Poovathingal and Gunawan, 2010):					
$\text{PC}_3 \overset{k_1}{\rightarrow} \text{PC}_3 + \text{mRNA}_G$	N_{PC}	$e^{\mathcal{N}}(\mu = 50, \sigma = 10)$	k_1	$5 \cdot 10^{-1}$	$5.1 \cdot 10^{-1}$
$\text{mRNA}_G \overset{k_2}{\rightarrow} \varnothing$	N_{mRNA_G}	0	k_2	5	5.1
$\text{mRNA}_G \overset{k_3}{\rightarrow} \text{mRNA}_G + \text{yEGFP}$	N_{yEGFP}	0	k_3	1	1
$\text{yEGFP} \overset{k_4}{\rightarrow} \varnothing$	N_{mRNA_R}	0	k_4	$2 \cdot 10^{-2}$	$2.1 \cdot 10^{-2}$
$\varnothing \overset{k_5}{\rightarrow} \text{mRNA}_R$	N_{TetR}	0	k_5	$5 \cdot 10^{-1}$	$5.3 \cdot 10^{-1}$
$\text{mRNA}_R \overset{k_6}{\rightarrow} \varnothing$			k_6	$2 \cdot 10^{-1}$	$2 \cdot 10^{-1}$
$\text{mRNA}_R \overset{k_7}{\rightarrow} \text{mRNA}_R + \text{TetR}$			k_7	1	1
$\text{TetR} \overset{k_8}{\rightarrow} \varnothing$			k_8	$2 \cdot 10^{-2}$	$2.1 \cdot 10^{-2}$

Fan, S., Geissmann, Q., Lakatos, E., Lukauskas, S., Ale, A., Babtie, A., Kirk, P., and Stumpf, M. (2016). Means: Python package for moment expansion approximation, inference and simulation. *Bioinformatics*, **32**(18), 2863–2865.

Flassig, R. J. and Sundmacher, K. (2012). Optimal design of stimulus experiments for robust discrimination of biochemical reaction networks. *Bioinformatics*, **28**(23), 3089–3096.

Fu, J., Wu, S., and Petzold, L. R. (2013). Time dependent solution for acceleration of tau-leaping. *Journal of Computational Physics*, **235**, 446–457.

Gillespie, D., Hellander, A., and Petzold, L. (2013). Perspective: Stochastic algorithms for chemical kinetics. *Journal of Chemical Physics*, **138**(17).

Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, **81**(25), 2340–2361.

Gupta, A. and Rawlings, J. B. (2014). Comparison of parameter estimation methods in stochastic chemical kinetic models: Examples in systems biology. *AIChE Journal*, **60**(4), 1253–1268.

Julier, S., Uhlmann, J., and Durrant-Whyte, H. F. (2000). A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, **45**(3), 477–482.

Kazeev, V., Khammash, M., Nip, M., and Schwab, C. (2014). Direct solution of the chemical master equation using quantized tensor trains. *PLoS Computational Biology*, **10**(3).

Lakatos, E., Ale, A., Kirk, P., and Stumpf, M. (2015). Multivariate moment closure techniques for stochastic kinetic models. *Journal of Chemical Physics*, **143**(9), 094107.

Lee, C. H., Kim, K.-H., and Kim, P. (2009). A moment closure method for stochastic reaction networks. *Journal of Chemical Physics*, **130**(13).

Lencastre Fernandes, R., Nierychlo, M., Lundin, L., Pedersen, A. E., Puentes Tellez, P. E., Dutta, A., Carlquist, M., Bolic, A., Schäpper, D., Brunetti, A. C., Helmark, S., Heins, A.-L., Jensen, A. D., Nopens, I., Rottwitt, K., Szita, N., van Elsas, J. D., Nielsen, P. H., Martinussen, J., Sørensen, S. J., Lantz, A. E., and Gernaey, K. V. (2011). Experimental methods and modeling techniques for description of cell population heterogeneity. *Biotechnology Advances*, **29**(6), 575–599.

Marchetti, L., Priami, C., and Thanh, V. H. (2016). HRSSA - efficient hybrid stochastic simulation for spatially homogeneous biochemical reaction networks. *Journal of Computational Physics*, **317**, 301–317.

Michaelis, L. and Menten, M. L. (1913). Die kinetik der invertinwirkung. *Biochemische Zeitschrift*, **49**, 333–369.

Munsky, B. and Khammash, M. (2006). The finite state projection algorithm for the solution of the chemical master equation. *Journal of Chemical Physics*, **124**(4).

Patnaik, P. R. (2006). External, extrinsic and intrinsic noise in cellular systems: analogies and implications for protein synthesis. *Biotechnology and Molecular Biology Reviews*, **1**(4), 121–127.

Poovathingal, S. K. and Gunawan, R. (2010). Global parameter estimation methods for stochastic biochemical systems. *BMC Bioinformatics*, **11**.

Sanft, K. V., Wu, S., Roh, M., Fu, J., Lim, R. K., and Petzold, L. R. (2011). Stochkit2: Software for discrete stochastic simulation of biochemical systems with events. *Bioinformatics*, **27**(17), 2457–2458.

Schenkendorf, R., Kremling, A., and Mangold, M. (2009). Optimal experimental design with the sigma point method. *IET Systems Biology*, **3**(1), 10–23.

Schlögl, F. (1972). Chemical reaction models for non-equilibrium phase transitions. *Zeitschrift für Physik*, **253**(2), 147–161.

Somogyi, E., Bouteiller, J.-M., Glazier, J., König, M., Medley, J., Swat, M., and Sauro, H. (2015). LibRoadRunner: A high performance SBML simulation and analysis library. *Bioinformatics*, **31**(20), 3315–3321.

Spiller, D. G., Wood, C. D., Rand, D. A., and White, M. R. H. (2010). Measurement of single-cell dynamics. *Nature*, **465**(7299), 736–745.

Toni, T. and Tidor, B. (2013). Combined model of intrinsic and extrinsic variability for computational network design with application to synthetic biology. *PLoS Computional Biology*, **9**(3), e1002960.