Hybrid models for chemical reaction networks: Multiscale theory and application to gene regulatory systems

Stefanie Winkelmann, and Christof Schütte

Hybrid models for chemical reaction networks: Multiscale theory and application to gene regulatory systems

Stefanie Winkelmann[1,2,a)] and Christof Schütte[1,2,b)]
[1]*Zuse Institute Berlin (ZIB), Takustr. 7, 14195 Berlin, Germany*
[2]*Department of Mathematics, Freie Universität Berlin, Arnimallee 6, 14195 Berlin, Germany*

Well-mixed stochastic chemical kinetics are properly modeled by the chemical master equation (CME) and associated Markov jump processes in molecule number space. If the reactants are present in large amounts, however, corresponding simulations of the stochastic dynamics become computationally expensive and model reductions are demanded. The classical model reduction approach uniformly rescales the overall dynamics to obtain deterministic systems characterized by ordinary differential equations, the well-known mass action reaction rate equations. For systems with multiple scales, there exist hybrid approaches that keep parts of the system discrete while another part is approximated either using Langevin dynamics or deterministically. This paper aims at giving a coherent overview of the different hybrid approaches, focusing on their basic concepts and the relation between them. We derive a novel general description of such hybrid models that allows expressing various forms by one type of equation. We also check in how far the approaches apply to model extensions of the CME for dynamics which do not comply with the central well-mixed condition and require some spatial resolution. A simple but meaningful gene expression system with negative self-regulation is analysed to illustrate the different approximation qualities of some of the hybrid approaches discussed. Especially, we reveal the cause of error in the case of small volume approximations. *Published by AIP Publishing.* https://doi.org/10.1063/1.4986560

I. INTRODUCTION

Stochastic modeling of biochemical reaction networks, such as gene expression systems, is of fundamental interest in molecular biology.[1–4] Of central relevance is the description by the chemical master equation (CME) and associated Markov jump processes where the discrete states refer to numbers for all species of reactant molecules. The dynamics can be simulated by Gillespie's stochastic simulation algorithm[5–7] that generates exact realizations of the underlying jump process. However, the simulations may get extremely time-consuming, especially if parts of the molecular population appear in high copy numbers or reaction rates are high. For such situations, various approximation methods and reduced models have been developed.

Beside the numerical schemes to speed up simulations (like τ-leaping methods[8]), there exist analytical approaches to reduce the system's complexity by defining simplified models. If the copy number is large for all involved species, the CME can be approximated by stochastic differential equations or ordinary differential equations (ODE's)—this is a classical result proven by Kurtz.[9–11] The approximation is based on an equal scaling of all molecule numbers. In many applications—as in genetic networks—however, some of the involved species (e.g., proteins) are typically present in much greater abundance than others (e.g., DNA), which makes an equal scaling inappropriate. This leads to hybrid methods where only a

part of the system is approximated by continuous dynamics, whereas species in small amounts remain discrete. The most prominent hybrid approach is given by piecewise deterministic Markov processes (PDMP's) which were introduced by Davis in 1984.[12] As the name suggests, these processes move deterministically for some period of time before the flow is interrupted by a random jump. PDMP's have been extensively studied in Refs. 13–15, including a convergence analysis and an application to autocatalytic genetic networks. Similar methods have been proposed in Refs. 16–18, with error estimates conducted in Ref. 19.

Other existing hybrid approaches are based on a partition of the set of reactions into fast and slow ones, modeling the fast reactions by continuous Markov processes while the slow reactions remain discrete.[20–25] Although the resulting hybrid algorithms significantly reduce the computational time, the underlying separation of reactions might be ambiguous and will not always provide a convenient description of the given dynamics. For many applications, the more natural partition is given in terms of population size levels.

In this paper, we focus on such natural partitions, i.e., species of reacting molecules will be partitioned into two or more regimes according to their abundance and not according to the speed of the reaction channels they are involved in. We give a short description of PDMP's and expand the model formulation by allowing noise around the deterministic trends of the rescaled species, which leads to piecewise chemical Langevin equations (PCLE's), also called "hybrid diffusions."[26–28] As a consequence, we arrive at a novel joint equation for all hybrid models with (a) low copy number

a)stefanie.winkelmann@fu-berlin.de
b)schuette@mi.fu-berlin.de

species with discrete molecule number states and stochastic jump process description, (b) medium copy number species with continuous molecule density states and stochastic PCLE-type description, mixed with (c) high copy number species with continuous molecule density states and deterministic reaction rate equation description, given by

$$
\begin{aligned}
\boldsymbol{Y}(t) = \boldsymbol{Y}(0) &+ \sum_{k \in \mathcal{K}_{low}} \mathcal{P}_k \left(\int_0^t \tilde{\alpha}_k(\boldsymbol{Y}(s))\, ds \right) \tilde{\boldsymbol{\nu}}_k \\
&+ \sum_{k \in \mathcal{K}_{medium}} \int_0^t \tilde{\alpha}_k(\boldsymbol{Y}(s)) \tilde{\boldsymbol{\nu}}_k\, ds + \int_0^t \Xi(\boldsymbol{Y}(s))\, d\boldsymbol{W}(s) \\
&+ \sum_{k \in \mathcal{K}_{high}} \int_0^t \tilde{\alpha}_k(\boldsymbol{Y}(s)) \tilde{\boldsymbol{\nu}}_k\, ds,
\end{aligned} \tag{1}
$$

where the state $\boldsymbol{Y}(t)$ is composed of the molecule numbers as well as molecule densities in the system at time t, the reactions are numbered by k and grouped according to their nature (i.e., their impact onto the three classes of species) into sets $\mathcal{K}_{low}$, $\mathcal{K}_{medium}$, and $\mathcal{K}_{high}$, while $\tilde{\alpha}_k$ and $\tilde{\boldsymbol{\nu}}_k$ denote the correspondingly scaled propensity functions and stochiometric vectors of the reactions, $\mathcal{P}_k$ are standard Poisson clock processes, Ξ defines appropriately scaled noise intensities, and $\boldsymbol{W}(s)$ denote standard vector-valued white noise processes. The details of how to derive all these quantities for a given chemical reaction network are explained in Sec. III.

For illustration, we investigate a gene expression system with negative self-regulation which very well demonstrates the different approximation properties of PDMP's and PCLE's; especially it reveals that PCLE's are better suited to reproduce empirical distributions and sizes of stochastic events like protein bursts, which are a typical characteristic of self-repressing gene expression systems. Moreover, the proposed genetic network provides a basis for clarifying the cause of a small volume failure: What exactly goes wrong if the approximative rescaling is applied to species with low copy numbers?

When considering the gene expression system within a eukaryotic cell, some spatial resolution is required because the individual reactions may be subject to local restrictions. This motivates to extend the analysis to reaction-diffusion systems described by spatiotemporal CME's. We will give a short introduction to such extended CME's and show that the associated reaction-diffusion dynamics can suitably be approximated by the presented hybrid methods, as well.

The article intends to give a structured overview and joint description of the existing model formulations and to illustrate the relation between them. However, also novel recombinations of the models will be presented and analysed. In Sec. II, we review the CME formalism and its approximation by chemical Langevin equations (CLE's) and reaction-rate equations (RRE's). Also the spatiotemporal chemical master equation (ST-CME) is depicted as a model extension for non-well-mixed reaction systems. Hybrid models for multiscale dynamics are presented in Sec. III, where we derive consistent time-change representations of the respective combined processes. Finally, Sec. IV contains the application to a gene regulatory system, including a detailed error analysis giving new insights into the structure of multiscale approximations.

II. FROM CME TO ODE: MODELING REACTION NETWORKS

In the following, we give a short review of several modeling approaches for chemical reaction kinetics. Starting with the chemical master equation as the most accurate model for well-mixed reaction systems in Sec. II A, we motivate the relevance of approximative modeling approaches, some of which are presented in Sec. II C. Finally, we show in Sec. II D how the CME-model can be extended to situations where the well-mixed condition is broken and some spatial resolution is required to capture the characteristics of the dynamics.

A. The chemical master equation

We consider a system of $L \in \mathbb{N}$ chemical species $S_1, \ldots, S_L$ which behave well-mixed in a fixed space of motion Ω. The system is affected by $K \in \mathbb{N}$ reaction channels $R_1, \ldots, R_K$. The state of the system at time $t \geq 0$ is given by $\boldsymbol{X}(t) = (X_1(t), \ldots, X_L(t)) \in \mathbb{N}_0^L$ with $X_l(t)$ denoting the number of molecules of species S_l at time t. The effect of reaction R_k is given by the state-change vector $\boldsymbol{\nu}_k = (\nu_{1k}, \ldots, \nu_{Lk}) \in \mathbb{Z}^L$ where ν_{lk} defines the net change in the number of molecules of species S_l due to reaction R_k. If the reaction is represented by the stoichiometric equation

$$
R_k : \quad a_{1k}S_1 + \cdots + a_{Lk}S_L \rightarrow a'_{1k}S_1 + \cdots + a'_{Lk}S_L, \tag{2}
$$
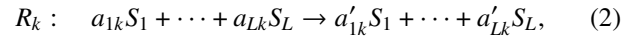
with coefficients $a_{lk}, a'_{lk} \in \mathbb{N}_0$, the related state-change vector has entries $\nu_{lk} = a'_{lk} - a_{lk}$.[29]

Given $\boldsymbol{X}(t) = \boldsymbol{x} = (x_1, \ldots, x_L)$, the corresponding transition

$$
\boldsymbol{x} \rightarrow \boldsymbol{x} + \boldsymbol{\nu}_k
$$

occurs at rate $\alpha_k(\boldsymbol{x}) > 0$ where α_k is a well-defined propensity function. According to the law of mass action, the propensity is a function of the corresponding macroscopic rate constant $\gamma_k > 0$ and the number of molecules involved in the reaction. For example, for a unimolecular reaction by species S_l, it holds that $\alpha_k(\boldsymbol{x}) = \gamma_k x_l$ and for a bimolecular reaction of two species $S_l, S_{l'}$, $l \neq l'$, it holds that $\alpha_k(\boldsymbol{x}) = \frac{\gamma_k}{V} x_l x_{l'}$, where V is the system's volume times the Avogadro constant.[30] The inverse scaling by the volume for the bimolecular reaction reflects the fact that inside a larger volume, it takes more time for two reactant molecules to find each other. Analogue arguments for a reaction as in (2) lead to the general propensity function

$$
\alpha_k(\boldsymbol{x}) = \gamma_k V \prod_{l=1}^{L} \frac{x_l!}{(x_l - a_{lk})! V^{a_{lk}}},
$$

see Ref. 31 for details.

In order to state the time evolution of $\boldsymbol{X}(t)$, let $\mathcal{R}_k(t)$ denote the number of times the reaction R_k has occurred by time t. Then, obviously,

$$
\boldsymbol{X}(t) = \boldsymbol{X}(0) + \sum_{k=1}^{K} \mathcal{R}_k(t)\boldsymbol{\nu}_k.
$$

Given that the reaction propensities depend only on the current value $\boldsymbol{X}(t)$ of the system, the model satisfies the Markov property and the counting processes $\mathcal{R}_k$ can be represented in terms of Poisson processes. This gives the *time-change representation*

$$X(t) = X(0) + \sum_{k=1}^{K} \mathcal{P}_k \left(\int_0^t \alpha_k(X(s)) ds \right) \nu_k \qquad (3)$$

of the continuous-time Markov process $X(t)$, with $\mathcal{P}_k$, $k = 1, \ldots, K$, denoting independent, unit-rate Poisson processes.[32]

As an alternative to this time-change representation, the dynamics can be characterized by the Kolmogorov forward equation for the distribution of $X(t)$. Defining

$$P(x, t) := \text{Prob}(X(t) = x | X(0) = x_0)$$

for some initial state $x_0 \in \mathbb{N}_0^L$, the system of differential equations

$$\frac{\partial P(x,t)}{\partial t} = \sum_{k=1}^{K} (\alpha_k(x - \nu_k) P(x - \nu_k, t) - \alpha_k(x) P(x, t)) \quad (4)$$

is fulfilled, which in the context of reaction networks is called the *chemical master equation* (CME).

B. Stochastic simulation

The CME (4) is the Fokker-Planck equation associated with the Markov jump process (3). In general, its state space is not finite, that is, the CME consists of a countable (but infinite) set of coupled ODE's which, even for rather simple chemical reaction systems, becomes very complex. Solving the CME analytically is impossible in almost all situations of practical interest. Instead, numerical solution methods can be applied, with the methods falling into two main categories: numerical integration methods for (4) that aim at approximating the probability distribution $P(x, t)$ and sampling techniques that compute an ensemble of realization of the jump process (3) and then approximate expectation values and higher moments of the probability distribution $P(x, t)$ by statistical means. Gillespie's stochastic simulation algorithm (SSA)[7] is the most prominent sampling technique. In its basic form, it uses the fact that given a state $x \in \mathbb{N}_0^L$, the random waiting time for the next reaction to occur follows an exponential distribution with mean $1/\alpha_0(x)$, where

$$\alpha_0(x) := \sum_{k'=1}^{K} \alpha_{k'}(x).$$

The index k of the next reaction is a statistically independent integer random variable with point probability $\alpha_k(x)/\alpha_0(x)$. In total, the stochastic simulation algorithm (SSA) is the following:

1. Initialize time $t = t_0$ and state $x = x_0$ and choose a time horizon $T > t_0$.
2. Calculate $\alpha_k(x)$ for all $k = 1, \ldots, K$ and their sum $\alpha_0(x)$.
3. Generate two random numbers r_1, r_2 from independent uniform distributions in $[0, 1]$, set

$$\tau = \frac{1}{\alpha_0(x)} \ln \left(\frac{1}{r_1} \right)$$

   and choose k to be the smallest integer satisfying $\sum_{k'=}^{k} \alpha_{k'}(x) > r_2 \alpha_0(x)$.
4. Execute the next reaction by replacing $t \leftarrow t + \tau$ and $x \leftarrow x + \nu_k$.
5. Return to 2 or end the simulation in case of $t \geq T$.

Obviously, the SSA is easy to implement and results in exact realizations of the jump process. There exist reformulations of it which reduce the computational effort per iteration step (i.e., per reaction event), see, e.g., the next-reaction method by Gibson and Bruck[33] or other methods in Refs. 34 and 35. Such alternative versions of the SSA are useful; however, they do not decrease the number of iteration steps necessary (on average) to reach a given time horizon. In the case of closely following reaction events, the so-called *dense accumulations*, the time increase τ in each iteration step is small—which induces a long runtime, no matter how effective the reaction events are determined in each step. This motivates to forgo the perfect exactness of the SSA in order to find faster approximate simulation strategies. One of them is the so-called τ-leaping method which is based on the idea of aggregating several reaction events per iteration step.[7] The analytical background is the following. Given a state $X(t) = x$ at time t, let $\tau > 0$ be small enough such that during the time interval $[t, t + \tau]$ the propensity functions of all reactions are likely to stay roughly constant. Thereby, the number of times reaction R_k fires during $[t, t + \tau]$ follows a Poisson distribution with mean $\alpha_k(x)\tau$, and the state at time $t + \tau$ can be approximated by

$$X(t + \tau) \approx x + \sum_{k=1}^{K} \mathcal{P}_k(\alpha_k(x)\tau)\nu_k, \qquad (5)$$

where $\mathcal{P}_k(\mu_k)$ are statistically independent Poisson variables with mean (and variance) μ_k. Equation (5) is an approximation of (3) and obviously suggests to apply basic numerical schemes—like the explicit Euler method—for approximative simulations of $X(t)$: Given a lag-time $\tau > 0$ as well as a state $X(t_j) = x_j$ at some time $t_j \geq 0$, generate for each k an independent Poisson random number $\mathcal{P}_k(\alpha_k(x_j)\tau)$ and set $x_{j+1} = x_j + \sum_{k=1}^{K} \mathcal{P}_k(\alpha_k(x)\tau)\nu_k$ as an approximation of $X(t_{j+1})$. For more details and practical issues (like how to choose the lag-time τ), we refer to Refs. 7, 32, and 36.

C. Approximation methods

Dense accumulations of reaction events occur for any reaction rate constants γ_k if the number of reactant molecules is large. In this case, a direct analytical approximation of the CME by a chemical Langevin equation or an ODE is reasonable not only to speed up numerical simulations but also to uncover the structure of the chemical system. We give a short summary of the stepwise approximation, for details see, e.g., Refs. 7, 9–11, and 30.

As before, let $\tau > 0$ be small enough such that an approximation of the dynamics by (5) is justified. If, on the other hand, τ is also large enough for all reaction channels to fire several times [i.e., $\alpha_k(x)\tau \gg 1$ for all k], then the Poisson variables can further be approximated by normal random variables, and Eq. (5) is approximated by

$$X(t + \tau) \approx x + \sum_k \mathcal{N}_k(\alpha_k(x)\tau, \alpha_k(x)\tau)\nu_k, \qquad (6)$$

where $\mathcal{N}_k(\mu_k, \sigma_k^2)$ are independent normal random variables with mean μ_k and variance σ_k^2. Reordering Eq. (6) yields

$$X(t + \tau) = x + \sum_k \alpha_k(x)\tau\nu_k + \sum_k \sqrt{\alpha_k(x)\tau} \mathcal{N}_k(0, 1)\nu_k,$$

which, in the form of an Itô stochastic differential equation, gives the *chemical Langevin equation* (CLE)

$$dX(t) = \sum_k \alpha_k(X(t))\nu_k \, dt + \sum_k \sqrt{\alpha_k(X(t))}\nu_k \, dW_k(t), \quad (7)$$

where $W_k(t)$ denotes a one-dimensional Wiener process (standard Brownian motion) for each k.[37] While the process defined by (3) or (4) runs within $\mathbb{N}_0^L$, the state space for the approximative process (7) extends to $\mathbb{R}^L$. Its integral notation

$$X(t) = X(0) + \sum_{k=1}^K \int_0^t \alpha_k(X(s))\nu_k \, ds$$
$$+ \sum_{k=1}^K \int_0^t \sqrt{\alpha_k(X(s))}\nu_k \, dW_k(s)$$

clearly exhibits the approximate character in comparison to (3).

In the thermodynamic limit where both the system's volume V and the volume-dependent molecule numbers $X_k^{(V)}$ approach infinity in a way that the concentrations $X_k^{(V)}/V$ stay constant, the noise terms in the chemical Langevin equation become negligible, leading to a deterministic system described by a set of ODE's. This is due to the fact that the propensity functions grow in direct proportion to the volume.[5–7,38] Actually, for a first-order reaction of species S_l,

$$\alpha_k(x) = \gamma_k x_l = V\gamma_k \frac{x_l}{V} = V\gamma_k c_l,$$

where $c_l := \frac{x_l}{V}$ is the concentration in moles per unit volume; for a second-order reaction of two different species S_l and $S_{l'}$,

$$\alpha_k(x) = \frac{\gamma_k}{V} x_l x_{l'} = V\gamma_k c_l c_{l'}.$$

Thus, by defining the volume-scaled propensities

$$\tilde{\alpha}_k(c) := \frac{\alpha_k(x)}{V}$$

for the vector of concentrations $c = (c_l)_{l=1,\ldots,L}$, multiplication of (7) by $\frac{1}{V}$ leads to

$$dC(t) = \sum_k \tilde{\alpha}_k(C(t))\nu_k \, dt + \sum_k \frac{1}{\sqrt{V}} \sqrt{\tilde{\alpha}_k(C(t))}\nu_k \, dW_k(t),$$

which, for $V \to \infty$, reduces to a simple system of ODE's for the concentrations

$$dC(t) = \sum_k \tilde{\alpha}_k(C(t))\nu_k \, dt,$$

see Refs. 9, 11, and 18 for details. A scalar transformation back to numbers rather than concentrations gives the *reaction-rate equation* (RRE),

$$dX(t) = \sum_k \alpha_k(X(t))\nu_k \, dt, \quad (8)$$

a deterministic equation characterizing the dynamics in the large population/large volume limit where the fluctuations become negligible.

Figure 1 shows trajectories of the dynamics characterized by the CME and its approximation equations for a simple system of binding and unbinding.

Formally, the approximation of Markov jump processes by Langevin dynamics or deterministic dynamics can be obtained by the Kramers-Moyal expansion:[39–41] In the first order, one gets the deterministic dynamics, and in the second order, one gets the Langevin dynamics. For linear reaction networks, it is well known that the first-order moments of all three approaches (CME, RRE, and CLE) match, that the CME and CLE match to the second moment, and that all approaches diverge at third order moments.[42]

Although the relation between the discrete stochastic model and its continuous approximations has been understood in every detail, the applicability is limited to systems with uniform population scaling. If only part of the system scales with the volume while some relevant species appear in low copy numbers independent of the volume, more sophisticated models are necessary. These hybrid models will be presented in Sec. III. To further motivate the relevance of multiscale molecular populations, we now present a common extension of the CME to systems with spatial resolution.

D. Spatial extension: The spatiotemporal CME

Modeling a reaction system by means of a CME (or by the approximative CLE or RRE) assumes the dynamics to be "well-mixed" in space which means that the diffusion of the molecules is fast compared to their reaction propensities. This implies that the spatial positions of the molecules are negligible, and the state of the system is fully described by

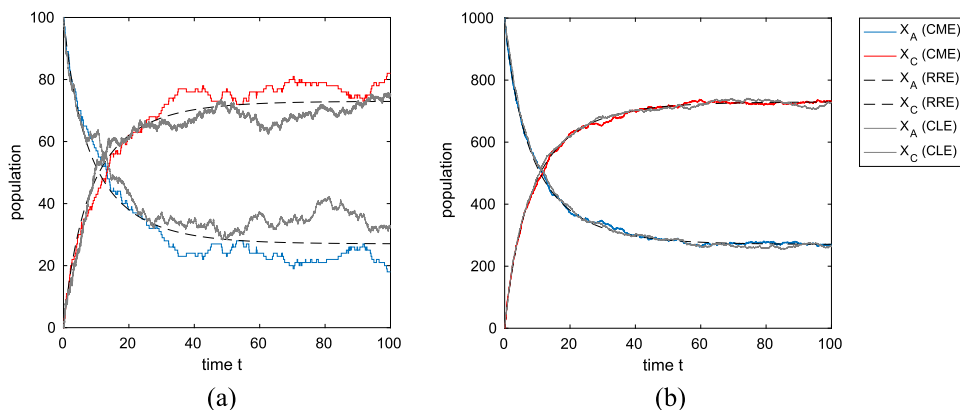FIG. 1. Gillespie simulation of the CME (red/blue), Euler simulation of the corresponding CLE (gray), and deterministic time-evolution determined by the RRE (black dashed) for a system of binding $A + B \to C$ and unbinding $C \to A + B$ of the chemical species $S_1 = A$, $S_2 = B$, $S_3 = C$. Initial states $X_C(0) = 0$ and $X_A(0) = X_B(0) = V$ with $V \in \{10^2, 10^3\}$. (a) $V = 10^2$ and (b) $V = 10^3$.

total numbers of molecules. If this well-mixed condition is broken, models with spatial resolution are required. The most detailed standard approach consists of individual molecule tracking where molecules are typically modeled as points or spheres undergoing spatially continuous Brownian motion, and bimolecular chemical reactions occur when reactive molecules pass within specified reaction-radii.[43–45] Alternatively, there exist approaches which discretize the space into a collection of non-overlapping compartments and approximate diffusion by jumps between the compartments. Interpreting the jumps as first-order reactions, this leads to an extended CME where the states have an additional spatial interpretation.[46–50] This second type of approach is considered in the following.

As in the setting of the CME, there are L species $S_1, \ldots, S_L$ which are governed by K reaction channels $R_1, \ldots, R_K$. Given a space discretization

$$\Omega = \Omega_1 \dot\cup \ldots \dot\cup \Omega_M$$

into $M \geq 2$ compartments, let $X_{lr}(t) \in \mathbb{N}_0$ denote the number of molecules of species S_l in compartment Ω_r at time $t \geq 0$. We use the notation $X_r(t) := (X_{1r}(t), \ldots, X_{Lr}(t))$ and identify the overall state of the system by the matrix $X(t) = (X_r(t))_{r=1,\ldots,M} \in \mathbb{N}_0^{L,M}$ containing the number of molecules of each species in each compartment. This state is affected both by chemical reactions and by diffusive transitions between the compartments.

Assuming that within each compartment the dynamics actually behave well-mixed, a local description by a CME is appropriate. Denoting again the net effect of reaction R_k onto the molecular population by $\nu_k = (\nu_{1k}, \ldots, \nu_{Lk}) \in \mathbb{Z}^L$, the emergence of this reaction in the rth compartment given the actual state $X(t) = x \in \mathbb{N}_0^{L,M}$ refers to the transition

$$x \rightarrow x + \nu_k \mathbf{1}_r,$$

where $\mathbf{1}_r \in \mathbb{N}^M$ is a row vector with the value 1 at entry r and zeros elsewhere. [$\nu_k \mathbf{1}_r$ is a $L \times M$-matrix whose rth column is equal to ν_k while all other columns contain zeros and simply indicate that $X(t)$ changes in compartment Ω_r.] The reaction propensities may now depend on the compartment and are given by functions $\alpha_k^r(x)$ which denote the probability per unit of time for reaction R_k to occur in compartment Ω_r given that $X_r(t) = x_r$, i.e., $\alpha_k^r(x)$ depends only on the values of x referring to compartment Ω_r.

These local reaction dynamics are combined with diffusive jumps of the molecules between the compartments. A jump of a molecule of species S_l from compartment Ω_r to compartment Ω_s, $s \neq r$, given the actual state $X(t) = x \in \mathbb{N}_0^{L,M}$ is described by the transition

$$x \rightarrow x + \mathbf{1}_s^l - \mathbf{1}_r^l,$$

where $\mathbf{1}_r^l$ is a matrix whose elements are all zero except the entry (l, r) which is one. Let λ_{rs}^l denote the rate for each individual molecule of species S_l to perform this jump. Since all molecules are assumed to diffuse independently of each other and jumps are treated as first-order reactions, the probability per unit of time for such a jump to occur at time t is $\lambda_{rs}^l X_{lr}(t)$.

As in the setting of the CME, let

$$P(x, t) = \mathbb{P}(X(t) = x | X(0) = x_0)$$

be the probability that the process is in state $x \in \mathbb{N}_0^{L,M}$ at time t given an initial state $X(0) = x_0$. Analogously to (4), the time evolution of $P(x, t)$ is now characterized by

$$
\begin{aligned}
\frac{dP(x, t)}{dt} = &\sum_{\substack{r,s=1 \\ r \neq s}}^{M} \sum_{l=1}^{L} \Big(\lambda_{sr}^l (x_{ls} + 1) P(x + \mathbf{1}_s^l - \mathbf{1}_r^l, t) \\
&- \lambda_{rs}^l x_{lr} P(x, t) \Big) + \sum_{r=1}^{M} \sum_{k=1}^{K} \Big(\alpha_k^r(x - \nu_k \mathbf{1}_r) \\
&\times P(x - \nu_k \mathbf{1}_r, t) - \alpha_k^r(x) P(x, t) \Big),
\end{aligned}
\tag{9}
$$

where the first line refers to the diffusive part while the second line describes the chemical reactions within the compartments. This system of ODE's is named *reaction-diffusion master equation* (RDME) or *spatiotemporal chemical master equation* (ST-CME)—depending on the underlying coarse-graining of space. The latter term indicates that the spatial discretization is not given by a regular Cartesian lattice but decomposes the state into areas of metastability, assuming that the considered reaction-diffusion process naturally exhibits a metastable behavior.[46] A typical example is given by reaction-diffusion systems within eukaryotic cells with a natural split-up into nucleus and cytoplasm; this setting will be considered in Sec. IV D.

No matter which type of coarse-graining is chosen, the main insight is that Eq. (9)–even if it looks more complex than the CME (4)—actually is just a CME with increased dimensionality. Both the simulation tools of Sec. II B and the approximation by chemical Langevin equations may directly be transferred to this spatial setting. The consideration of the thermodynamic limit to obtain deterministic dynamics, however, is based on a scaling by the volume which—in such a spatial consideration—gets a new and more concrete meaning. Relating the volume to the size of the spatial compartments means that an increasing volume affects the transition rates λ_{sr}^l between the compartments. The simple interpretation of jumps as first-order reactions with rates independent of the volume is not sustainable any more, and a general statement about how these rates scale with the volume is impossible. Here, a more specific analysis is required. On the other hand, the RRE as a limit model to describe the evolution of a reaction system also makes sense for systems with spatial interpretation. With regard to a fixed volume, the RRE simply provides an approximative description of the dynamics defined by the spatiotemporal chemical master equation as long as the considered compartments and their respective molecular populations are not too small.

The possible presence of multiple scales within the dynamics becomes even more evident in such a spatially inhomogeneous environment: While a huge molecular population in one compartment may permit approximations by Langevin dynamics or RRE's, a small population in another compartment may require the maintenance of the stochastic, discrete nature of a Markov jump process. Also the spatial inhomogeneity of reaction propensities easily leads to cascades of reaction speeds—asking for hybrid methods to handle multiscale dynamical systems.

III. HYBRID MODELS FOR MULTISCALE SYSTEMS

The approximation of the CME (4) by the RRE (8) requires all species to appear in high copy numbers. If some species remain in small copy numbers, an equal scaling of the whole system is inappropriate and the approach may fail. A typical example is given by the process of gene expression in single cells where the produced proteins may appear in large copy numbers while the template for transcription, the DNA, is present only in one or two copies. By describing the whole system in terms of RRE's, the stochastic nature of the system completely gets lost. Such situations require multiscale modeling approaches which take into account the different scaling levels. One prominent example of such multiscale approaches is given by piecewise deterministic Markov processes (PDMP's) which consider the large copy number limit for parts of the molecular population only while other parts are kept stochastic and discrete (the low copy number species). We give a short introduction to PDMP's, followed by an extension to hybrid diffusions where the high copy number species are described by CLE's rather than by RRE's.

A. Piecewise deterministic Markov processes

1. Special case: Piecewise deterministic reaction processes

As in Sec. II, we consider a system of L chemical species $S_1, \ldots, S_L$ undergoing K reaction channels $R_1, \ldots, R_K$ which are specified by their state-change vectors $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_K$. Given the system's volume $V > 0$, let $X_l^{(V)}(t) \in \mathbb{N}_0$ denote the number of molecules of species S_l at time $t \geq 0$. In contrast to the derivation of the RRE (8), we assume that only part of the molecular population scales with the system's volume, while some of the species appear in low copy numbers independent of V. Without loss of generality, let $S_1, \ldots, S_d$ for some $d < L$ be these low copy number species. For partially approximating the dynamics, only the components $X_{d+1}^{(V)}(t), \ldots, X_L^{(V)}(t)$ referring to the high abundant species are scaled by $1/V$ while the others are left unscaled, i.e., we consider the process $\boldsymbol{Y}^{(V)}(t) = \left(Y_l^{(V)}(t) \right)_{l=1,\ldots,L}$ defined by

$$Y_l^{(V)}(t) := \begin{cases} X_l^{(V)}(t), & l = 1, \ldots, d, \\ \frac{1}{V} X_l^{(V)}(t), & l = d+1, \ldots, L. \end{cases}$$

Analogously to the time-change representation (3) of the unscaled process, the dynamics of the partially scaled process $(\boldsymbol{Y}^{(V)}(t))$ are characterized by

$$\boldsymbol{Y}^{(V)}(t) = \boldsymbol{Y}^{(V)}(0) + \sum_{k=1}^{K} \mathcal{P}_k \left(\int_0^t \alpha_k^{(V)} \left(\boldsymbol{Y}^{(V)}(s) \right) ds \right) \boldsymbol{\nu}_k^{(V)} \quad (10)$$

for suitable partially scaled state-change vectors $\boldsymbol{\nu}_k^{(V)}$ defined by

$$\nu_{lk}^{(V)} := \begin{cases} \nu_{lk}, & l = 1, \ldots, d, \\ \frac{1}{V} \nu_{lk}, & l = d+1, \ldots, L \end{cases}$$

and adapted reaction propensity functions $\alpha_k^{(V)}$. As before, $\mathcal{P}_k$, $k = 1, \ldots, K$, denote independent, unit-rate Poisson processes.

Let now $\mathcal{K} \subset \{1, \ldots, K\}$ be the subset of those indices belonging to reactions which do *not* affect the low copy number

species, i.e.,

$$\mathcal{K} := \{k \in \{1, \ldots, K\} | \nu_{lk} = 0 \ \forall l = 1, \ldots, d\}.$$

Given this separation of reaction channels, we assume that there exist non-zero limit vectors $\tilde{\boldsymbol{\nu}}_k$ with

$$V \cdot \boldsymbol{\nu}_k^{(V)} \xrightarrow{V \to \infty} \tilde{\boldsymbol{\nu}}_k, \quad \text{for } k \in \mathcal{K} \quad (11)$$

and

$$\boldsymbol{\nu}_k^{(V)} \xrightarrow{V \to \infty} \tilde{\boldsymbol{\nu}}_k, \quad \text{for } k \notin \mathcal{K}. \quad (12)$$

This implies that the state-change vectors $\boldsymbol{\nu}_k$ do not scale with V such that, in the limit model, the dynamics of the high abundant species will be globally continuous. Generalizations of the model allow also jump in the high abundant species; however, in the interest of transparency, we stick to this more special setting. Assuming that also for the propensity functions there exist suitable limits $\tilde{\alpha}_k$ with $\frac{1}{V} \alpha_k^{(V)} \xrightarrow{V \to \infty} \tilde{\alpha}_k$ uniformly on compacts for $k \in \mathcal{K}$ and $\alpha_k^{(V)} \xrightarrow{V \to \infty} \tilde{\alpha}_k$ uniformly on compacts for $k \notin \mathcal{K}$, it has been shown in Ref. 15 that the sequence $\left(\boldsymbol{Y}^{(V)} \right)_{V \geq 1}$ of scaled Markov jump processes solving (10) converges to a stochastic process given by

$$\boldsymbol{Y}(t) = \boldsymbol{Y}(0) + \sum_{k \notin \mathcal{K}} \mathcal{P}_k \left(\int_0^t \tilde{\alpha}_k(\boldsymbol{Y}(s)) \, ds \right) \tilde{\boldsymbol{\nu}}_k + \int_0^t F(\boldsymbol{Y}(s)) \, ds,$$

$$(13)$$

in the sense of $\boldsymbol{Y}^{(V)}(t) \xrightarrow{V \to \infty} \boldsymbol{Y}(t)$ almost surely for each $t \geq 0$, with the vector field $F : \mathbb{R}^L \to \mathbb{R}^L$ given by

$$F(\boldsymbol{y}) := \sum_{k \in \mathcal{K}} \tilde{\alpha}_k(\boldsymbol{y}) \tilde{\boldsymbol{\nu}}_k. \quad (14)$$

That is, we have

$$\boldsymbol{Y}(t) = \boldsymbol{Y}(0) + \sum_{k \notin \mathcal{K}} \mathcal{P}_k \left(\int_0^t \tilde{\alpha}_k(\boldsymbol{Y}(s)) \, ds \right) \tilde{\boldsymbol{\nu}}_k$$

$$+ \sum_{k \in \mathcal{K}} \int_0^t \tilde{\alpha}_k(\boldsymbol{Y}(s)) \tilde{\boldsymbol{\nu}}_k \, ds \quad (15)$$

in comparison to (3) which clearly shows that it is a mixture of the jump process and deterministic dynamics: The process $\boldsymbol{Y}(t)$ exhibits a piecewise deterministic behavior; after starting at time $t = 0$ in $\boldsymbol{Y}(0)$, it follows a deterministic motion determined by the vector field F until t reaches the first jump time of any of the Poisson processes $\mathcal{P}_k$, $k \in \{1, \ldots, K\} \setminus \mathcal{K}$. Then the respective Poisson process jumps and the deterministic evolution restarts with a new initial state. Note that by definition of $\mathcal{K}$, it holds that $(F(\boldsymbol{y}))_l = 0$ for $l = 1, \ldots, d$ and any $\boldsymbol{y} \in \mathbb{R}^L$ which means that the deterministic flow is constrained to the components Y_l, $l = d+1, \ldots, L$, of large copy number species.

Equation (15) defines the piecewise deterministic reaction process as a hybrid Markov process using time-change representations. In another branch of the literature, the corresponding CME was discussed and derived, see, e.g., Refs. 16 and 18. It can be shown that the reduced CME associated with (15) allows for approximation of expectation values of the probability distribution governed by the full CME (4) up to an error of order $1/V$.[18] This implies that a statistically representative path ensemble of the process given by (15) would

allow for the same order of accuracy (if one can control the additional statistical error originating from the fact that one can only compute a finite sample of the path ensemble).

2. General case: Piecewise-deterministic Markov processes

The process $Y(t)$ characterized by (13) or (15) belongs to a general class of hybrid processes, called *piecewise deterministic Markov processes* (PDMP's), that are defined in a more general setting:[12,15,51] The state of a PDMP at some time $t \geq 0$ is given by a couple $Y(t) = (I(t), C(t))$ containing a discrete variable $I(t) \in \mathbb{N}_0^d$ and a continuous variable $C(t) \in \mathbb{R}^n$. As the name suggests, the process $Y(t)$ follows a deterministic motion which is interrupted by random jumps. The lengths of the time intervals between two successive jumps are random variables that will be called *waiting times* τ_j. The jump times $(t_j)_{j \in \mathbb{N}_0}$ are recursively defined by $t_0 = 0$ and $t_{j+1} = t_j + \tau_j$. Within each of the time intervals $[t_j, t_{j+1})$, the discrete variable remains constant, while the continuous variable evolves according to a given ODE. More precisely, consider the state space

$$\mathcal{S} = \mathbb{N}_0^d \times \mathbb{R}^n,$$

provided with the Borel σ-algebra $\mathcal{B}(\mathcal{S})$, and define the dynamics of the PDMP $Y(t) = (I(t), C(t))$ within $\mathcal{S}$ by the following characteristics.

1. The *deterministic flow*: For each $i \in \mathbb{N}_0^d$, there is a continuous vector field $f^{(i)} : \mathbb{R}^n \to \mathbb{R}^n$ defining the deterministic flow of the continuous variable $C(t)$ by the differential equation

$$dC(t) = f^{(i)}(C(t)) \, dt. \tag{16}$$

2. The *jump rate function* $\lambda : \mathcal{S} \to [0, \infty)$ which determines the distribution of waiting times of the process within a branch $\{i\} \times \mathbb{R}^n$ of the state space.

3. The *transition kernel* $Q : \mathcal{S} \times \mathcal{B}(\mathcal{S}) \to [0, 1]$ specifying the distribution of the process after a jump. We assume $Q(y, \{y\}) = 0$ for all $y \in \mathcal{S}$ which means that only "proper" jumps are considered as jumps.

With these ingredients, a sample path $(Y(t))_{t \geq 0} = (I(t), C(t))_{t \geq 0}$ of the PDMP, given an initial state $y_0 = (i_0, c_0) \in \mathcal{S}$ at time $t_0 = 0$, is recursively constructed as follows. Given the state $Y(t_j) = (i_j, c_j) \in \mathcal{S}$ of the process at time t_j, $j \in \mathbb{N}_0$, let $\Phi_{i_j}(c_j, t)$ be the solution of the initial value problem

$$d\Phi_{i_j}(c_j, t) = f^{(i_j)}(\Phi_{i_j}(c_j, t)) \, dt, \quad \Phi_{i_j}(c_j, 0) = c_j. \tag{17}$$

We assume this solution to exist and to be unique for $t \geq 0$. Define the next jump time by $t_{j+1} := t_j + \tau_j$, where the waiting time τ_j has the distribution

$$\mathbb{P}(\tau_j > t) = \exp\left(-\int_0^t \lambda(i_j, \Phi_{i_j}(c_j, s)) \, ds\right).$$

The deterministic evolution within the time interval $[t_j, t_{j+1})$ is given by

$$I(t_j + s) := i_j, \quad C(t_j + s) := \Phi_{i_j}(c_j, s) \quad \text{for } s \in [0, \tau_j).$$

Then, the post-jump state $Y(t_{j+1}) = (i_{j+1}, c_{j+1})$ at time $t_{j+1} = t_j + \tau_j$ is selected independently according to the distribution

$$\mathbb{P}(Y(t_{j+1}) \in A) = Q((i_j, \Phi_{i_j}(c_j, \tau_j)), A), \quad A \in \mathcal{B}(\mathcal{S}).$$

Thus, on the time interval $[t_j, t_{j+1}]$, the process $Y(t)$ is defined by

$$Y(t_j + s) = (I(t_j + s), C(t_j + s)) = \begin{cases} (i_j, \Phi_{i_j}(c_j, s)), & s \in [0, \tau_j), \\ (i_{j+1}, c_{j+1}), & s = \tau_j, \end{cases}$$

with the post-jump value c_{j+1} of the continuous component serving as an initial value for the next time interval.

At the jump times t_j, both the discrete and the continuous variables can instantaneously change their values by the transitions

$$i_j \to i_{j+1}, \quad \Phi_{i_j}(c_j, \tau_j) \to c_{j+1}.$$

If $c_{j+1} = \Phi_{i_j}(c_j, \tau_j)$ for all $j \in \mathbb{N}_0$, the trajectories of the continuous variable are globally continuous and the only effect of a jump is a change in the regime by a new value of the discrete variable. If, on the other hand, $c_{j+1} \neq \Phi_{i_j}(c_j, \tau_j)$, the trajectories of the continuous variable are only piecewise continuous.

3. Relation

The relation between the general construction and the special case given by (13) for reaction processes is as follows. The discrete process $I(t)$ is given by the first d (unscaled, integer-valued) components $(Y_1(t), \ldots, Y_d(t))$, while the continuous process is given by $C(t) = (Y_{d+1}(t), \ldots, Y_L(t)) \in \mathbb{R}^n$ with $n = L - d$. In short, we denote states as $y = (i, c)$. The vector field $f^{(i)}$ which works on the subspace of the continuous component is related to the vector field F via

$$F(y) = F((i, c)) = (0, f^{(i)}(c)).$$

The jump rate function λ is determined by the reaction propensities $\tilde{\alpha}$ via

$$\lambda(y) = \sum_{k \notin \mathcal{K}} \tilde{\alpha}_k(y), \tag{18}$$

and the transition kernel Q is given by

$$Q(y, \{y + \tilde{\nu}_k\}) = \frac{\tilde{\alpha}_k(y)}{\sum_{k' \notin \mathcal{K}} \tilde{\alpha}_{k'}(y)} \tag{19}$$

for each $k \in \{1, \ldots, K\} \setminus \mathcal{K}$. This relation connects the considerations for piecewise deterministic reaction processes to the more general literature on piecewise deterministic Markov processes and opens up the manifold of deep results on convergence and numerical approximation.

B. Hybrid diffusion

Between two consecutive jump times t_j and t_{j+1}, the continuous variable $C(t)$ of a PDMP is given by the deterministic flow $C(t) = \Phi_i(c, t)$ with $I(t_j) = i$ and $C(t_j) = c$ and satisfies the ODE $dC(t) = f^{(i)}(C(t)) \, dt$, see Eq. (16). In Ref. 26, Crudu *et al.* propose to consider an Itô stochastic differential equation instead of this ODE, thus adding a diffusion term to produce noise in the flow of the continuous variable, which leads to processes called *hybrid diffusion*. Similar to a PDMP, such a hybrid diffusion is given by a process $Y(t) = (I(t), C(t))$ which consists of a discrete component $I(t) \in \mathbb{N}_0^d$ and a continuous component $C(t) \in \mathbb{R}^n$. Both components have piecewise continuous trajectories interrupted by jumps at jump times t_j, $j \in \mathbb{N}_0$. Again, the jump dynamics are specified by a rate function λ and a transition kernel Q. Instead of the deterministic flow (16) between two jump times, however, the dynamics

of the continuous variable $C(t)$ are given by the stochastic differential equation

$$dC(t) = f^{(i)}(C(t)) dt + \xi^{(i)}(C(t)) dW(t), \qquad (20)$$

where $W(t)$ is an m-dimensional Wiener process, $f^{(i)} : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous vector field, and $\xi^{(i)} : \mathbb{R}^n \to \mathbb{R}^{n,m}$ is a diffusion matrix for each $i \in \mathbb{N}_0^d$.

Sampling the hybrid diffusion follows the same procedure as sampling the PDMP, with the deterministic flow (16) replaced by the randomized flow (20). That is, in between two jump times t_j, t_{j+1}, the process is defined by

$$I(t_j + s) := i_j, \quad C(t_j + s) := \Phi_{i_j}(c_j, s) \quad \text{for } s \in [0, \tau_j),$$

where $\Phi_{i_j}(c_j, s)$ is now the solution of

$$d\Phi_{i_j}(c_j, t) = f^{(i_j)}(\Phi_{i_j}(c_j, t)) dt + \xi^{(i_j)}(\Phi_{i_j}(c_j, t)) dW(t),$$
$$\Phi_{i_j}(c_j, 0) = c_j. \qquad (21)$$

Equation (21) is the analogue to (17) which defines the dynamics of the continuous variable in the piecewise deterministic setting.

In the context of a multiscale reaction system described by the special kind of PDMP (13), a nearby approach is to replace the deterministic flow of the species arising in high copy numbers by chemical Langevin dynamics as defined in (7). Again, we assume a two-scale separation of the species' abundance and subdivide the given reaction channels $R_1, \ldots, R_K$ according to their effect onto the classes of species, with the index subset $\mathcal{K} \subset \{1, \ldots, K\}$ separating those reactions which affect only high copy number species. Without loss of generality, we here assume that this subset $\mathcal{K}$ consists of the first m indices $1, \ldots, m$ ($m \leq K$), which will simplify the notation in the following description. As a direct analogue to the time-change representation (13) of the related piecewise deterministic reaction process, we define the *piecewise chemical Langevin equation* (PCLE)

$$Y(t) = Y(0) + \sum_{k \notin \mathcal{K}} \mathcal{P}_k \left(\int_0^t \tilde{\alpha}_k(Y(s)) ds \right) \tilde{\nu}_k$$
$$+ \int_0^t F(Y(s)) ds + \int_0^t \Xi(Y(s)) dW(s), \qquad (22)$$

with $F(y) := \sum_{k \in \mathcal{K}} \tilde{\alpha}_k(y) \tilde{\nu}_k$ as in (13), $W(s)$ denoting an m-dimensional Wiener process, and the diffusion matrix $\Xi : \mathbb{R}^L \to \mathbb{R}^{L,m}$ is given by

$$(\Xi(y))_{lk} := \sqrt{\tilde{\alpha}_k(y)} \tilde{\nu}_{lk}, \quad l = 1, \ldots, L, \ k = 1, \ldots, m. \qquad (23)$$

In between the random jump times which are determined by the Poisson processes $\mathcal{P}_k$, the dynamics of $Y(t)$ follow an Itô diffusion process given by the last two summands of (22). These two summands are just an integral version of a chemical Langevin equation like (7).

The process defined by the piecewise chemical Langevin equation (22) is a special type of hybrid diffusion. Just as in the setting of piecewise deterministic dynamics, the first d components of $Y(t)$ referring to low copy number species define the discrete variable $I(t)$, while the remaining components define the continuous variable $C(t)$. By definition of $\mathcal{K}$, the first d rows

of $\Xi(y)$ contain only zeros; the other rows form the diffusion matrix $\xi^{(i)}(c)$ for $y = (i, c)$, i.e., we have

$$\Xi(y) = \Xi((i, c)) = \begin{pmatrix} 0 \\ \xi^{(i)}(c) \end{pmatrix}.$$

The vector fields are again related by $F(y) = F((i, c)) = (0, f^{(i)}(c))$, and the jump rate function and the transition kernel are given by Eqs. (18) and (19), respectively. In contrast to the piecewise deterministic process, the dynamics defined by this piecewise chemical Langevin equation keep some randomness also for the species arising in high copy numbers.

By the noise term, both the CLE and the PCLE can induce negative values in the population, which of course is an unphysical prediction. This problem is addressed in Ref. 52. However, we stress that only the dynamics of medium or high abundant species are approximated by Langevin dynamics, such that negative values become extremely unlikely.

C. Joint equation

We note that hybrid diffusions contain piecewise deterministic Markov processes as a special case of the diffusion intensities $\xi^{(i)}$ vanishing to zero. This means that—without further work—we can also consider a hybrid model that mixes all three approaches: the description by discrete jump processes for low copy number species, chemical Langevin dynamics containing stochastic fluctuation for medium copy number species, and deterministic reaction kinetics for large copy number species. We arrive at the joint equation (1) that was announced in the Introduction. In direct analogy to the partition of reactions considered for the two-scale systems before, the index set $\mathcal{K}_{high}$ is given by those reactions whose effect is restricted to high copy number species (i.e., $\nu_{lk} = 0$ for l referring to a low or medium copy number species), while reactions in $\mathcal{K}_{medium}$ may affect both medium and high copy number species (but *not* low copy number species), and $\mathcal{K}_{low}$ contains the remaining reactions which affect at least one low copy number species.

The underlying idea is that the medium copy number species possess a volume scaling different from one of the high copy number species. For example, given that the high copy number species scale linear in V (as assumed in Secs. III A and III B), the medium copy number species could scale like $\sqrt{V}$. An adapted definition of the partially scaled process $Y^{(V)}$—in this specific case,

$$Y_l^{(V)}(t) := \begin{cases} X_l^{(V)}(t) & , \ l = 1, \ldots, d, \\ \frac{1}{\sqrt{V}} X_l^{(V)}(t), & l = d+1, \ldots, d', \\ \frac{1}{V} X_l^{(V)}(t) & , \ l = d'+1, \ldots, L, \end{cases}$$

where $S_1, \ldots, S_d$ are the low copy number species and $S_{d+1}, \ldots, S_{d'}$ are the medium copy number species for some $1 < d < d' < L$, and likewise for the corresponding state-change vectors $\nu_k^{(V)}$, it then again leads to suitable limit vectors $\tilde{\nu}_k$, in this case, given by

$$\tilde{\nu}_k := \begin{cases} \lim_{V \to \infty} V \cdot \nu_k^{(V)} & , \ k \in \mathcal{K}_{low}, \\ \lim_{V \to \infty} \sqrt{V} \cdot \nu_k^{(V)}, & k \in \mathcal{K}_{medium}, \\ \lim_{V \to \infty} \nu_k^{(V)} & , \ k \in \mathcal{K}_{high}. \end{cases}$$

By the multiple scaling, the limit state-change vectors $\tilde{\nu}_k$ of reactions $k \in \mathcal{K}_{medium}$ have zero entries for high copy number species (i.e., $\tilde{\nu}_{lk} = 0$ for $l = d' + 1, \ldots, L$, $k \in \mathcal{K}_{medium}$), such that the noise induced by the diffusion matrix Ξ is restricted to the medium copy number species, see its definition (23). The dynamics of the high copy number species are thus solely defined by the third line of (1), which is an ODE describing deterministic dynamics, but with coefficients that possibly depend not only on the discrete state of the low copy number species but also on the continuous stochastic flow of the medium copy number species.

The explicit form of the limit state-change vectors $\tilde{\nu}_k$ and the corresponding propensity functions $\tilde{\alpha}_k$ depends on the concrete scaling of the different classes of species, which in turn depends on the application at hand. The joint equation (1) is quite general in the sense that its concrete shape and the resulting dynamics depend on the chosen scaling. In particular, it comprises the two hybrid models of the PDMP and PCLE as special cases. All the hybrid models enable to reduce the complexity of a given reaction system while keeping its stochastic nature at a characteristic level. This is demonstrated in Sec. IV by an application to a multilevel gene expression system.

The asymptotic behavior of multiscale processes is investigated in Ref. 53, with several results on the weak convergence of these processes toward PDMP's. We note that in general the multiscale approaches will be able to reproduce the distribution of a multiscale process also in cases where a uniform rescaling and approximation by a RRE or CLE might fail to capture the process' characteristics (like in settings of multistability, see Ref. 54). Error bounds are given in Refs. 18 and 19, stating that the hybrid model of a PDMP approximates the marginal distributions of the low abundant species and the conditional moments of the high abundant species up to an error of order $1/V$—a statement which directly can be transferred to the joint equation (1) with three levels. The dependence of the approximation quality on the system's volume will be illustrated and analysed in the context of the following application.

IV. APPLICATION: AUTOREPRESSIVE GENETIC SYSTEM

A set of meaningful examples for multiscale reactive systems is given by gene regulatory networks.[55] While the template for transcription, the DNA, is present only in one or two copies which can be active or inactive, the synthesized messenger RNA (mRNA) and protein molecules may reach high population levels. There exist many network models for the interactions between the involved species and the regulation of gene activity (see, e.g., Ref. 26 for a set of more or less complex gene network models). In Refs. 13–15, an *autocatalytic* network of gene expression is investigated as an example for PDMP's. In such a network, the proteins regulate the transcription by a positive feedback law, i.e., the proteins activate the gene by binding to its binding sites. It has been shown that for such a system the proteins become extinct in finite time; and the point mass in zero is the only stationary distribution of the system.

However, many genes are controlled by *negative* self-regulating transcription factors, especially in prokaryotes like, e.g., *E. coli* organisms.[56] For such autorepressive genetic systems, gene expression typically occurs in bursts. Bursts can only appear when part of the system (at least the DNA) is treated as a discrete random variable. In the limit of describing the whole system by ODE's, the bursts naturally get lost (because the activity state of DNA is averaged). Bursting of gene expression can arise at the transcriptional or the translational level,[55] or the bursts are modeled as individual stochastic events by which a certain number of products are instantaneously introduced to the system.[57] We will here describe the transcriptional bursting where the gene switches between activity and inactivity.

A. Model

We consider the gene expression system depicted in Fig. 2. If the gene (DNA) is active, messenger RNA (mRNA) is transcribed at rate γ_1 and translated into proteins P at rate γ_2. Each protein can interrupt this stepwise production by repressing the gene and causing its inactivity at rate γ_3. The gene is activated again at a rate γ_4. Both within time periods of gene activity and within time periods of gene inactivity, mRNA is degraded at rate γ_5 and proteins are degraded at rate γ_6. All individual reactions are displayed in the reactions scheme (24) with DNA_0 denoting the repressed (inactive) gene,

$$
\begin{aligned}
DNA &\xrightarrow{\gamma_1} DNA + mRNA, \\
mRNA &\xrightarrow{\gamma_2} mRNA + P, \\
DNA + P &\xrightarrow{\gamma_3} DNA_0, \\
DNA_0 &\xrightarrow{\gamma_4} DNA + P, \\
mRNA &\xrightarrow{\gamma_5} \emptyset, \\
P &\xrightarrow{\gamma_6} \emptyset.
\end{aligned}
\tag{24}
$$
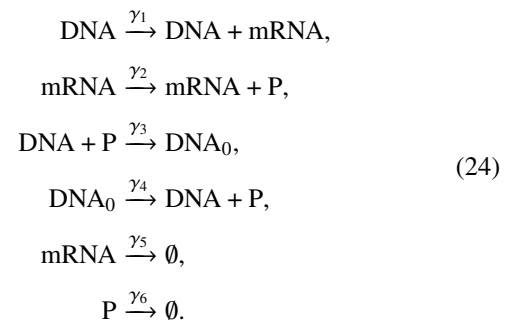
For the CME, we denote the state of the gene expression system at time t by $X(t) = (D(t), M(t), P(t)) \in \{0, 1\} \times \mathbb{N}_0^2$
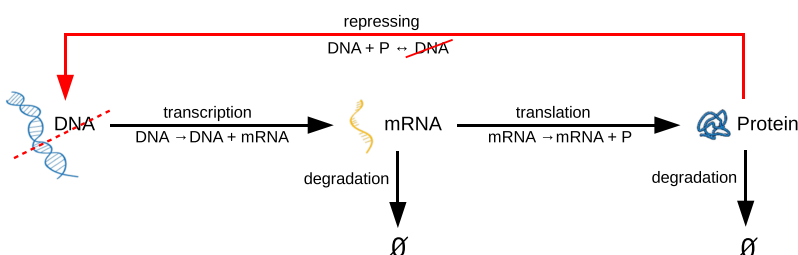

FIG. 2. Gene expression with negative self-regulation: By transcription, the active gene (DNA) produces messenger RNA (mRNA) which is translated into proteins. Both mRNA and proteins degrade with time. The proteins repress the production by deactivating the gene. Repressing is reversible: The gene is activated again at a constant rate.

with $D(t) \in \{0, 1\}$ referring to the number of active DNA and $M(t) \in \mathbb{N}_0$ and $P(t) \in \mathbb{N}_0$ giving the number of mRNA and proteins, respectively. The net change vectors of the reactions listed in (24) are given by

$$
\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}, \\
\mathbf{v}_4 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_5 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_6 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix},
$$
(25)

and the related propensity functions read as

$$
\alpha_1(\mathbf{x}) = V\gamma_1 x_1, \quad \alpha_2(\mathbf{x}) = \gamma_2 x_2, \quad \alpha_3(\mathbf{x}) = \frac{\gamma_3}{V} x_1 x_3, \\
\alpha_4(\mathbf{x}) = \gamma_4(1 - x_1), \quad \alpha_5(\mathbf{x}) = \gamma_5 x_2, \quad \alpha_6(\mathbf{x}) = \gamma_6 x_3.
$$
(26)

In this case, the population size of both mRNA and proteins scales with the volume V. In the hybrid models, DNA is therefore treated as the stochastic, discrete variable, and mRNA and proteins are treated as continuous variables. The state space of the PDMP is given by

$$
\mathcal{S} = \{(i, \mathbf{c}) : i \in \{0, 1\}, \mathbf{c} \in \mathbb{R}^2\}.
$$

For each $t \geq 0$, the state of the PDMP is denoted by $(I(t), \mathbf{C}(t)) \in \mathcal{S}$ with $I(t) \in \{0, 1\}$ defining the state of the DNA [$I(t) = 1$ referring to active DNA and $I(t) = 0$ referring to repressed DNA] and $\mathbf{C}(t) = (C_1(t), C_2(t)) \in \mathbb{R}^2$ defining the concentration of mRNA and proteins. The index set $\mathcal{K}$ of reactions that do not affect the low copy number "species" DNA is given by $\mathcal{K} = \{1, 2, 5, 6\}$. After the suitable partial scaling of the state-change vectors, given by $\mathbf{v}_1^{(V)} = \left(0, \frac{1}{V}, 0\right)$, $\mathbf{v}_2^{(V)} = \left(0, 0, \frac{1}{V}\right)$, $\mathbf{v}_3^{(V)} = \left(-1, 0, -\frac{1}{V}\right)$, $\mathbf{v}_4^{(V)} = \left(1, 0, \frac{1}{V}\right)$, $\mathbf{v}_5^{(V)} = \left(0, -\frac{1}{V}, 0\right)$, $\mathbf{v}_6^{(V)} = \left(0, 0, -\frac{1}{V}\right)$, the limit vectors defined by (11) and (12) read as

$$
\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \\
\mathbf{v}_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_5 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_6 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.
$$

The limit propensities depending on the state $\mathbf{y} = (i, (c_1, c_2))$ are given by

$$
\tilde{\alpha}_1(\mathbf{y}) = \gamma_1 i, \quad \tilde{\alpha}_2(\mathbf{y}) = \gamma_2 c_1, \quad \tilde{\alpha}_3(\mathbf{y}) = \gamma_3 i c_2,
$$

$$
\tilde{\alpha}_4(\mathbf{y}) = \gamma_4(1 - i), \quad \tilde{\alpha}_5(\mathbf{y}) = \gamma_5 c_1, \quad \tilde{\alpha}_6(\mathbf{y}) = \gamma_6 c_2.
$$

By definition (14), we obtain

$$
F(\mathbf{y}) = \gamma_1 i \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \gamma_2 c_1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \gamma_5 c_1 \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + \gamma_6 c_2 \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix},
$$

such that depending on the state $I(t) = i \in \{0, 1\}$ of the DNA, the deterministic flow for the mRNA and protein concentration is given by the vector field

$$
f^{(i)} : \mathbb{R}^2 \to \mathbb{R}^2, \quad f^{(i)}(\mathbf{c}) = \begin{pmatrix} \gamma_1 i \\ 0 \end{pmatrix} + \begin{pmatrix} -\gamma_5 & 0 \\ \gamma_2 & -\gamma_6 \end{pmatrix} \mathbf{c}.
$$
(27)

The rate function for jumps of the DNA takes the form

$$
\lambda(i, \mathbf{c}) = \gamma_3 i c_2 + \gamma_4(1 - i),
$$

and the transition kernel is deterministic in the sense of $Q((i, \mathbf{c}), ((1 - i), \mathbf{c})) = 1$ for $i \in \{0, 1\}, \mathbf{c} \in \mathbb{R}^2$. This means that for each time t, the stochastic repressing reaction takes place at rate $\gamma_3 I(t) C_2(t)$ while activation takes place at rate $\gamma_4(1 - I(t))$, giving

$$
\frac{d}{dt}\mathbb{P}(I(t) = 1) = \gamma_4 \mathbb{P}(I(t) = 0) - \gamma_3 C_2(t)\mathbb{P}(I(t) = 1)
$$

$$
= -\frac{d}{dt}\mathbb{P}(I(t) = 0).
$$

With $m = |\mathcal{K}| = 4$, the diffusion function of the corresponding piecewise chemical Langevin dynamics turns out to be

$$
\xi^{(i)}(\mathbf{c}) = \begin{pmatrix} \sqrt{\gamma_1 i} & 0 & -\sqrt{\gamma_5 c_1} & 0 \\ 0 & \sqrt{\gamma_2 c_1} & 0 & -\sqrt{\gamma_6 c_2} \end{pmatrix},
$$

meaning independent noise in the flows of mRNA and proteins. In this special case, the resulting flow equation (20) for the continuous component $\mathbf{C}(t)$ can be written as a coupled set of stochastic differential equations

$$
dC_1(t) = (\gamma_1 I(t) - \gamma_5 C_1(t)) \, dt + \sqrt{\gamma_1 I(t) + \gamma_5 C_1(t)} \, dW_1(t)
$$

for mRNA, and

$$
dC_2(t) = (\gamma_2 C_1(t) - \gamma_6 C_2(t)) \, dt + \sqrt{\gamma_2 C_1(t) + \gamma_6 C_2(t)} \, dW_2(t)
$$
(28)

for proteins, with independent Wiener processes $W_1(t)$ and $W_2(t)$. Here, we used the fact that given two independent Wiener processes $W_1(t)$ and $W_2(t)$, the weighted sum $\frac{aW_1(t) + bW_2(t)}{\sqrt{a^2 + b^2}}$ is another Wiener process.

B. Simulation results and model comparison

We compare the dynamics defined by the three models (CME, PDMP, and PCLE) for the self-regulated gene expression system with rates

$$
\begin{array}{lll}
\gamma_1 = 10^{-2}, & \gamma_2 = 0.5, & \gamma_3 = 0.1, \\
\gamma_4 = 10^{-2}, & \gamma_5 = 5 \cdot 10^{-3}, & \gamma_6 = 0.2
\end{array}
$$
(29)

and a large volume $V = 100$. These are artificial values without any claim to capture reality. Anyway, the parameter values in gene expression systems vary a lot and depend on the organism under consideration. The rates proposed here induce high population levels for both mRNA and proteins and comparatively long time periods of active DNA and are thereby well suited to demonstrate the relation between the three models under consideration—which is the main purpose of this section.

Sample paths of the CME are generated computationally by the stochastic simulation algorithm (see Sec. II B). For the PCLE and PDMP, the Euler-Maruyama method and Euler method are used, combined with a check-up for a switch in DNA in each iteration step. Trajectories of all three models are given in Fig. 3. One can observe comparatively long periods of time with active DNA (marked by gray areas) where both populations show a positive trend, while a repressed DNA (white areas) induces negative trends. The Gillespie simulation of the CME shows fluctuations around the trends which are well reproduced by the PCLE. The high population levels of mRNA and proteins induce long run times for the simulations of the CME; the approximations by the PDMP or PCLE significantly reduce the computational effort.
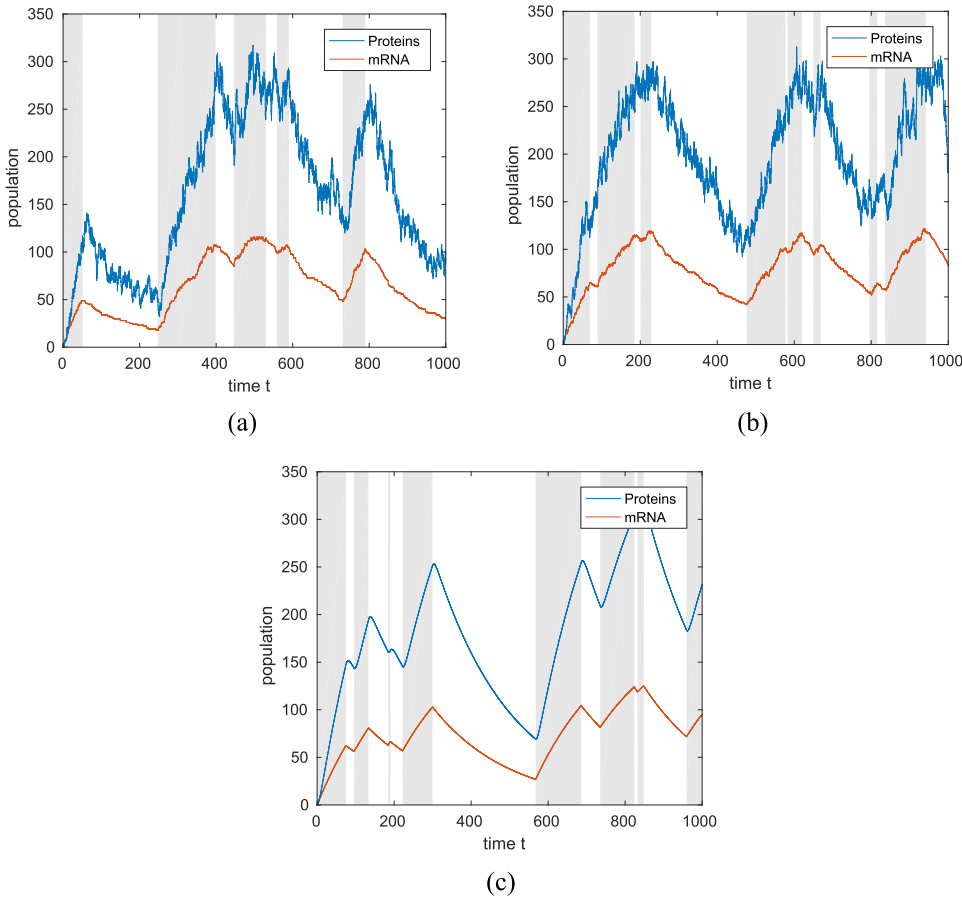
(a)

(b)

(c)

FIG. 3. Independent simulations of the CME, the PCLE, and the PDMP for parameter set (29) and initial state given by active DNA ($D(0) = I(0) = 1$) and absence of mRNA and proteins ($M(0) = C_1(0) = 0$, $P(0) = C_2(0) = 0$). The gray areas indicate the time periods of active DNA ($D(t) = I(t) = 1$). In the interest of comparability with the CME, we plot the abundance $V \cdot C_1(t)$ and $V \cdot C_2(t)$ for the PCLE and PDMP ($V = 100$). (a) CME, (b) PCLE, and (c) PDMP.

In order to reveal the differences in the approximation properties of the PDMP and PCLE we consider not only first and second order moments but also the distribution of the molecular populations and their local maxima—referring to protein bursts—which are typical for gene expression systems with negative self-regulation. What concretely do we consider as a burst? In the stochastic systems (CME and PCLE) "bursts" (large numbers of molecules) can appear by chance at any time. It is the piecewise deterministic model which clarifies what kind of bursts we are interested in: the maxima in the protein population after each individual repressing reaction, see Fig. 3(c). This motivates to define the burst size as the maximum number of protein molecules within a time interval of inactive DNA. Note that by this definition also small peaks in the molecular population are counted, namely, if repressing

takes place although the number of proteins is comparatively small (which can occur in all three models).

While the average long-term dynamics of all three models agree very well (see Table I), the size of the protein bursts is much better reproduced by the PCLE than by the PDMP (see Table II) which is due to the lack of noise in the PDMP.

Next, we compare the time-dependent evolution of the system starting with an initial state given by active DNA ($D(0) = 1$) and absence of mRNA and proteins ($M(0) = 0$, $P(0) = 0$). Based on Monte Carlo simulations (10^4 runs for each model), we draw histograms of the protein population at different points in time (Fig. 4) and calculate the time-dependent empirical averages of the three populations as an approximation of the first-order moments (Fig. 5).

TABLE I. Empirical averages ($\pm$ sample standard deviation) over long-term simulation ($T = 10^6$) for the gene expression system with parameter set (29). Lines 1-3: long-term averages of the DNA-process $D(t)$ (or $I(t)$), the mRNA-abundance $M(t)$ (or $V \cdot C_1(t)$), and the protein-abundance $P(t)$ (or $V \cdot C_2(t)$), respectively. Lines 4-5: mean length of time-intervals with active DNA ($D(t) = 1$) and time-intervals with inactive/repressed DNA ($D(t) = 0$). Size of the time step for the PCLE and PDMP: 0.1.

		CME	PCLE	PDMP
1	Mean DNA	0.34 ($\pm$0.47)	0.34 ($\pm$0.47)	0.34 ($\pm$0.47)
2	Mean mRNA	67.60 ($\pm$30.76)	67.51 ($\pm$31.20)	67.52 ($\pm$29.83)
3	Mean protein	169.03 ($\pm$77.64)	168.71 ($\pm$78.74)	168.80 ($\pm$74.41)
4	Mean active time	52.19 ($\pm$43.15)	51.37 ($\pm$43.16)	51.30 ($\pm$42.77)
5	Mean inactive time	101.46 ($\pm$101.49)	101.05 ($\pm$101.78)	100.63 ($\pm$100.01)

TABLE II. Protein bursts: results of long-term simulation ($T = 10^6$) for the gene expression system with parameter set (29).

	CME	PCLE	PDMP
Mean burst size	250.43	250.08	223.13
Maximum burst size	500	500.55	420.84
Minimum burst size	39	32.16	20.25

While the average dynamics agree very well for all the three approaches, the histogram of the PDMP reveals a clear deviation from the histograms of the other two approaches, see the spike at ~190 in Fig. 4(a). The reason for this deviation is the following. At time $t = 100$, there are about 35% of trajectories with "non-stop" gene activity (i.e., the DNA has never been repressed within the time interval [0, 100]). For the PDMP—due to its determinism within this period of time—all these trajectories are completely consistent on the time interval [0, 100]. Especially, the populations of mRNA and proteins agree with a value of ~190 proteins at time $t = 100$. For the CME and the PCLE, the proportion of non-stop activity at time $t = 100$ is the same, but the number of molecules varies for these non-stop activity trajectories due to randomness. The stochastic noise is able to create protein abundance up to 250 at time $t = 100$ which is simply impossible in the setting of the PDMP. As time passes, the deviation of the PDMP vanishes because the DNA-process $D(t)$ or $I(t)$ reaches equilibrium.

As for computational effort, the two hybrid methods outperform the Gillespie simulations by several orders of magnitude in the case of large volume systems. Simulations of the PCLE take longer than simulations of the PDMP because random noise terms have to be drawn in each iteration step. In return, the PCLE-simulations reproduce the variance in the dynamics. The three-level approach given by the joint equation (1) is the optimal combination of both hybrid approaches for situations of multiple levels. For illustration, consider the same genetic network as before with the propensity functions given in (26), but replace the propensity for transition by $\alpha_1(\boldsymbol{x}) = \sqrt{V}\gamma_1 x_1$ and the propensity for translation by $\alpha_2(\boldsymbol{x}) = \sqrt{V}\gamma_2 x_2$. This induces a mRNA-abundance of order $\sqrt{V}$ and a protein abundance of order V, see Fig. 6(a). The three-scale approach couples stochastic jumps of the DNA-state to Langevin dynamics of the medium abundant mRNA-population and deterministic dynamics of high abundant proteins. A simulation of the corresponding joint equation is given in Fig. 6(b). The dynamics could also be approximated by a PCLE or PDMP, but the joint three-level equation combines both of their advantages: The runtime is reduced in comparison to PCLE-simulations (because noise is only calculated for mRNA), and in contrast to the PDMP, there is stochastic noise in the mRNA-process which—by translation—also induces some local variance in the protein dynamics.

C. Small volume failure

The preceding investigations demonstrate that for high levels of mRNA and protein populations, the hybrid models deliver good approximations of the CME in terms of long-term averages and that the PCLE is also able to properly reproduce bursts and distributions of the number of molecules. In the case
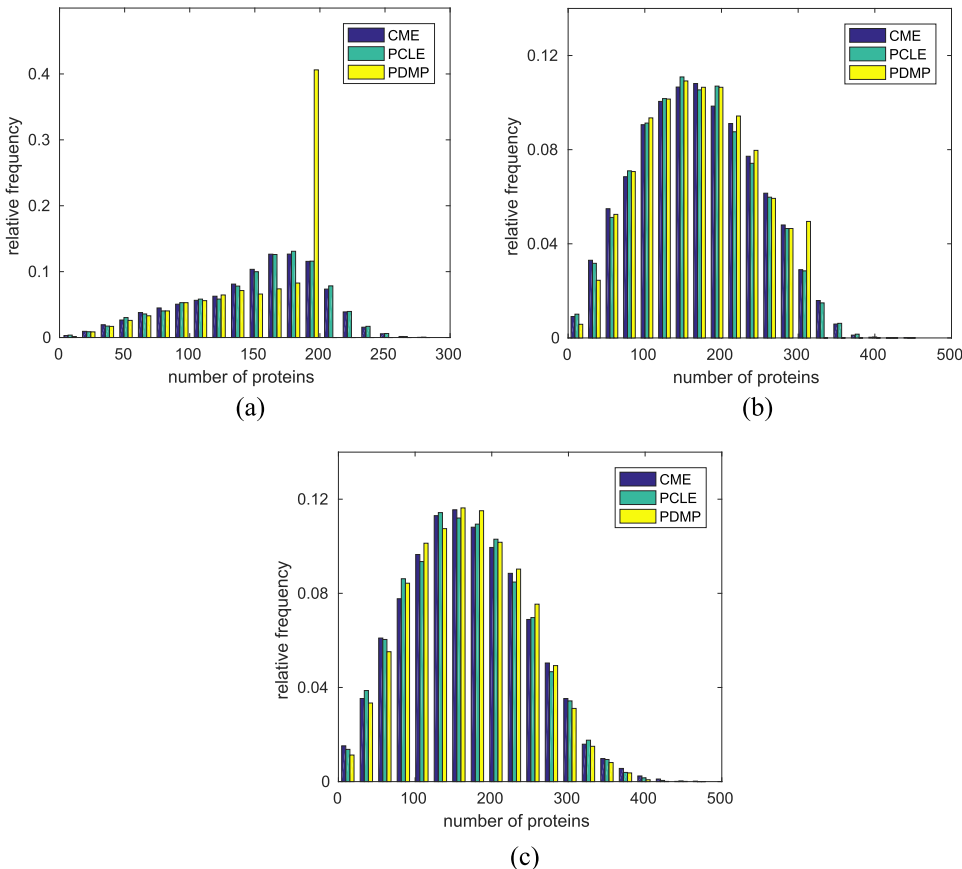

FIG. 4. Empirical distribution of proteins at times $t \in \{100, 200, 500\}$ taken from 10^4 Monte Carlo simulations of each model, all with initial states $D(0) = 1$, $M(0) = 0$, $P(0) = 0$ (or $I(0) = 1$, $C_1(0) = 0$, $C_2(0) = 0$, respectively). (a) $t = 100$, (b) $t = 200$, and (c) $t = 500$.
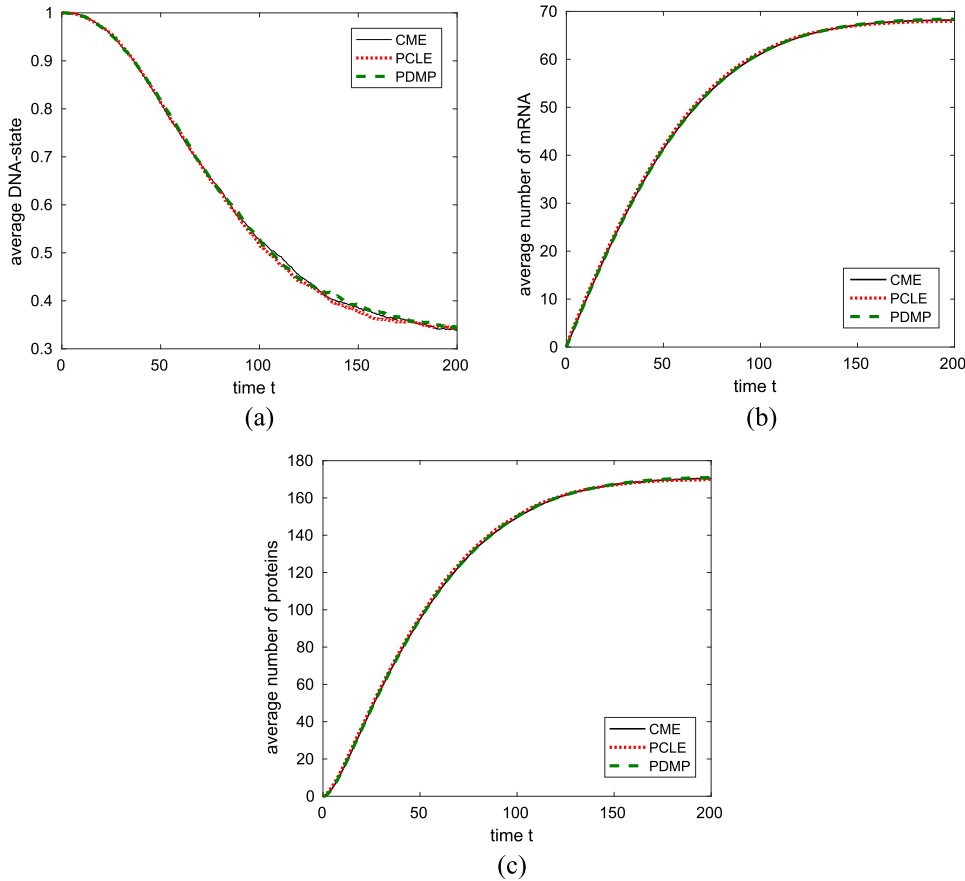
FIG. 5. Average dynamics for large volume $V = 100$. (a) Empirical mean of the DNA-state $D(t)$ or $I(t)$, respectively depending on time. This empirical mean also approximates the probability $\mathbb{P}(D(t) = 1)$ or $\mathbb{P}(I(t) = 1)$, respectively, which defines the marginal distribution of the DNA-state. [(b) and (c)] Empirical means of the number of mRNA and proteins depending on time. Results from Monte Carlo simulations with 10^4 runs for each model (CME, PCLE, and PDMP) given the initial states $D(0) = 1$, $M(0) = 0$, $P(0) = 0$ (or $I(0) = 1$, $C_1(0) = 0$, $C_2(0) = 0$, respectively). (a) DNA, (b) mRNA, and (c) proteins.

of parameter values which induce small population levels for mRNA and proteins, however, the approximation methods may fail. This is shown in Fig. 7 where the overall empirical averages of the populations disagree in the case of $V = 1$ [with propensity functions given by (26) and rate constants given by (29)]. Both the PDMP and PCLE show a decreased DNA-activity [see Fig. 7(a)] combined with lower levels of mRNA and protein abundance [see Figs. 7(b) and 7(c)]. On the other hand, the conditional expectations of the mRNA and protein population given a fixed value of the DNA-state are still consistent, see Fig. 8 where the evolution of the mean number of mRNA-molecules and proteins conditioned on $D(t) = 1$ (or $I(t) = 1$) $\forall t$ is shown. This consistency directly results from the fact that the conditional dynamics only contain first-order reactions. The only second-order reaction of the system (repressing) induces the switching between

the conditions and is ignored for the purpose of constant DNA-activity. Given the fixed DNA-state, the approximation of the mRNA and protein dynamics complies with a one-level approximation by a CLE or RRE (see Sec. II C) such that the corresponding approximation properties apply and first-order moments are perfectly reproduced for any volume.

Consequently, the divergence of the overall averages results from the fact that the marginal distributions of the DNA-state do not coincide, see Fig. 7(a). This deviation of the marginal distributions is caused by the second-order repressing reaction $DNA + P \rightarrow DNA_0$. The following short analysis clarifies this aspect.

Consider at first the PDMP with initial state $I(0) = 1$ (active DNA) and $\boldsymbol{C}(0) = (C_1(0), C_2(0)) = \boldsymbol{c}_0 \in \mathbb{R}_+^2$. Let T^{PDMP} denote the random time of first repression, i.e.,
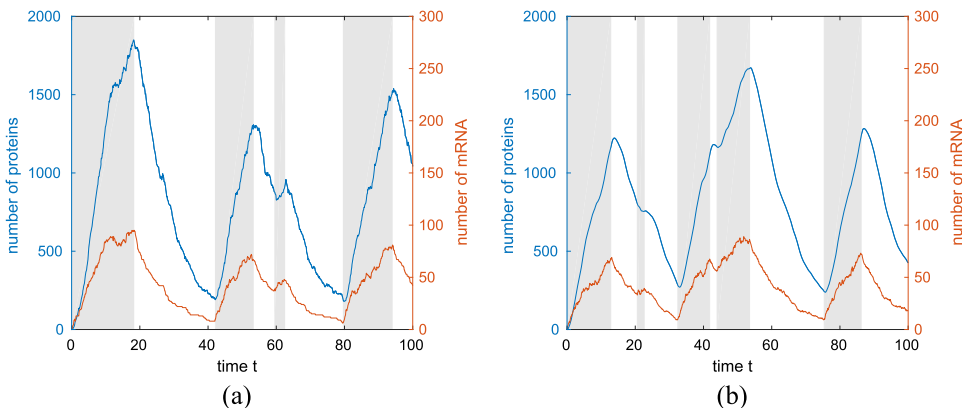

FIG. 6. Independent simulations of the CME and the joint equation (1) for scaled propensity functions $\alpha_1(\boldsymbol{x}) = \sqrt{V}\gamma_1 x_1$ and $\alpha_2(\boldsymbol{x}) = \sqrt{V}\gamma_2 x_2$ and $V = 100$. Parameter values: $\gamma_1 = 1$, $\gamma_2 = 1$, $\gamma_3 = 10^{-2}$, $\gamma_4 = 0.1$, $\gamma_5 = 0.1$, $\gamma_6 = 0.5$. The initial state given by active DNA and absence of mRNA and proteins. The gray areas indicate the time periods of active DNA. (a) CME and (b) joint equation.
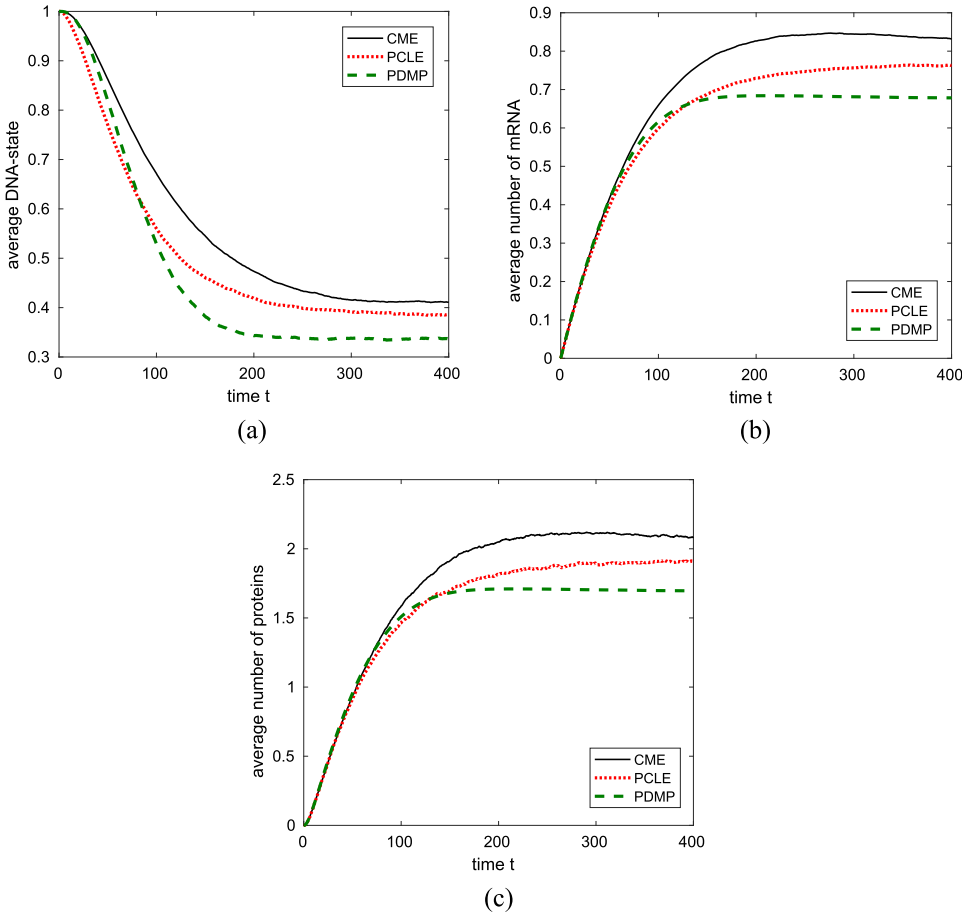
FIG. 7. Average dynamics for small volume $V = 1$. (a) Empirical mean of the DNA-state $D(t)$ or $I(t)$, respectively, depending on time. This empirical mean also approximates the probability $\mathbb{P}(D(t) = 1)$ or $\mathbb{P}(I(t) = 1)$, respectively, which defines the marginal distribution of the DNA-state. [(b) and (c)] Empirical means of the number of mRNA and proteins depending on time. Results from Monte Carlo simulations with 10^5 runs for each model (CME, PCLE, and PDMP) given the initial states $D(0) = 1$, $M(0) = 0$, $P(0) = 0$ (or $I(0) = 1$, $C_1(0) = 0$, $C_2(0) = 0$, respectively). (a) DNA, (b) mRNA, and (c) proteins.

$$T^{PDMP} := \inf\{t > 0 | I(t) = 0\}, \tag{30}$$

which will be called *repressing time* in the following. The repressing time T^{PDMP} is exponentially distributed with time dependent rates $\lambda_t^{PDMP} \geq 0$ given by $\lambda_t^{PDMP} := \gamma_3 c_2(t)$ where $c_2(t)$ is the deterministic flow of the protein concentration during the active period of time $[0, T^{PDMP}]$ given the initial state $\boldsymbol{c}_0$. More precisely, it holds that

$$\mathbb{P}\left(T^{PDMP} \leq t\right) = 1 - e^{-\Lambda_t^{PDMP}},$$

with $\Lambda_t^{PDMP} := \int_0^t \lambda_s^{PDMP} ds = \int_0^t \gamma_3 c_2(s) ds$. As for the PCLE, the rates for repression are not only time dependent but also

random, given by $\lambda_t^{PCLE} := \gamma_3 C_2(t)$ with $C_2(t)$ denoting the random flow of proteins defined in (28) given some initial state $I(0) = 1$, $\boldsymbol{C}(0) = \boldsymbol{c}_0 \in \mathbb{R}_+^2$. The corresponding distribution of the first repressing time T^{PCLE} [again defined by (30)] is given by

$$\mathbb{P}\left(T^{PCLE} \leq t\right) = 1 - \mathbb{E}\left(e^{-\Lambda_t^{PCLE}}\right),$$

with $\Lambda_t^{PCLE} := \int_0^t \lambda_s^{PCLE} ds = \int_0^t \gamma_3 C_2(s) ds$. Equivalently, for the CME, the repressing time $T^{CME} := \inf\{t > 0 | D(t) = 0\}$ has the distribution

$$\mathbb{P}\left(T^{CME} \leq t\right) = 1 - \mathbb{E}\left(e^{-\Lambda_t^{CME}}\right),$$
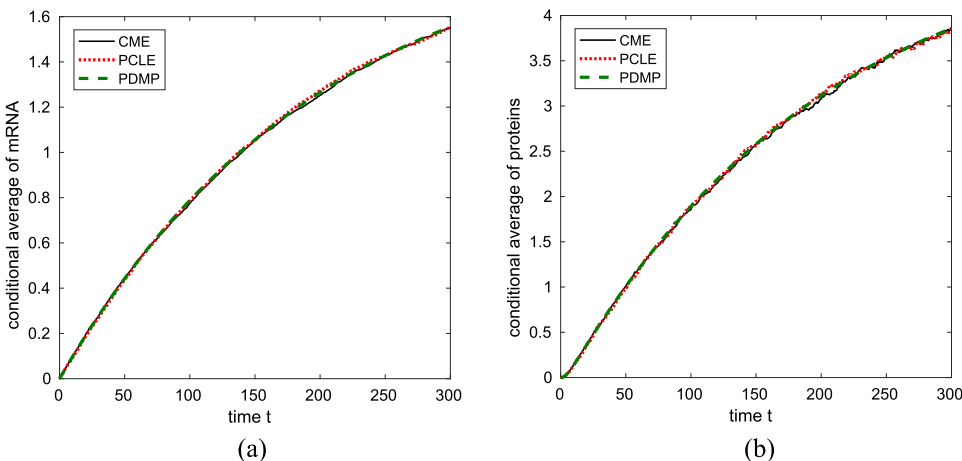

FIG. 8. *Conditional* average dynamics for small volume $V = 1$. [(a) and (b)] Empirical mean of mRNA and proteins depending on time *conditioned on* $D(t) = 1$ or $I(t) = 1$ for all t, respectively. Results from Monte Carlo simulations with 10^5 runs for each model (CME, PCLE, and PDMP) given the initial states $D(0) = 1$, $M(0) = 0$, $P(0) = 0$ (or $I(0) = 1$, $C_1(0) = 0$, $C_2(0) = 0$, respectively). (a) mRNA and (b) proteins.

with $\Lambda_t^{CME} := \int_0^t \lambda_s^{CME} \, ds$ and $\lambda_t^{CME} := \frac{\gamma_3}{V} P(t)$ where $P(t) \in \mathbb{N}_0$ gives the number of proteins at time t in the jump process defined by the CME, again with initially active DNA, i.e., $D(0) = 1$.

Now, although the first-order moments of the number of proteins agree for all three models, i.e., it holds that

$$c_2(t) = \mathbb{E}(C_2(t)) = \mathbb{E}\left(\frac{P(t)}{V}\right),$$

for all $t \in [0, T]$, the distributions of the repressing time diverge due to the general inequality

$$e^{-\Lambda_t^{PDMP}} \neq \mathbb{E}\left(e^{-\Lambda_t^{PCLE}}\right) \neq \mathbb{E}\left(e^{-\Lambda_t^{CME}}\right).$$

The fluctuations in the protein abundance have a non-linear impact on the repression times such that their zero averages are not carried over. Note that in the small volume setting, the PCLE is even likely to produce negative molecule numbers—which is not only inappropriate for interpretations of the underlying real-world application but also has a disproportional impact on the repressing propensities.

In summary, the different quality of fluctuations in the number of proteins during DNA-activity (no fluctuations for the PDMP, white noise for the PCLE, and Poisson-like variance for the CME) causes the deviations in the repressing times and with it the deviations of the marginal distributions of the DNA-state. Although, even in this small volume setting, the *conditional* dynamics within time periods of constant DNA agree for all three approaches with respect to the first-order moments, the switching times between these periods diverge—and with it the overall population averages. However, for large population levels of proteins, fluctuations become proportionally small and their impact vanishes—and with it the deviations in the repressing times.

D. Extension to eukaryotes

When considering the process of gene expression modeled in Fig. 2 within a eukaryotic cell containing a nucleus, the well-mixed assumption on the level of the overall cell is naturally broken. At least a minimal spatial differentiation is necessary because the reduced permeability of the nuclear membrane restricts the free flow through the cell and decomposes it into two metastable compartments: the nucleus and the cytoplasm. Moreover, some of the involved reactions [displayed in (24)] are restricted to either one of the compartments: jhile transcription takes place in the nucleus, translation typically arises within the cytoplasm. The time it takes first for the mRNA to leave the nucleus and second for the produced proteins to enter the nucleus induces a delayed repressing of the DNA. The protein population still shows bursts, but these bursts occur with time delay.

In order to demonstrate that the hybrid methods of Sec. III can directly be applied to the spatial extensions of the CME presented in Sec. II D, we consider the process of gene expression in a eukaryotic cell and approximate parts of the system by chemical Langevin dynamics or RRE's. In contrast to Sec. IV B where we intended to uncover the differences between the two approximation methods, we now choose more realistic parameter values based on the data given in Refs. 56 and 58. Especially we take into account that the number of proteins typically exceeds the number of mRNA by several orders of magnitude.

Based on the split-up into two compartments, the state of the ST-CME is given by the matrix $X(t) \in \mathbb{N}_0^{3,2}$ with the entry $X_{lr}(t)$ denoting the number of molecules of species l ($l = 1$: DNA, $l = 2$: mRNA, $l = 3$: proteins) in compartment r ($r = 1$: nucleus, $r = 2$: cytoplasm). Although it holds that $X_{11}(t) \in \{0, 1\}$ and $X_{12}(t) = 0$ for all t (because there is at most one active DNA which is located in the nucleus), we use this larger state space in order to stick to the general notation of Sec. II D. The net change vectors of the reactions are as in (25), and the compartment-dependent reaction propensities are

$$\alpha_1^r(\boldsymbol{x}) = \begin{cases} \gamma_1 x_{11}, & r = 1, \\ 0, & r = 2; \end{cases} \quad \alpha_2^r(\boldsymbol{x}) = \begin{cases} 0, & r = 1, \\ \gamma_2 x_{22}, & r = 2; \end{cases} \quad \alpha_3^r(\boldsymbol{x}) = \begin{cases} \frac{\gamma_3}{V_1} x_{11} x_{21}, & r = 1, \\ 0, & r = 2; \end{cases}$$

$$\alpha_4^r(\boldsymbol{x}) = \begin{cases} \gamma_4(1 - x_{11}), & r = 1, \\ 0, & r = 2; \end{cases} \quad \alpha_5^r(\boldsymbol{x}) = \begin{cases} \gamma_5 x_{21}, & r = 1, \\ \gamma_5 x_{22}, & r = 2; \end{cases} \quad \alpha_6^r(\boldsymbol{x}) = \begin{cases} \gamma_6 x_{31}, & r = 1, \\ \gamma_6 x_{32}, & r = 2, \end{cases}$$

where V_1 is the volume of the nucleus set to $V_1 = 1$ and the rate constants are

$$\begin{aligned} \gamma_1 &= 0.1, \quad \gamma_2 = 0.1, \quad \gamma_3 = 10^{-2}, \\ \gamma_4 &= 10^{-2}, \quad \gamma_5 = 5 \cdot 10^{-3}, \quad \gamma_6 = 2 \cdot 10^{-4}. \end{aligned} \quad (31)$$

As for the transition rates, we assume that nuclear mRNA-molecules enter the cytoplasm at rate $\lambda_{12}^2 = 10^{-3}$, while their return to the nucleus is precluded ($\lambda_{21}^2 = 0$). Vice versa, proteins switch from cytoplasm to nucleus at rate $\lambda_{21}^3 = 10^{-3}$ but

not back ($\lambda_{12}^3 = 0$). For DNA, we naturally have $\lambda_{12}^1 = \lambda_{21}^1 = 0$. We assume these transition rates to be independent of the volume.

Simulations of the resulting ST-CME reveal that for the chosen rates, the number of mRNA-molecules in both compartments is much smaller than the number of proteins. This suggests keeping the discrete, stochastic nature for mRNA while approximating only the proteins by continuous dynamics. The state space of the PDMP and the hybrid diffusion is
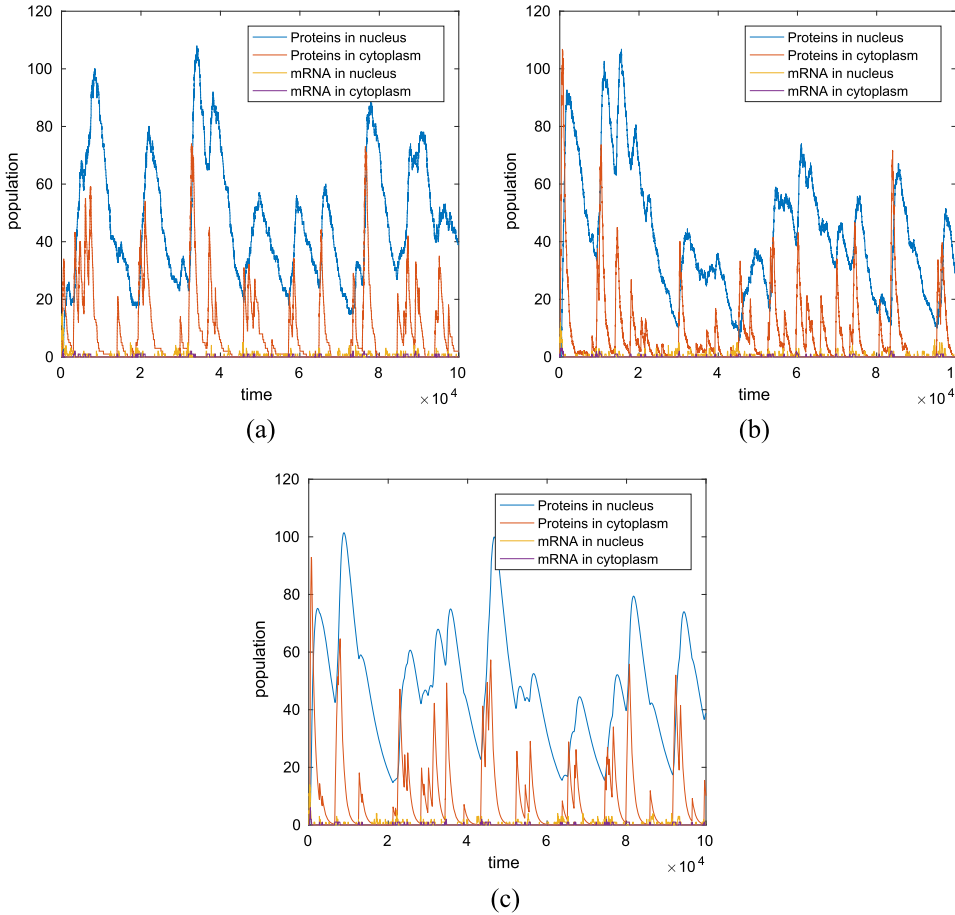
(a)

(b)

(c)

FIG. 9. Independent simulations of the ST-CME, the spatial PCLE, and the spatial PDMP for parameter set given in (31) and initial state given by active DNA and absence of mRNA and proteins. (a) ST-CME, (b) PCLE, and (c) PDMP.

given by

$$\mathcal{S} = \{(\boldsymbol{i}, \boldsymbol{c}) : \boldsymbol{i} \in \mathbb{N}_0^{2,2}, \boldsymbol{c} \in \mathbb{R}^2\},$$

with $i_{11} \in \{0, 1\}$ denoting the state of DNA, $i_{21} \in \mathbb{N}_0$ and $i_{22} \in \mathbb{N}_0$ denoting the number of mRNA-molecules in nucleus and cytoplasm, respectively, and $\boldsymbol{c} \in \mathbb{R}^2$ describing the protein concentration in nucleus and cytoplasm. (It is $i_{12} = 0$ as this refers to DNA in the cytoplasm.)

The resulting index set of reactions keeping the discrete species unchanged then reduces to $\mathcal{K} = \{2, 6\}$. By interpreting again the jumps as first-order reactions, also the jumps of proteins belong to this group of reactions.

This time, the deterministic flow of the two-component protein concentration $\boldsymbol{c} \in \mathbb{R}^2$ depends on the number $i_{22} \in \mathbb{N}_0$

of mRNA-molecules in the cytoplasm and is given by the vector field

$$f^{(\boldsymbol{i})} : \mathbb{R}^2 \to \mathbb{R}^2, \quad f^{(\boldsymbol{i})}(\boldsymbol{c}) = \begin{pmatrix} 0 \\ \gamma_2 i_{22} \end{pmatrix} + \begin{pmatrix} -\gamma_6 & \lambda_{21}^3 \\ 0 & -\gamma_6 - \lambda_{21}^3 \end{pmatrix} \boldsymbol{c}.$$

The jump rate function is given by

$$\lambda(\boldsymbol{i}, \boldsymbol{c}) = \gamma_1 i_{11} + \gamma_3 i_{11} c_1 + \gamma_4 (1 - i_{11}) + \gamma_5 i_{21} + \gamma_5 i_{22} + \lambda_{12}^2 i_{21}$$

and the transition kernel is determined by the respective fractions.

Individual simulations of the jump process defined by the ST-CME and the approximative PCLE and PDMP with spatial split-up are given in Fig. 9. In contrast to the dynamics

TABLE III. Empirical averages (± sample standard deviation) over long-term simulations ($T = 10^7$) for the gene expression system in eukaryotes with parameter set given in (31). Lines 1-5: long-term averages of the DNA-process $X_{11}(t)$ (or $I_{11}(t)$), the mRNA-abundance $X_{2r}(t)$ (or $I_{2r}(t)$) in both compartments, and the protein-abundance $X_{3r}(t)$ (or $V \cdot C_r(t)$) in both compartments, respectively. Lines 6-7: mean length of time-intervals with active DNA ($X_{11}(t) = 1$ or $I_{11}(t) = 1$, respectively) and time-intervals with inactive/repressed DNA ($X_{11}(t) = 0$ or $I_{11}(t) = 0$, respectively). Step size for the PCLE and PDMP: 0.1.

		ST-CME	Spatial PCLE	Spatial PDMP
1	Mean DNA	0.03 (±0.17)	0.03 (±0.17)	0.03 (±0.17)
2	Mean mRNA (nucleus)	0.49 (±0.84)	0.49 (±0.84)	0.49 (±0.84)
3	Mean mRNA (cytoplasm)	0.10 (±0.32)	0.10 (±0.32)	0.10 (±0.32)
4	Mean proteins (nucleus)	40.21 (±19.31)	40.30 (±19.34)	40.89 (±19.07)
5	Mean proteins (cytoplasm)	8.04 (±11.91)	8.09 (±11.83)	8.15 (±11.68)
6	Mean active time	2.99 (±3.69)	3.14 (±3.98)	3.05 (±3.76)
7	Mean inactive time	100.05 (±100.16)	99.88 (±100.15)	100.55 (±100.22)

considered in Sec. IV B, the time periods of active DNA are very short, such that an illustration by gray background areas is inappropriate. Instead, the trends of the protein populations within both compartments are determined by the actual random and discrete number of mRNA-molecules. A comparison of the statistics in Table III shows again a very good agreement of all three modeling approaches with respect to long-term averages.

V. CONCLUSION

The CME is of crucial importance for accurately modeling discrete stochastic reaction kinetics. However, with increasing dimension, its complexity renders existing solution methods infeasible. This motivates the construction of reduced systems for approximating the dynamics. We presented several hybrid approaches for modeling multiscale reaction dynamics based on rescaling methods. We considered hybrid diffusions as an extension of piecewise deterministic Markov processes and formulated both models within the specific framework of two-scale reaction dynamics. Moreover, we showed how theses approaches can be combined to a generalized joint model for systems with more than two population scales. In order to complete the model overview, we also mentioned spatial extensions of the CME which are suitable for systems that do not fulfil the central well-mixed condition. The main insight is that each of the approximation methods can as well be transferred to this more complex setting of spatiotemporal dynamics.

We applied some of the hybrid approaches to a multiscale genetic network, considering both a well-mixed system within a prokaryotic cell and a two-compartment system within a eukaryotic cell described by the spatiotemporal CME. Due to a negative feedback law, the produced proteins repress their own transcription by deactivating the DNA, such that the overall system switches between time periods of active DNA with positive trends for the gene products and time periods of inactive DNA with negative trends for the gene products. In contrast to uniform approximation methods, the hybrid approaches are able to maintain this characteristic behavior. For the case of large copy numbers of the products, a comparison of simulation statistics revealed good approximation properties of the hybrid models in terms of first-order moments. By the additional noise of the piecewise chemical Langevin dynamics, also protein bursts and empirical distributions have accurately been reproduced. The joint equation (1) combines the two-scale hybrid models to an ideal framework for three-level systems. We showed that in the small volume setting, the hybrid models under consideration fail by distorting the repression event: although the conditional dynamics within the time periods of active or inactive DNA are still congruent with respect to averages, the switching times between them follow different distributions.

This paper gives a coherent overview of several modeling approaches for stochastic reaction kinetics and reveals their interrelation as well as possible recombinations. We only considered time-homogeneous systems where the classification of species according to their copy number levels is fixed for all times. In many applications, however, also switches between different levels—induced by the occurrence of reactions—are plausible. For the future, an investigation of flexible, time-adaptive model combinations, as proposed, e.g., in Ref. 59, will be of special interest.

[1] W. J. Blake, M. Kærn, C. R. Cantor, and J. J. Collins, "Noise in eukaryotic gene expression," Nature **422**(6932), 633–637 (2003).

[2] R. Bundschuh, F. Hayot, and C. Jayaprakash, "Fluctuations and slow variables in genetic networks," Biophys. J. **84**(3), 1606–1615 (2003).

[3] T. B. Kepler and T. C. Elston, "Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations," Biophys. J. **81**(6), 3116–3136 (2001).

[4] B. B. Kaufmann and A. van Oudenaarden, "Stochastic gene expression: From single molecules to the proteome," Curr. Opin. Genet. Dev. **17**(2), 107–112 (2007).

[5] D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," J. Comput. Phys. **22**(4), 403–434 (1976).

[6] D. T. Gillespie, "A rigorous derivation of the chemical master equation," Phys. A **188**(1), 404–425 (1992).

[7] D. T. Gillespie, "Stochastic simulation of chemical kinetics," Annu. Rev. Phys. Chem. **58**, 35–55 (2007).

[8] D. T. Gillespie, "Approximate accelerated stochastic simulation of chemically reacting systems," J. Chem. Phys. **115**(4), 1716–1733 (2001).

[9] T. G. Kurtz, "Solutions of ordinary differential equations as limits of pure jump Markov processes," J. Appl. Probab. **7**(1), 49–58 (1970).

[10] T. G. Kurtz, "Limit theorems for sequences of jump Markov processes approximating ordinary differential processes," J. Appl. Probab. **8**(2), 344–356 (1971).

[11] T. G. Kurtz, "The relationship between stochastic and deterministic models for chemical reactions," J. Chem. Phys. **57**(7), 2976–2978 (1972).

[12] M. H. A. Davis, "Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models," J. R. Stat. Soc. B **46**(3), 353–388 (1984).

[13] S. Zeiser, U. Franz, and V. Liebscher, "Autocatalytic genetic networks modeled by piecewise-deterministic Markov processes," J. Math. Biol. **60**(2), 207–246 (2010).

[14] S. Zeiser, U. Franz, O. Wittich, and V. Liebscher, "Simulation of genetic networks modelled by piecewise deterministic Markov processes," IET Syst. Biol. **2**(3), 113 (2008).

[15] U. Franz, V. Liebscher, and S. Zeiser, "Piecewise-deterministic Markov processes as limits of Markov jump processes," Adv. Appl. Probab. **44**(03), 729–748 (2012).

[16] T. Jahnke, "On reduced models for the chemical master equation," SIAM Multiscale Model. Simul. **9**(4), 1646–1676 (2011).

[17] S. Smith, C. Cianci, and R. Grima, "Model reduction for stochastic chemical systems with abundant species," J. Chem. Phys. **143**(21), 214105 (2015).

[18] S. Menz, J. Latorre, C. Schütte, and W. Huisinga, "Hybrid stochastic-deterministic solution of the chemical master equation," SIAM Multiscale Model. Simul. **10**(4), 1232–1262 (2012).

[19] T. Jahnke and M. Kreim, "Error bound for piecewise deterministic processes modeling stochastic reaction systems," SIAM Multiscale Model. Simul. **10**(4), 1119–1147 (2012).

[20] S. L. Cotter, K. C. Zygalakis, I. G. Kevrekidis, and R. Erban, "A constrained approach to multiscale stochastic simulation of chemically reacting systems," J. Chem. Phys. **135**(9), 094102 (2011).

[21] E. L. Haseltine and J. B. Rawlings, "Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics," J. Chem. Phys. **117**(15), 6959–6969 (2002).

[22] K. Ball, T. G. Kurtz, L. Popovic, G. Rempala *et al.*, "Asymptotic analysis of multiscale approximations to reaction networks," Ann. Appl. Probab. **16**(4), 1925–1961 (2006).

[23] A. Alfonsi, E. Cancès, G. Turinici, B. Di Ventura, and W. Huisinga, "Adaptive simulation of hybrid stochastic and deterministic models for biochemical systems," in *ESAIM: Proceedings* (EDP Sciences, 2005), Vol. 14, pp. 1–13.

[24]A. Alfonsi, E. Cances, G. Turinici, B. Di Ventura, and W. Huisinga, "Exact simulation of hybrid stochastic and deterministic models for biochemical systems," Ph.D. thesis, INRIA, 2004.

[25]H. Salis and Y. Kaznessis, "Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions," J. Chem. Phys. **122**(5), 054103 (2005).

[26]A. Crudu, A. Debussche, and O. Radulescu, "Hybrid stochastic simplifications for multiscale gene networks," BMC Syst. Biol. **3**(1), 89 (2009).

[27]M. Ghosh and A. Bagchi, "Modeling stochastic hybrid systems," in *System Modeling and Optimization* (Springer, 2005), pp. 269–280.

[28]G. Pola, M. L. Bujorianu, J. Lygeros, and M. D. Di Benedetto, "Stochastic hybrid models: An overview," in *Proceedings of the IFAC Conference on Analysis and Design of Hybrid Systems* (Elsevier, 2003), pp. 45–50.

[29]A. N. Gorban and I. V. Karlin, *Invariant Manifolds for Physical and Chemical Kinetics* (Springer, 2004).

[30]D. F. Anderson and T. G. Kurtz, "Continuous time Markov chain models for chemical reaction networks," in *Design and Analysis of Biomolecular Circuits* (Springer, 2011), pp. 3–42.

[31]N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, 1981).

[32]Y. Hu, T. Li, and B. Min, "A weak second order tau-leaping method for chemical kinetic systems," J. Chem. Phys. **135**(2), 024113 (2011).

[33]M. A. Gibson and J. Bruck, "Efficient exact stochastic simulation of chemical systems with many species and many channels," J. Phys. Chem. A **104**(9), 1876–1889 (2000).

[34]Y. Cao, H. Li, and L. Petzold, "Efficient formulation of the stochastic simulation algorithm for chemically reacting systems," J. Chem. Phys. **121**(9), 4059–4067 (2004).

[35]L. Lok and R. Brent, "Automatic generation of cellular reaction networks with moleculizer 1.0," Nat. Biotechnol. **23**(1), 131–136 (2005).

[36]D. F. Anderson, A. Ganguly, and T. G. Kurtz, "Error analysis of tau-leap simulation methods," Ann. Appl. Probab. **21**(6), 2226–2262 (2011).

[37]D. T. Gillespie, "The chemical Langevin equation," J. Chem. Phys. **113**(1), 297–306 (2000).

[38]D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," J. Phys. Chem. **81**(25), 2340–2361 (1977).

[39]H. A. Kramers, "Brownian motion in a field of force and the diffusion model of chemical reactions," Physica **7**(4), 284–304 (1940).

[40]J. E. Moyal, "Stochastic processes and statistical physics," J. R. Stat. Soc. B **11**(2), 150–210 (1949).

[41]H. Risken, "Fokker-Planck equation," in *The Fokker-Planck Equation* (Springer, 1984), pp. 63–95.

[42]A. Ghosh, A. Leier, and T. T. Marquez-Lago, "The spatial chemical Langevin equation and reaction diffusion master equations: Moments and qualitative solutions," Theor. Biol. Med. Modell. **12**(1), 5 (2015).

[43]S. Andrews and D. Bray, "Stochastic simulation of chemical reactions with spatial resolution and single molecule detail," Phys. Biol. **1**(3), 137 (2004).

[44]J. Schöneberg and F. Noé, "ReaDDy—A software for particle-based reaction-diffusion dynamics in crowded cellular environments," PLoS One **8**(9), e74261 (2013).

[45]J. S. van Zon and P. R. ten Wolde, "Simulating biochemical networks at the particle level and in time and space: Green's function reaction dynamics," Phys. Rev. Lett. **94**, 128103 (2005).

[46]S. Winkelmann and C. Schütte, "The spatiotemporal master equation: Approximation of reaction-diffusion dynamics via Markov state modeling," J. Chem. Phys. **145**(21), 214107 (2016).

[47]R. Erban and S. J. Chapman, "Stochastic modelling of reaction-diffusion processes: Algorithms for bimolecular reactions," Phys. Biol. **6**(4), 046001 (2009).

[48]S. Hellander, A. Hellander, and L. Petzold, "Reaction rates for mesoscopic reaction-diffusion kinetics," Phys. Rev. E **91**(2), 023312 (2015).

[49]S. A. Isaacson, "The reaction-diffusion master equation as an asymptotic approximation of diffusion to a small target," SIAM J. Appl. Math. **70**(1), 77–111 (2009).

[50]S. A. Isaacson and C. S. Peskin, "Incorporating diffusion in complex geometries into stochastic chemical kinetics simulations," SIAM J. Sci. Comput. **28**(1), 47–74 (2006).

[51]M. H. A. Davis, *Markov Models and Optimization* (CRC Press, 1993), Vol. 49.

[52]D. Schnoerr, G. Sanguinetti, and R. Grima, "The complex chemical Langevin equation," J. Chem. Phys. **141**(2), 024103 (2014).

[53]A. Crudu, A. Debussche, A. Muller, O. Radulescu *et al.*, "Convergence of stochastic gene networks to hybrid piecewise deterministic processes," Ann. Appl. Probab. **22**(5), 1822–1859 (2012).

[54]A. Duncan, S. Liao, T. Vejchodský, R. Erban, and R. Grima, "Noise-induced multistability in chemical systems: Discrete versus continuum modeling," Phys. Rev. E **91**(4), 042111 (2015).

[55]P. Bokes, J. R. King, A. Wood, and M. Loose, "Multiscale stochastic modelling of gene expression," J. Math. Biol. **65**(3), 493–520 (2012).

[56]D. J. Stekel and D. J. Jenkins, "Strong negative self regulation of prokaryotic transcription factors increases the intrinsic noise of protein expression," BMC Syst. Biol. **2**(1), 6 (2008).

[57]N. Friedman, L. Cai, and X. S. Xie, "Linking stochastic dynamics to population distribution: An analytical framework of gene expression," Phys. Rev. Lett. **97**(16), 168302 (2006).

[58]R. Milo and R. Phillips, *Cell Biology by the Numbers* (Garland Science, 2015).

[59]A. Duncan, R. Erban, and K. Zygalakis, "Hybrid framework for the simulation of stochastic chemical kinetics," J. Comput. Phys. **326**, 398–419 (2016).