HYBRID STOCHASTIC–DETERMINISTIC SOLUTION OF THE CHEMICAL MASTER EQUATION*

STEPHAN MENZ†, JUAN C. LATORRE†, CHRISTOF SCHÜTTE†, AND WILHELM HUISINGA‡

Abstract. The chemical master equation (CME) is the fundamental evolution equation of the stochastic description of biochemical reaction kinetics. In most applications it is impossible to solve the CME *directly* due to its high dimensionality. Instead *indirect* approaches based on realizations of the underlying Markov jump process are used such as the stochastic simulation algorithm (SSA). In the SSA, however, every reaction event has to be resolved explicitly such that it becomes numerically inefficient when the system's dynamics include fast reaction processes or species with high population levels. In many hybrid approaches, such fast reactions are approximated as continuous processes or replaced by quasi-stationary distributions either in a stochastic or deterministic context. Current hybrid approaches, however, almost exclusively rely on the computation of ensembles of stochastic realizations. We present a novel hybrid stochastic–deterministic approach to solve the CME *directly*. Starting point is a partitioning of the molecular species into discrete and continuous species that induces a partitioning of the reactions into discrete–stochastic and continuous–deterministic. The approach is based on a WKB approximation of a conditional probability distribution function (PDF) of the continuous species (given a discrete state) combined with a multiscale expansion of the CME. The black resulting hybrid stochastic–deterministic evolution equations comprise a CME with averaged propensities for the PDF of the discrete species that is coupled to an evolution equation of the partial expectation of the continuous species for each discrete state. In contrast to indirect hybrid methods, the impact of the evolution of discrete species on the dynamics of the continuous species has to be taken into account explicitly. The proposed approach is efficient whenever the number of discrete molecular species is small. We illustrate the performance of the new hybrid stochastic–deterministic approach in application to model systems of biological interest.

Key words. chemical master equation, hybrid model, multiscale analysis, partial averaging, asymptotic approximation, WKB-ansatz

AMS subject classifications.

1. Introduction. The last decade has witnessed an increased interest in stochastic descriptions of biochemical reaction networks due to considerable experimental evidence indicating that stochastic effects can play a crucial role in many cellular processes like gene expression and regulation [1, 2, 3, 4], where constituents are present in small numbers. The fundamental equation of stochastic chemical kinetics is the chemical master equation (CME). It defines the temporal evolution of the probability density function (PDF) of the system's state describing the number of molecules of each species at a given time. Only few approaches exist that directly solve for the PDF [5, 6, 7, 8]. The main problem is that the state space grows exponentially with the number of species, which renders most direct approaches computationally infeasible for larger reaction networks.

There exist a number of *approximate* solutions techniques to the CME [9, 10, 11, 12, 13]. Most are based on Monte Carlo (MC) simulations of the Markov jump process underlying the CME [9, 10], such as the stochastic simulation algorithm (SSA) [9]. An overview of existing methods is given in [14]. We call these MC-based methods *indirect* in the following. The advantage of indirect approaches is their easy applicability. They share, however, common disadvantages with MC-based approaches: There is always a *sampling error* since the PDF has to be approximated by a statistically significant ensemble of realizations. The number of realization required

*

†Department of Mathematics and Computer Science, Freie Universität Berlin, Germany
‡Institute of Mathematics, University of Potsdam, Germany

to meet a certain accuracy can be very large, depending on the quantities of interest (expectation values, higher moments or the PDF itself). Convergence can be slow, despite the fact that the approximation error decays like $Cn^{-1/2}$ with the number n of realizations. In addition, the constant C can be exponentially large if the system exhibits switching behavior [15].

Apart from the problem of judging the required number of realization to build-up sufficient statistics, the computational costs of each *single* realization of indirect methods generally depend on the number of reactions and the reaction events, i.e., the number of firings of a reaction channel. This property renders indirect methods numerically impracticable whenever the system dynamics includes rapidly firing reaction channels. Common approaches to avoid the explicit realization of every single reaction event are based on a τ-leap condition [16, 17] that allows one to approximate the reaction extents as independent Poisson random variables, or hybrid formulations [18, 19, 20], where state changes resulting from fast reactions are approximated as continuous processes or are incorporated via their conditional invariant measures (quasi-stationary PDFs) in the fast subspace [21, 22, 23]. The applicability of these indirect hybrid approaches depend on the existence of a *time scale gap* that allows one to distinguish between fast and slow reaction channels.

In this article, we propose and theoretically justify a novel *direct* hybrid stochastic–deterministic approach to solve the CME. Starting point is a partitioning of the molecular species into discrete and continuous species that induces a partitioning of the reactions into discrete–stochastic and continuous–deterministic. Using Bayes' formula, we decompose the full PDF $P(X,Z;t) = P(X|Z;t) \cdot P(Z;t)$ into the part $P(Z;t)$ related to the discrete species and the conditional PDF $P(X|Z;t)$ for the species to be modeled continuously. Based on some scaling parameter ε motivated by large population levels or fast reactions, we approximate the conditional PDF using the Wentzel-Kramers-Brillouin (WKB) approximation [24] and derive evolution equations based on a multiscale expansion of the scaled CME. This approach can be interpreted as taking partial expectations over the continuous species. If the entire state space is scaled, our approach resembles the well-known limit from the CME to the deterministic formulation of chemical reaction networks [25]. A first heuristic hybrid approach to directly solve the CME based on a partitioning of the state space was proposed in [26], where the coupling is realized in two separate steps: (i) propagation of the continuous variables and discrete distribution for a given time step; (ii) distribution of the continuous variables according to the change in the discrete distribution. In contrast to [26], we present a closed hybrid discrete PDE–ODE approach to the CME that implicitly integrates the propagation and distribution steps continuously. In contrast to indirect hybrid methods, the impact of the evolution of discrete species on the dynamics of the continuous species is taken into account explicitly.

The derivation of our hybrid approach neither requires a time scale gap nor does the resulting method suffer from the aforementioned disadvantages of indirect methods. Our approach will be more efficient than indirect approaches if the reaction system comprises a few molecular species in low quantities and the remaining species in larger quantities or associated with rapidly firing reaction channels. This is often the case for systems comprising gene regulation, transcription and metabolic regulatory networks.

The paper is organized as follows: First we introduce the general setting, including a brief background on the stochastic and deterministic formulation of reaction

networks and their relation. In the main part, we derive the proposed hybrid model, based on an asymptotic approximation of the PDF. We further illustrate the hybrid approach and study its efficiency, approximation error and applicability by three numerical examples.

2. The Chemical Master Equation and Mass Action Kinetics. Consider a system of N species S_i $(i = 1,\dots,N)$ that interact through M reaction channels R_μ $(\mu = 1,\dots,M)$ of the type

$$R_\mu: \quad s^r_{\mu 1} S_1 + \cdots + s^r_{\mu N} S_N \xrightarrow{k_\mu} s^p_{\mu 1} S_1 + \cdots + s^p_{\mu N} S_N \quad (\mu = 1,\dots,M), \tag{2.1}$$

where $s^r_{\mu i}$ and $s^p_{\mu i} \in \mathbb{N}_0$ are the stoichiometric coefficients of the reactant and product species S_i, respectively, and $k_\mu \in \mathbb{R}^+$ denotes the macroscopic rate constant of R_μ. The state of the system at time t is describe by the process $\boldsymbol{Y}(t) \in \mathbb{N}_0^N$ with

$$\boldsymbol{Y}_i(t) := \text{number of entities of species } S_i \text{ at time } t \quad (i = 1,\dots,N). \tag{2.2}$$

Firing of a reaction channel R_μ causes a net change $\upsilon_\mu \in \mathbb{Z}^N$ in the state of the system $\boldsymbol{Y}(t) \leftarrow \boldsymbol{Y}(t) + \upsilon_\mu$ with

$$\upsilon_\mu := \left(s^p_{\mu i} - s^r_{\mu i}\right)_{i=1}^N \quad (\mu = 1,\dots,M). \tag{2.3}$$

In the discrete–stochastic formulation of biochemical reaction kinetics, the state of the system is modeled as a continuous-time, discrete-state space Markov jump process $\boldsymbol{Y}(t)$. The probability that a channel R_μ fires and the process changes from state Y to state $Y + \upsilon_\mu$ in the next infinitesimal time interval $[t, t + \mathrm{d}t)$ is given by the propensity function a_μ according to

$$\mathbb{P}\Big[\boldsymbol{K}_\mu(t + \mathrm{d}t) - \boldsymbol{K}_\mu(t) = 1 \,|\, \boldsymbol{Y}_1(t) = Y_1,\dots,\boldsymbol{Y}_N(t) = Y_N\Big] = a_\mu(Y)\,\mathrm{d}t + o(\mathrm{d}t), \tag{2.4}$$

TABLE 2.1

Relevant elementary reactions and their propensity functions with respect to macroscopic rate constants k_μ and the corresponding conversion factor Ω, e.g., the system volume times the Avogadro constant $N_A \approx 6 \times 10^{23}\ mol^{-1}$.

Order	Reaction	Propensity
0^{th}	$\varnothing \xrightarrow{k_0} \cdots$	$a_0(Y) = k_0\Omega$
1^{st}	$S_i \xrightarrow{k_1} \cdots$	$a_1(Y) = k_1 Y_i$
2^{nd}	$S_i + S_j \xrightarrow{k_{2\text{a}}} \cdots$	$a_{2\text{a}}(Y) = \frac{k_{2\text{a}}}{\Omega} Y_i Y_j$
	$2S_i \xrightarrow{k_{2\text{b}}} \cdots$	$a_{2\text{b}}(Y) = \begin{cases} \frac{k_{2\text{b}}}{\Omega} Y_i (Y_i - 1) & \text{if } Y_i \geq 1 \\ 0 & \text{otherwise} \end{cases}$

where $o(\mathrm{d}t)$ refers to unspecified terms which satisfy $o(\mathrm{d}t)/\mathrm{d}t \to 0$ as $\mathrm{d}t \to 0$, and $\boldsymbol{K}_\mu(t)$ denotes the number of occurrences of reaction R_μ at time t. $\boldsymbol{K}_\mu$ is also a random variable (see for instance [27, 28] for more details). For an elementary reaction R_μ, the propensity is of the form

$$a_\mu(Y) = \begin{cases} \frac{k_\mu}{\Omega^{|s^r_\mu|-1}} \prod_{i=1}^{N} \frac{Y_i!}{(Y_i - s^r_{\mu i})!} & \text{if } Y_i \geq s^r_{\mu i} \text{ for all } i = 1,\dots,N, \\ 0 & \text{otherwise,} \end{cases} \tag{2.5}$$

with Ω denoting a factor related to conversion of the macroscopic rate constant k_μ, e.g., the system volume, the Avogadro constant or the product of both. The sum of all stoichiometric coefficients $|s^r_\mu| = \sum_{i=1}^{N} s^r_{\mu i}$ of an elementary reaction—specifying the number of reacting entities—is called reaction order. In many reaction systems, only zero, first and second order reactions are considered. Their propensities are given in Table 2.1. The time evolution of the probability density function (PDF)

$$P(Y;t) = \mathbb{P}[\boldsymbol{Y}_1(t) = Y_1, \dots, \boldsymbol{Y}_N(t) = Y_N], \tag{2.6}$$

is given by the chemical master equation (CME) [27]

$$\frac{\partial}{\partial t} P(Y;t) = \sum_{\mu=1}^{M} a_\mu\left(Y - \upsilon_\mu\right) P\left(Y - \upsilon_\mu; t\right) - a_\mu(Y) P(Y;t). \tag{2.7}$$

The CME (2.7) may be considered as a discrete partial differential equation (PDE), or, equivalently, as a countable system of ordinary differential equations (CODEs) [6].

In classical formulation of biochemical reaction kinetics, the state of the system at time t is approximated by a deterministic process $\boldsymbol{y}(t)$ on a continuous state space $\mathbb{R}^N_0$. The states y of the deterministic model are related to the states Y in the stochastic formulation by $y = Y/\Omega$; for instance with Ω denoting the system volume times the Avogadro constant in a model based on amount concentrations of the species. The time evolution of $\boldsymbol{y}(t)$ is given by the system of ordinary differential equations (ODEs):

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{y}(t) = \sum_{\mu=1}^{M} \upsilon_\mu \alpha_\mu(\boldsymbol{y}(t)), \tag{2.8}$$

with $\alpha_\mu(y) := a_\mu(Y)/\Omega$ denoting the Ω-scaled propensity of a reaction. For an elementary reaction R_μ it is

$$\begin{aligned} a^\Omega_\mu(Y) &= \frac{k_\mu}{\Omega^{|s^r_\mu|-1}} \prod_{i=1}^{N} \frac{Y_i!}{\left(Y_i - s^r_{\mu i}\right)!} = \Omega \frac{k_\mu}{\Omega^{|s^r_\mu|}} \prod_{i=1}^{N} \prod_{s=0}^{s^r_{\mu i}-1} (Y_i - s) \\ &= \Omega k_\mu \prod_{i=1}^{N} \prod_{s=0}^{s^r_{\mu i}-1} \left(y_i - \frac{s}{\Omega}\right) = \Omega \alpha_\mu(y), \end{aligned} \tag{2.9}$$

if $y_i \geq s^r_{\mu i}/\Omega$ for all $i = 1,\dots,N$, and $\alpha_\mu(y) = 0$ otherwise.

Remark: Relation to Reaction Rates. Based on the law of mass action, a deterministic model typically incorporates reaction rates $v_\mu(y) := k_\mu \prod_{i=1}^{N} y_i^{s^r_{\mu i}}$ instead of propensities α_μ. For elementary reactions R_μ with $s^r_{\mu i} \leq 1$ for all $i = 1,\ldots,N$ (e.g., zero and first order reactions), both, the rate v_μ and the propensity α_μ, are identical. In case of more complex reactions (any $s^r_{\mu i} > 1$), the rate v_μ approximates α_μ for $\Omega \gg 1$, since then

$$\alpha_\mu(y) = k_\mu \prod_{i=1}^{N} \prod_{s=0}^{s^r_{\mu i}-1} \left(y_i - \frac{s}{\Omega}\right) = k_\mu \prod_{i=1}^{N} y_i^{s^r_{\mu i}} + \mathcal{O}\left(\Omega^{-1}\right). \tag{2.10}$$

In the sequel, we will solely use reaction propensities.

As shown by T. G. Kurtz [29], in the thermodynamic limit, i.e., the number of entities of all species and the volume of the system approach infinity ($Y_0 \to \infty, \Omega \to \infty$) while the species concentrations converge to some value y_0, the continuous deterministic process $\boldsymbol{y}(t)$ given by eq. (2.8) approaches the discrete stochastic process that underlies the CME (2.7) for every finite time t. Hence, if species are present in larger numbers, the deterministic mass action kinetics is a good approximation to the CME. This well-known property is one of the key facts to be exploited by the herein proposed hybrid approach for solving the CME. While in the above results the thermodynamic limit is applied to the reaction system as a whole, the idea of our hybrid approach is to apply only a partial limit to those species that are present in large quantities. Since a partial volume limit is hard to justify for obvious reasons, we pursue a multiscale expansion approach with respect to a parameter $\varepsilon \ll 1$. This 'artificial' parameter ε will be linked to the abundance of species S_i with $Y_i \gg 1$ and plays a similar role as Ω^{-1} in the classical deterministic limit.

For our derivation it is instructive to link the mass action kinetics model to the evolution of expected values of the probability density function of the CME. The expected value of the stochastic process $\boldsymbol{Y}(t)$ at a specific time t is defined as

$$\mathbb{E}[\boldsymbol{Y}(t)] := \sum_{Y} Y \cdot P(Y;t). \tag{2.11}$$

Multiplication of the CME (2.7) with Y and summation over all possible states yields

$$\sum_{Y} Y \cdot \frac{\partial}{\partial t} P(Y;t) = \sum_{Y} Y \cdot \sum_{\mu=1}^{M} a_\mu\left(Y - \upsilon_\mu\right) P\left(Y - \upsilon_\mu; t\right) - a_\mu(Y)\, P(Y;t). \tag{2.12}$$

Exchange summations and exploiting the fact that $\boldsymbol{Y}(t)$ is non-negative allows us to re-index the first sum, yielding

$$\frac{\partial}{\partial t}\mathbb{E}[\boldsymbol{Y}(t)] = \sum_{\mu=1}^{M} \upsilon_\mu \mathbb{E}\left[a_\mu(\boldsymbol{Y}(t))\right]. \tag{2.13}$$

For zero and first order reactions R_μ, the propensity a_μ is a linear function and $\mathbb{E}[a_\mu(\boldsymbol{Y}(t))] = a_\mu(\mathbb{E}[\boldsymbol{Y}(t)])$ holds, resulting in

$$\frac{\partial}{\partial t}\mathbb{E}[\boldsymbol{Y}(t)] = \sum_{\mu=1}^{M} \upsilon_\mu a_\mu(\mathbb{E}[\boldsymbol{Y}(t)]). \tag{2.14}$$

For second and higher order reactions, the propensity is non-linear and, as it is well-known, $\mathbb{E}[a_\mu(\boldsymbol{Y}(t))] \neq a_\mu(\mathbb{E}[\boldsymbol{Y}(t)])$. However, from results of T. G. Kurtz we infer that $\mathbb{E}[a_\mu(\boldsymbol{Y}(t))] \approx a_\mu(\mathbb{E}[\boldsymbol{Y}(t)])$ is expected to hold for general reaction systems where molecular species are present in large numbers (close to the thermodynamic limit).

3. Derivation of the Hybrid CME–ODE Method. In the following, we derive a general hybrid description of the system dynamics where the time evolution of the probability density $P(Z;t)$ of all species present in small numbers Z is coupled to the time evolution of the partial expectations $\mathbb{E}_Z[\boldsymbol{X}] := \sum_Z X \cdot P(X,Z;t)$ of species with population levels X adequate for an approximation by deterministic mass action kinetics. The derivation will be based on a WKB-ansatz for the conditional probability $P(X|Z;t)$ and the resulting implications in computing expected values of $\boldsymbol{X}$. In general, the WKB-technique is a powerful method for approximating the solution of a linear differential equation whose highest derivative is multiplied by a small parameter ε, for more information see for instance [24]. In the context of the CME, the leading order WKB-approximation (eikonal function) is known to describe the mode of the probability function in the basin of an attractor [30, 31] that naturally leads to the corresponding deterministic formulation of the reaction processes [25]. In order to maintain the effects of discrete, stochastic fluctuations on the system dynamics, we apply this approximation only partially.

3.1. Definition of Discrete and Continuous Species and Reactions. We partition the system with respect to the species and their expected number of entities. Assume that for a given reaction network it can be distinguished between:

(i) 'Continuous' species S_i^c $(i = 1,\dots,N^c)$ whose changes in number of entities are approximated by a continuous, deterministic process.
(ii) 'Discrete' species S_i^d $(i = 1,\dots,N^d)$ whose changes in number of entities retain a discrete, stochastic description.

This partitioning is disjoint: $N^c + N^d = N$, and we rearrange the species-related variables accordingly:

$$Y = (X,Z)^{\mathrm{T}}, \quad \text{and} \quad \upsilon_\mu = \left(\nu_\mu, \zeta_\mu\right)^{\mathrm{T}} \quad (\mu = 1,\dots,M), \tag{3.1}$$

where X and ν_μ denote the number of entities and net changes of all continuous species S_i^c, and Z and ζ_μ denote the number of entities and net changes of all discrete species S_i^d, respectively. Finally, we represent the joint probability function $P(X,Z;t)$ using conditional probabilities as

$$P(X,Z;t) = P(X|Z;t)\,P(Z;t). \tag{3.2}$$

To correctly account for stochastic fluctuations in the discrete variable Z, all reaction channels R_μ that act on a discrete species S_i^d are modeled as a discrete, stochastic process. All other reactions influence only the continuous species. Based on the above assumption, those reactions are approximated as a continuous, deterministic process. Hence, the discrete–continuous partitioning of the species induces a corresponding partition of the reaction channels:

(i) 'Continuous' reactions do not change the number of entities of any discrete species S_i^d, i.e.,

$$\zeta_\mu = 0 \quad \text{for all } \mu \in M^c, \tag{3.3}$$

where M^c denotes the subset of all continuous reactions.

(ii) 'Discrete' reactions change the number of entities of at least one discrete species S_i^d, i.e.,

$$\zeta_\mu \neq 0 \quad \text{for all } \mu \in M^d, \tag{3.4}$$

where M^d denotes the subset of all discrete reactions.

As a consequence of the above partitioning, the net changes of all reaction channels can be rearranged as

$$\begin{array}{c} \\ \text{species} \begin{array}{c} \text{cont.} \\ \text{disc.} \end{array} \end{array} \begin{array}{c} \begin{array}{cc} \text{reactions} & \\ \text{cont.} & \text{disc.} \end{array} \\ \begin{pmatrix} \nu_{M^c} & \nu_{M^d} \\ 0 & \zeta_{M^d} \end{pmatrix} \end{array}. \tag{3.5}$$

Grouping terms together, we thus obtain

$$\begin{aligned} &\frac{\partial}{\partial t}\left[P(X|Z;t)\,P(Z;t)\right] \\ &\quad = P(Z;t)\sum_{\mu\in M^c} a_\mu\left(X-\nu_\mu,Z\right)P\left(X-\nu_\mu|Z;t\right) - a_\mu(X,Z)\,P(X|Z;t) \\ &\qquad + \sum_{\mu\in M^d}\Big[a_\mu\left(X-\nu_\mu,Z-\zeta_\mu\right)P\left(X-\nu_\mu|Z-\zeta_\mu;t\right)P\left(Z-\zeta_\mu;t\right) \\ &\qquad\qquad -a_\mu(X,Z)\,P(X|Z;t)\,P(Z;t)\Big]. \end{aligned} \tag{3.6}$$

3.2. ε-Scaling of the Continuous Species and Reactions. According to our assumption on continuous species, we scale their population levels with a factor $\varepsilon \ll 1$, i.e.,

$$x := \varepsilon \cdot X. \tag{3.7}$$

The parameter ε is related to the abundance of the continuous species S_i^c and used in the following asymptotic approximation to derive a partial limit of reaction kinetics. The exact value of ε may not be required, since the final equations in the scaled state space can be transformed back to the original unscaled state space. However, as the following hybrid approach gives an asymptotic approximation, the resulting error depends on the validity of this partial continuous–deterministic approximation and only vanish in the limit as $\varepsilon \to 0$.

In order to keep the probability invariant under the change of variables (3.7), the PDF of the scaled population levels is given by

$$P_\varepsilon(x,Z;t) = P_\varepsilon(x|Z;t)\,P_\varepsilon(Z;t) := \varepsilon^{N^c} P(X|Z;t)\,P(Z;t), \tag{3.8}$$

where N^c denotes the number of continuous species. Hence, with respect to the

scaled levels, the partitioned CME. (3.6) reads

$$\begin{aligned}\frac{\partial}{\partial t}&\left[P_\varepsilon(x|Z;t)\,P_\varepsilon(Z;t)\right]\\ &= P_\varepsilon(Z;t)\sum_{\mu\in M^c} a_\mu^\varepsilon\left(x-\varepsilon\nu_\mu,Z\right)P_\varepsilon\left(x-\varepsilon\nu_\mu|Z;t\right)-a_\mu^\varepsilon(x,Z)\,P_\varepsilon(x|Z;t)\\ &\quad+\sum_{\mu\in M^d}\Big[a_\mu^\varepsilon\left(x-\varepsilon\nu_\mu,Z-\zeta_\mu\right)P_\varepsilon\left(x-\varepsilon\nu_\mu|Z-\zeta_\mu;t\right)P_\varepsilon\left(Z-\zeta_\mu;t\right)\\ &\qquad\qquad-a_\mu^\varepsilon(x,Z)\,P_\varepsilon(x|Z;t)\,P_\varepsilon(Z;t)\Big],\end{aligned}\tag{3.9}$$

where $a_\mu^\varepsilon(x,Z)=a_\mu(X=x/\varepsilon,Z)$ for all $\varepsilon>0$. Intuitively it is clear that the intensity of a reaction process does not depend on the scale of the reactant levels, see also [28]. For example, one might interpret the scaling as some transformation of units.

In accordance with the definition of continuous species, we assume that the scaling of their levels imposes a corresponding scaling of the continuous reaction channels, analogously to the classical formulation of reaction kinetics. If $N^c=N$ and thus all levels are scaled, we require that our hybrid approach coincides with the purely deterministic limit. In this case, the parameter ε can be linked to the Ω-scaling in classical reaction kinetics via $\varepsilon=\Omega^{-1}$. Therefore, in line with eq. (2.9), we assume a corresponding ε-scaling of the propensities of all continuous reactions

$$\alpha_\mu(x,Z):=\varepsilon\cdot a_\mu^\varepsilon(x,Z)=\varepsilon\cdot a_\mu(X=x/\varepsilon,Z)\quad\text{for all }\mu\in M^c.\tag{3.10}$$

In our context, however, this scaling is only applied to the subset M^c of the reaction system, since the firing of a discrete channel results in changes of the process $\boldsymbol{Z}(t)$, which, by definition, necessitates stochastic reaction kinetics. Hence, the propensities of all discrete reactions are assumed to satisfy

$$\alpha_\mu(x,Z):=a_\mu^\varepsilon(x,Z)=a_\mu(X=x/\varepsilon,Z)\quad\text{for all }\mu\in M^d.\tag{3.11}$$

Under assumptions (3.10) and (3.11), eq. (3.9) becomes

$$\begin{aligned}\frac{\partial}{\partial t}&\left[P_\varepsilon(x|Z;t)\,P_\varepsilon(Z;t)\right]\\ &=\frac{1}{\varepsilon}P_\varepsilon(Z;t)\sum_{\mu\in M^c}\alpha_\mu\left(x-\varepsilon\nu_\mu,Z\right)P_\varepsilon\left(x-\varepsilon\nu_\mu|Z;t\right)-\alpha_\mu(x,Z)\,P_\varepsilon(x|Z;t)\\ &\quad+\sum_{\mu\in M^d}\Big[\alpha_\mu\left(x-\varepsilon\nu_\mu,Z-\zeta_\mu\right)P_\varepsilon\left(x-\varepsilon\nu_\mu|Z-\zeta_\mu;t\right)P_\varepsilon\left(Z-\zeta_\mu;t\right)\\ &\qquad\qquad-\alpha_\mu(x,Z)\,P_\varepsilon(x|Z;t)\,P_\varepsilon(Z;t)\Big].\end{aligned}\tag{3.12}$$

In the following, we seek an approximate solution of the ε-scaled CME (3.12) in the form of a multiscale expansion. We assume that the conditional probability $P_\varepsilon(x|Z;t)$ can be represented in a WKB-like series expansion with respect to the spatial coordinate, i.e.,

$$P_\varepsilon(x|Z;t)=\frac{1}{\sqrt{\varepsilon}}\exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}\left(U_0(x|Z;t)+\varepsilon U_1(x|Z;t)+\ldots\right),\tag{3.13}$$

where the factor $\varepsilon^{-1/2}$ is related to the normalization of $P_\varepsilon(x|Z;t)$. We represent the PDF $P_\varepsilon(Z;t)$ of the discrete species in an asymptotic series with respect to the spatial coordinate of the form

$$P_\varepsilon(Z;t) = P_0(Z;t) + \varepsilon P_1(Z;t) + \dots. \tag{3.14}$$

We assume that the functions s_0, U_i and P_i, $i = 0, 1, \dots$ in the above asymptotic expansion of the PDFs $P_\varepsilon(x|Z;t)$ and $P_\varepsilon(Z;t)$ are sufficiently continuously differentiable with respect to the arguments x and t.

3.3. Leading Order Approximation of Conditional PDF. We determine a solution of the conditional PDF $P_\varepsilon(x|Z;t)$ to its leading order $\mathcal{O}\left(\varepsilon^{-1}\right)$. Differentiation of eqs. (3.13) and (3.14) with respect to t gives

$$\begin{aligned}\frac{\partial}{\partial t}P_\varepsilon(x|Z;t) &= \frac{1}{\varepsilon}P_\varepsilon(x|Z;t)\frac{\partial}{\partial t}s_0(x|Z;t)\\ &\quad + \frac{1}{\sqrt{\varepsilon}}\exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}\left(\frac{\partial}{\partial t}U_0(x|Z;t) + \varepsilon\frac{\partial}{\partial t}U_1(x|Z;t) + \dots\right)\end{aligned} \tag{3.15}$$

and

$$\frac{\partial}{\partial t}P_\varepsilon(Z;t) = \frac{\partial}{\partial t}P_0(Z;t) + \varepsilon\frac{\partial}{\partial t}P_1(Z;t) + \dots. \tag{3.16}$$

Hence, on the left hand side of the ε-scaled CME (3.12) we find to leading order

$$\begin{aligned}&\frac{\partial}{\partial t}\left[P_\varepsilon(x|Z;t)\,P_\varepsilon(Z;t)\right]\\ &\quad= P_\varepsilon(Z;t)\frac{\partial}{\partial t}P_\varepsilon(x|Z;t) + P_\varepsilon(x|Z;t)\frac{\partial}{\partial t}P_\varepsilon(Z;t)\\ &\quad= \frac{1}{\sqrt{\varepsilon}}\left(\frac{1}{\varepsilon}P_0(Z;t)\exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}U_0(x|Z;t)\frac{\partial}{\partial t}s_0(x|Z;t) + \mathcal{O}(1)\right).\end{aligned} \tag{3.17}$$

In the right hand side of eq. (3.12), we first Taylor expand the eikonal function $s_0\left(x-\varepsilon\nu_\mu|Z;t\right)$ and $U_0\left(x-\varepsilon\nu_\mu|Z;t\right)$ around the state x, i.e.,

$$s_0\left(x-\varepsilon\nu_\mu|Z;t\right) = s_0(x|Z;t) - \varepsilon\nu_\mu^{\mathrm{T}}\nabla s_0(x|Z;t) + \mathcal{O}\left(\varepsilon^2\right) \tag{3.18}$$

and

$$U_0\left(x-\varepsilon\nu_\mu|Z;t\right) = U_0(x|Z;t) + \mathcal{O}(\varepsilon), \tag{3.19}$$

where $\nabla s_0(x|Z;t)$ denotes the gradient of s_0 with respect to x, defined as

$$\nabla s_0(x|Z;t) := \left(\frac{\partial}{\partial x_1}s_0(x|Z;t), \dots, \frac{\partial}{\partial x_{N^c}}s_0(x|Z;t)\right)^{\mathrm{T}}. \tag{3.20}$$

This implies the following expansion of the conditional PDF,

$$P_\varepsilon\left(x-\varepsilon\nu_\mu|Z;t\right) = \frac{1}{\sqrt{\varepsilon}}\exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}\exp\left\{-\nu_\mu^{\mathrm{T}}\nabla s_0(x|Z;t)\right\}\left(U_0(x|Z;t) + \mathcal{O}(\varepsilon)\right). \tag{3.21}$$

Similarly, we Taylor expand the propensities $\alpha_\mu\left(x-\varepsilon\nu_\mu, Z\right)$ in eq. (3.12) as

$$\alpha_\mu\left(x-\varepsilon\nu_\mu, Z\right) = \alpha_\mu(x,Z) + \mathcal{O}(\varepsilon). \tag{3.22}$$

Substituting both expansions (3.21) and (3.22) into the right hand side of the ε-scaled CME (3.12), using eq. (3.17) on the left hand side, multiplying with $\varepsilon^{-1/2}$ and comparing the terms of order $\mathcal{O}\left(\varepsilon^{-1}\right)$ on both sides, we obtain the leading order approximation of the continuous processes for a given discrete state Z as

$$\frac{\partial}{\partial t} s_0(x|Z;t) = \sum_{\mu\in M^c} \alpha_\mu(x,Z)\left[\exp\left\{-\nu_\mu^{\mathrm{T}}\nabla s_0(x|Z;t)\right\} - 1\right], \tag{3.23}$$

where we divided both sides by $P_0(Z;t)\exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}U_0(x|Z;t)$.[1] A convenient approach for the analysis of the eikonal function s_0 is to consider the PDE (3.23) as the Hamilton–Jacobi equation for the action of an auxiliary system with coordinates $\boldsymbol{x}(t|Z)$ and momenta $\boldsymbol{p}(t|Z) := \nabla s_0(x|Z;t)$. The corresponding Hamiltonian H of this system is defined as

$$H(\boldsymbol{x},\boldsymbol{p};t) := -\frac{\partial}{\partial t} s_0(x|Z;t) = -\sum_{\mu\in M^c} \alpha_\mu(\boldsymbol{x}(t|Z),Z)\left[\exp\left\{-\nu_\mu^{\mathrm{T}}\boldsymbol{p}(t|Z)\right\} - 1\right]. \tag{3.24}$$

The corresponding Hamiltonian equations of motion read

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}(t|Z) &:= \frac{\partial}{\partial \boldsymbol{p}} H(\boldsymbol{x},\boldsymbol{p};t) = \sum_{\mu\in M^c} \nu_\mu \alpha_\mu(\boldsymbol{x}(t|Z),Z)\exp\left\{-\nu_\mu^{\mathrm{T}}\boldsymbol{p}(t|Z)\right\}, \\
\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p}(t|Z) &:= -\frac{\partial}{\partial \boldsymbol{x}} H(\boldsymbol{x},\boldsymbol{p};t) = \sum_{\mu\in M^c} \nabla\alpha_\mu(\boldsymbol{x}(t|Z),Z)\left[\exp\left\{-\nu_\mu^{\mathrm{T}}\boldsymbol{p}(t|Z)\right\} - 1\right].
\end{aligned} \tag{3.25}$$

These equations of motions define the characteristics of the PDE (3.23) and their solution determines $s_0(x|Z;t)$ [32].

We assume that the initial function $s_0(x|Z;t=0)$ has a *unique* global extremum $x = \bar{\boldsymbol{x}}_0(Z)$ where

$$\bar{\boldsymbol{p}}_0(Z) \equiv \nabla s_0(x = \bar{\boldsymbol{x}}_0(Z)|Z;t=0) = 0. \tag{3.26}$$

Moreover, we will assume that this extremum is in fact the global maximum of $s_0(x|Z;t=0)$. In the regime $\varepsilon \ll 1$, this assumption ensures that the ε-parametrized family of distributions $P_\varepsilon(x|Z;t=0)$ has support concentrated around $\bar{\boldsymbol{x}}_0(Z)$ and width $\mathcal{O}\left(\sqrt{\varepsilon}\right)$.

Remark. One could choose, for instance, a Gaussian distribution of the form

$$P_\varepsilon(x|Z;t=0) = \frac{1}{\sqrt{\varepsilon}}\sqrt{\frac{\sigma^2}{(2\pi)^{N^c}}}\exp\left\{-\frac{1}{\varepsilon}\frac{|x-\bar{\boldsymbol{x}}_0(Z)|^2}{2\sigma^2}\right\}. \tag{3.27}$$

[1] If $P_0(Z;t)\exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}U_0(x|Z;t) = 0$, eqs. (3.12) and (3.17) result in an algebraic equation for $s_0(x|Z;t)$. This situation will be jointly dealt with the case where $P_0(Z;t) \ll 1$, see Section 3.7.

In this case,

$$s_0(x|Z;t=0) = -\frac{|x-\bar{\boldsymbol{x}}_0(Z)|^2}{2\sigma^2}, \tag{3.28}$$

and

$$U_0(x|Z;t=0) = \sqrt{\frac{\sigma^2}{(2\pi)^{N^c}}}, \quad U_i(x|Z;t=0) \equiv 0 \quad (i=1,2,\ldots). \tag{3.29}$$

The equations of motion (3.25) for the propagation in time of the extremum $\bar{\boldsymbol{x}}(t|Z)$ and corresponding $\bar{\boldsymbol{p}}(t|Z)$ are

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\boldsymbol{x}}(t|Z) = \sum_{\mu\in M^c} \nu_\mu \alpha_\mu(\bar{\boldsymbol{x}}(t|Z),Z) \quad \text{with} \quad \bar{\boldsymbol{x}}(t=0|Z) = \bar{\boldsymbol{x}}_0(Z), \tag{3.30}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\boldsymbol{p}}(t|Z) = 0 \quad \text{with} \quad \bar{\boldsymbol{p}}(t=0|Z) = \bar{\boldsymbol{p}}_0(Z) = 0. \tag{3.31}$$

Moreover, $\bar{\boldsymbol{x}}(t|Z)$ remains the unique maximum of s_0 for all $t > 0$.[2] The solution of eq. (3.30) hence gives the most probable values $\bar{\boldsymbol{x}}(t|Z)$ of the continuous species for a given discrete state Z, which is identical to the classical solution of biochemical reaction kinetics (compare eqs. (2.8) and (3.30)). However, it is important to realize that a solution of eq. (3.30), or more generally eq. (3.25), is only valid on the $\mathcal{O}(\varepsilon)$ scale, but the continuous processes also depend on the discrete–stochastic dynamics evolving on the $\mathcal{O}(1)$ scale. Hence, we wish to derive evolution equations that live on scales of order $\mathcal{O}(1)$. In *indirect* multiscale methods, for instance, eq. (3.30) is usually solved up to the predicted time of a next stochastic reaction event where the system is then updated accordingly and propagation of the characteristics continued for the corresponding new initial conditions [18, 19, 20] (see also [33] for an analysis on this type of multiscale numerical methods).

3.4. Laplace's Integral Approximation of Conditional PDF. In the previous section we derived the evolution equation (3.30) for the most probable value of $\boldsymbol{x}$ for each discrete state Z, as given by the maximum of the eikonal function s_0, and showed that this maximum will remain unique at $\bar{\boldsymbol{x}}(t|Z)$ for all times $t > 0$. As we demonstrate in this section, this result has important consequences for computing expectations of the process $\boldsymbol{x}$. Because of the special form of the ansatz (3.13), we can use the Laplace approximation to compute integrals of the form

$$\mathbb{E}^\varepsilon[f(x(t))|Z] = \int f(x)\,P_\varepsilon(x|Z;t)\,\mathrm{d}x. \tag{3.32}$$

[2] If there exists a solution $(\bar{\boldsymbol{y}}(t|Z), \bar{\boldsymbol{p}}(t|Z))$ of (3.25) with initial conditions $\bar{\boldsymbol{y}}_0(Z) \neq \bar{\boldsymbol{x}}_0(Z)$ and $\bar{\boldsymbol{p}}_0(Z) \neq 0$, but such that $\bar{\boldsymbol{p}}(t=T|Z) = 0$ for $T > 0$ (implying that $\bar{\boldsymbol{y}}(t=T|Z)$ is an extremum of $s_0(x|Z;t=T)$), then one could reverse time in eqs. (3.30)–(3.31) and conclude that necessarily $\bar{\boldsymbol{p}}_0(Z) = 0$, which yields a contradiction.

If the function $s_0(x|Z;t)$ has a unique maximum at the point $x = \bar{\boldsymbol{x}}(t|Z)$, then by Laplace's method [34], it is

$$\begin{aligned}
&\int f(x)\exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}\mathrm{d}x \\
&\qquad = \exp\left\{\frac{1}{\varepsilon}s_0(\bar{\boldsymbol{x}}(t|Z)\,|Z;t)\right\}\left(f(\bar{\boldsymbol{x}}(t|Z))\sqrt{\frac{\varepsilon(2\pi)^{N^c}}{\left|\nabla^2 s_0(\bar{\boldsymbol{x}}(t|Z);t)\right|}}+\mathcal{O}\left(\varepsilon^{3/2}\right)\right), \qquad (3.33)
\end{aligned}$$

where we have used that the Hessian matrix $\nabla^2 s_0(x = \bar{\boldsymbol{x}}(\cdot)\,|Z;t)$ is positive definite, since $\bar{\boldsymbol{x}}(\cdot)$ is a maximum of s_0. This result has an immediate consequence: Since the total probability $P_\varepsilon(x|Z;t)$ *has to* integrate to one, we can compute $s_0(x = \bar{\boldsymbol{x}}(\cdot)\,|Z;t)$ and $U_0(x = \bar{\boldsymbol{x}}(\cdot)\,|Z;t)$ directly. We have

$$\begin{aligned}
1 \equiv &\int P_\varepsilon(x|Z;t)\,\mathrm{d}x \\
&= \int \frac{1}{\sqrt{\varepsilon}}\exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}(U_0(x|Z;t)+\varepsilon U_1(x|Z;t)+\ldots)\,\mathrm{d}x \\
&= \frac{1}{\sqrt{\varepsilon}}\int \exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}(U_0(x|Z;t)+\mathcal{O}(\varepsilon))\,\mathrm{d}x \\
&= \frac{1}{\sqrt{\varepsilon}}\exp\left\{\frac{1}{\varepsilon}s_0(\bar{\boldsymbol{x}}(t|Z)\,|Z;t)\right\}\left(U_0(\bar{\boldsymbol{x}}(t|Z)\,|Z;t)\sqrt{\frac{\varepsilon(2\pi)^{N^c}}{\left|\nabla^2 s_0(\bar{\boldsymbol{x}}(t|Z)\,|Z;t)\right|}}+\mathcal{O}\left(\varepsilon^{3/2}\right)\right) \\
&= \exp\left\{\frac{1}{\varepsilon}s_0(\bar{\boldsymbol{x}}(t|Z)\,|Z;t)\right\}\left(U_0(\bar{\boldsymbol{x}}(t|Z)\,|Z;t)\sqrt{\frac{(2\pi)^{N^c}}{\left|\nabla^2 s_0(\bar{\boldsymbol{x}}(t|Z)\,|Z;t)\right|}}+\mathcal{O}(\varepsilon)\right). \qquad (3.34)
\end{aligned}$$

Since the above equation holds *for all* ε it follows that

$$s_0(\bar{\boldsymbol{x}}(t|Z)\,|Z;t) = 0, \qquad (3.35)$$

and

$$U_0(\bar{\boldsymbol{x}}(t|Z)\,|Z;t) = \sqrt{\frac{\left|\nabla^2 s_0(\bar{\boldsymbol{x}}(t|Z)\,|Z;t)\right|}{(2\pi)^{N^c}}}, \qquad (3.36)$$

for all Z and $t \geq 0$. Moreover, we find that the $\mathcal{O}(\varepsilon)$ corrections in eq. (3.34) must be identical to zero. In particular, this shows that the partial expectation of any function f of $\boldsymbol{x}$ with respect to $P_\varepsilon(\cdot, Z;t)$ can be approximated by

$$\begin{aligned}
\mathbb{E}_Z^\varepsilon[f(\boldsymbol{x})] := &\int f(x)\,P_\varepsilon(x,Z;t)\,\mathrm{d}x \\
&= \int f(x)\,P_\varepsilon(x|Z;t)\,\mathrm{d}x\cdot P_\varepsilon(Z;t) \\
&= \frac{1}{\sqrt{\varepsilon}}\int f(x)\exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}U_0(x|Z;t)\,\mathrm{d}x\cdot(P_0(Z;t)+\mathcal{O}(\varepsilon)) \\
&= f(\bar{\boldsymbol{x}}(t|Z))\,P_0(Z;t)+\mathcal{O}(\varepsilon), \qquad (3.37)
\end{aligned}$$

where we used Laplace's method and the above results for s_0 and U_0 (eqs. (3.35) and (3.36)) for the last equality. In a similar manner, we find that all expectation values with respect to the PDF $P_\varepsilon(\cdot, Z;t)$ can be computed from $\bar{\boldsymbol{x}}(t|Z)$ and $P_0(Z;t)$.[3]

3.5. Approximation of the PDF of Discrete Species. We use the results of the previous two sections to derive the evolution equation of $P_0(Z;t)$. Integration of the ε-scaled CME (3.12) over x results in

$$\begin{aligned}
\frac{\partial}{\partial t}P_\varepsilon(Z;t) &= \int \frac{\partial}{\partial t}\left[P_\varepsilon(x|Z;t)\,P_\varepsilon(Z;t)\right]\mathrm{d}x \\
&= \frac{1}{\varepsilon}P_\varepsilon(Z;t)\int \sum_{\mu\in M^c}\Big[\alpha_\mu\left(x-\varepsilon\nu_\mu,Z\right)P_\varepsilon\left(x-\varepsilon\nu_\mu|Z;t\right)-\alpha_\mu(x,Z)\,P_\varepsilon(x|Z;t)\Big]\mathrm{d}x \\
&\quad+\int \sum_{\mu\in M^d}\Big[\alpha_\mu\left(x-\varepsilon\nu_\mu,Z-\zeta_\mu\right)P_\varepsilon\left(x-\varepsilon\nu_\mu|Z-\zeta_\mu;t\right)P_\varepsilon\left(Z-\zeta_\mu;t\right) \\
&\qquad\qquad -\alpha_\mu(x,Z)\,P_\varepsilon(x|Z;t)\,P_\varepsilon(Z;t)\Big]\mathrm{d}x \\
&= 0+\int \sum_{\mu\in M^d}\Big[\alpha_\mu\left(x-\varepsilon\nu_\mu,Z-\zeta_\mu\right)P_\varepsilon\left(x-\varepsilon\nu_\mu|Z-\zeta_\mu;t\right)P_\varepsilon\left(Z-\zeta_\mu;t\right) \\
&\qquad\qquad -\alpha_\mu(x,Z)\,P_\varepsilon(x|Z;t)\,P_\varepsilon(Z;t)\Big]\mathrm{d}x.
\end{aligned} \tag{3.38}$$

This last equation implies that the evolution of $P_\varepsilon(Z;t)$ is given by scales of order $\mathcal{O}(1)$. Therefore, the dynamics of $P_\varepsilon(Z;t)$ can be expressed as

$$\begin{aligned}
\frac{\partial}{\partial t}P_\varepsilon(Z;t) &= \frac{\partial}{\partial t}P_0(Z;t)+\varepsilon\frac{\partial}{\partial t}P_1(Z;t)+\ldots \\
&= \frac{1}{\sqrt{\varepsilon}}\int \sum_{\mu\in M^d}\Big[\alpha_\mu\left(x,Z-\zeta_\mu\right)\exp\left\{\frac{1}{\varepsilon}s_0\left(x|Z-\zeta_\mu;t\right)\right\}U_0\left(x|Z-\zeta_\mu;t\right) \\
&\qquad\times\exp\left\{-\nu_\mu^{\mathrm{T}}\nabla s_0\left(x|Z-\zeta_\mu;t\right)\right\}P_0\left(Z-\zeta_\mu;t\right) \\
&\qquad -\alpha_\mu(x,Z)\exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}U_0(x|Z;t)\,P_0(Z;t)\Big]\mathrm{d}x+\mathcal{O}(\varepsilon),
\end{aligned} \tag{3.39}$$

where we consider the same expansions of $P_\varepsilon\left(x-\varepsilon\nu_\mu|Z;t\right)$ and $\alpha_\mu\left(x-\varepsilon\nu_\mu,Z\right)$ as already used for the eikonal approximation, see eqs. (3.21) and (3.22). Combing Laplace's integral approximation with the results for s_0 and U_0 given in eqs. (3.35) and (3.36), and comparing the terms of order $\mathcal{O}(1)$ on both sides of eq. (3.39) yields

[3] From this last result we can infer that the assumption that s_0 has a *unique* maximum at $t=0$ may be relaxed. If initially s_0 has many local maxima given by $\bar{\boldsymbol{x}}_0^i$, $i=1,2,\ldots$, these maxima will evolve independently in time according to eq. (3.30). The approximation of partial expectations of the form (3.37) will then be a superposition of terms of the form $f\left(\bar{\boldsymbol{x}}^i(t|Z)\right)$, as long as the distance between these maxima is $\mathcal{O}(1)$ so that Laplace's method can be applied.

the time evolution of $P_0(Z;t)$ as

$$\frac{\partial}{\partial t}P_0(Z;t) = \sum_{\mu\in M^d} \alpha_\mu\left(\bar{\boldsymbol{x}}\left(t\,|\,Z-\zeta_\mu\right), Z-\zeta_\mu\right) P_0\left(Z-\zeta_\mu;t\right) - \alpha_\mu\left(\bar{\boldsymbol{x}}(t\,|\,Z), Z\right) P_0(Z;t), \tag{3.40}$$

where we used the WKB-result that characteristics $\bar{\boldsymbol{x}}(t\,|\,Z)$ propagate the maximum of $P_\varepsilon(x\,|\,Z;t)$, and hence $\exp\left\{-\nu_\mu^{\mathrm{T}}\nabla s_0\left(\bar{\boldsymbol{x}}(t\,|\,Z)\,|\,Z;t\right)\right\} = 1$ for all Z and t.

3.6. Approximation of the Partial Expectation of Continuous Species. As we did in the previous section, we may derive the evolution of the leading order approximation for the partial expectations of the continuous species on the $\mathcal{O}(1)$ scale. As shown above, the partial expectation of any function f of $\boldsymbol{x}$ with respect to $P_\varepsilon(\cdot, Z;t)$ can be approximated by

$$\mathbb{E}_Z^\varepsilon\left[f(\boldsymbol{x})\right] = f\left(\bar{\boldsymbol{x}}(t\,|\,Z)\right) P_0(Z;t) + \mathcal{O}(\varepsilon). \tag{3.41}$$

This approximation is consistent with the derived evolution of $P_0(Z;t)$ in eq. (3.40). More importantly, it also allows us to derive the evolution of the leading order approximation of $\mathbb{E}_Z^\varepsilon[\boldsymbol{x}]$, i.e.,

$$\mathbb{E}_Z^0[\boldsymbol{x}] := \bar{\boldsymbol{x}}(t\,|\,Z)\, P_0(Z;t). \tag{3.42}$$

First note that multiplication of the ε-scaled CME (3.12) with x and subsequent integration gives

$$\begin{aligned}
\frac{\partial}{\partial t}\mathbb{E}_Z^\varepsilon[\boldsymbol{x}] &= \int x \frac{\partial}{\partial t}\left[P_\varepsilon(x\,|\,Z;t)\, P_\varepsilon(Z;t)\right]\mathrm{d}x \\
&= \frac{1}{\varepsilon}P_\varepsilon(Z;t)\int x \sum_{\mu\in M^c}\left[\alpha_\mu\left(x-\varepsilon\nu_\mu, Z\right)P_\varepsilon\left(x-\varepsilon\nu_\mu\,|\,Z;t\right) - \alpha_\mu(x,Z)\,P_\varepsilon(x\,|\,Z;t)\right]\mathrm{d}x \\
&\quad + \int x \sum_{\mu\in M^d}\Big[\alpha_\mu\left(x-\varepsilon\nu_\mu, Z-\zeta_\mu\right)P_\varepsilon\left(x-\varepsilon\nu_\mu\,|\,Z-\zeta_\mu;t\right)P_\varepsilon\left(Z-\zeta_\mu;t\right) \\
&\qquad\qquad - \alpha_\mu(x,Z)\,P_\varepsilon(x\,|\,Z;t)\,P_\varepsilon(Z;t)\Big]\mathrm{d}x \\
&= P_\varepsilon(Z;t)\int \sum_{\mu\in M^c}\left[\nu_\mu\alpha_\mu(x,Z)\,P_\varepsilon(x\,|\,Z;t)\right]\mathrm{d}x \\
&\quad + \int x \sum_{\mu\in M^d}\Big[\alpha_\mu\left(x-\varepsilon\nu_\mu, Z-\zeta_\mu\right)P_\varepsilon\left(x-\varepsilon\nu_\mu\,|\,Z-\zeta_\mu;t\right)P_\varepsilon\left(Z-\zeta_\mu;t\right) \\
&\qquad\qquad - \alpha_\mu(x,Z)\,P_\varepsilon(x\,|\,Z;t)\,P_\varepsilon(Z;t)\Big]\mathrm{d}x.
\end{aligned} \tag{3.43}$$

Once again, we observe that the terms of order $\mathcal{O}\left(\varepsilon^{-1}\right)$ do not appear in the right hand side of eq. (3.43), and therefore the dynamics of $\mathbb{E}_Z^{\varepsilon}[\boldsymbol{x}]$ can be approximated by

$$\begin{aligned}\frac{\partial}{\partial t}\mathbb{E}_Z^{\varepsilon}[\boldsymbol{x}] &= \frac{\partial}{\partial t}\mathbb{E}_Z^{0}[\boldsymbol{x}] + \mathcal{O}(\varepsilon)\\ &= P_0(Z;t)\frac{1}{\sqrt{\varepsilon}}\int \sum_{\mu\in M^c}\left[\nu_\mu \alpha_\mu(x,Z)\exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}U_0(x|Z;t)\right]\mathrm{d}x\\ &\quad + \frac{1}{\sqrt{\varepsilon}}\int x\sum_{\mu\in M^d}\Big[\alpha_\mu\left(x,Z-\zeta_\mu\right)\exp\left\{\frac{1}{\varepsilon}s_0\left(x|Z-\zeta_\mu;t\right)\right\}U_0\left(x|Z-\zeta_\mu;t\right)\\ &\qquad\qquad \times\exp\left\{-\nu_\mu^{\mathrm{T}}\nabla s_0\left(x|Z-\zeta_\mu;t\right)\right\}P_0\left(Z-\zeta_\mu;t\right)\\ &\qquad\qquad -\alpha_\mu(x,Z)\exp\left\{\frac{1}{\varepsilon}s_0(x|Z;t)\right\}U_0(x|Z;t)P_0(Z;t)\Big]\mathrm{d}x+\mathcal{O}(\varepsilon).\end{aligned} \tag{3.44}$$

Applying Laplace's method, using the results for s_0 and U_0 in eqs. (3.35) and (3.36), and comparing the terms of order $\mathcal{O}(1)$ on both sides, we find

$$\begin{aligned}\frac{\partial}{\partial t}\mathbb{E}_Z^{0}[\boldsymbol{x}] &= P_0(Z;t)\sum_{\mu\in M^c}\nu_\mu\alpha_\mu(\bar{\boldsymbol{x}}(t|Z),Z)\\ &\quad + \sum_{\mu\in M^d}\alpha_\mu\left(\bar{\boldsymbol{x}}\left(t|Z-\zeta_\mu\right),Z-\zeta_\mu\right)\mathbb{E}^0_{Z-\zeta_\mu}[\bar{\boldsymbol{x}}]-\alpha_\mu(\bar{\boldsymbol{x}}(t|Z),Z)\,\mathbb{E}_Z^0[\bar{\boldsymbol{x}}].\end{aligned} \tag{3.45}$$

In summary, eqs. (3.40) and (3.45) describe the evolutions of $P_0(Z;t)$ and $\mathbb{E}_Z^0[\boldsymbol{x}]$. It is worth mentioning that although $\mathbb{E}_Z^0[\boldsymbol{x}]$ is determined explicitly by $\bar{\boldsymbol{x}}(t|Z)$ and $P_0(Z;t)$, it cannot be recovered by direct integration of eqs. (3.30) and (3.40), since the evolution of $\bar{\boldsymbol{x}}(t|Z)$ as given by eq. (3.30) is only valid on the order $\mathcal{O}(\varepsilon)$. The evolution equation (3.45) of $\mathbb{E}_Z^0[\boldsymbol{x}]$, on the other hand, is valid on the order $\mathcal{O}(1)$, which is the scale we are interested in. The corresponding conditional levels $\bar{\boldsymbol{x}}(t|Z)$ can then be recovered by the relation (3.42), as we shall see in the next sections.

3.7. Final Equations of the Hybrid CME–ODE Approach. Summarizing, we derived the following hybrid system approximating the coupled dynamics of the stochastic and deterministic processes up to order $\mathcal{O}(1)$. The time evolution of the probability distribution $P_0(Z;t)$ of the discrete species is given by eq. (3.40), i.e.,

$$\frac{\partial}{\partial t}P_0(Z;t) = \sum_{\mu\in M^d} a_\mu\left(\bar{\boldsymbol{X}}\left(t|Z-\zeta_\mu\right),Z-\zeta_\mu\right)P_0\left(Z-\zeta_\mu;t\right) - a_\mu\left(\bar{\boldsymbol{X}}\left(t|Z\right),Z\right)P_0(Z;t), \tag{3.46}$$

where we re-substituted the functions α_μ by the original propensities a_μ for the discrete reactions using definition (3.11). By eq. (3.45), the partial expectations

$\mathbb{E}^0_Z[\boldsymbol{X}] = \varepsilon^{-1}\mathbb{E}^0_Z[\boldsymbol{x}]$ of the continuous species for a discrete state Z are defined by

$$\frac{\partial}{\partial t}\mathbb{E}^0_Z[\boldsymbol{X}] = \underbrace{P_0(Z;t)\sum_{\mu\in M^c}\nu_\mu a_\mu\left(\bar{\boldsymbol{X}}(t|Z),Z\right)}_{\text{impact of continuous reactions}} + \underbrace{\sum_{\mu\in M^d} a_\mu\left(\bar{\boldsymbol{X}}\left(t|Z-\zeta_\mu\right),Z-\zeta_\mu\right)\mathbb{E}^0_{Z-\zeta_\mu}[\boldsymbol{X}] - a_\mu\left(\bar{\boldsymbol{X}}(t|Z),Z\right)\mathbb{E}^0_Z[\boldsymbol{X}]}_{\text{impact of discrete reactions}}, \quad (3.47)$$

where we replaced all α_μ by the original propensities a_μ as given by eqs. (3.10) and (3.11). To obtain the solution of $\bar{\boldsymbol{X}}(t|Z)$ that is valid for scales of order $\mathcal{O}(1)$ from the above equations, we use definition (3.42), i.e., divide $\mathbb{E}^0_Z[\boldsymbol{X}] = \bar{\boldsymbol{X}}(t|Z)P_0(Z;t)$ by $P_0(Z;t) > 0$ or set $\bar{\boldsymbol{X}}(t|Z) = 0$ if $P_0(Z;t) = 0$, respectively.

As the size of the discrete, stochastic state space is drastically reduced a standard numerical integration of eqs. (3.46) and (3.47) becomes applicable even for more complex systems. During numerical integration of the hybrid system the actual propensities have to be evaluated at every integration step. This also requires the explicit computation of the conditional values $\bar{\boldsymbol{X}}(t|Z)$ of the continuous species from $P_0(Z;t)$ and $\mathbb{E}^0_Z[\boldsymbol{X}]$. To avoid numerical instabilities, we introduce a threshold $\delta \ll 1$: Whenever the actual value of $P_0(Z;t) \le \delta$, we set $P_0(Z;t) = 0$ and $\mathbb{E}^0_Z[\boldsymbol{X}] = 0$ in the evaluation of eqs. (3.46) and (3.47), i.e., the actual dynamics of the hybrid system will be constrained to those discrete states with $P_0(Z;t) > \delta$ during integration. For appropriate choices of δ, the additional error made with this criterion can be expected to be negligible. A reasonable choice for δ would be a value not greater than the allowed absolute error used in the numerical integration of the hybrid system.

3.8. Coarse Graining of the Continuous Processes. In many reaction systems, most reactions are of zero, first or second order, i.e., the propensity functions will depend on the level of a few species only, see Table 2.1. If the propensities of the continuous reactions are constant on a subset $\mathcal{Z}_k$ of discrete states, then we may associate a single ODE with the entire subset $\mathcal{Z}_k$ rather than an ODE for each element of the subset. We therefore seek subsets $\mathcal{Z}_k$ of discrete states such that for all $\mu \in M^c$:

$$a_\mu(\cdot, Z) = const \quad \text{for all } Z \in \mathcal{Z}_k. \quad (3.48)$$

If criterion (3.48) holds, we find that the evolution equations of the eikonal functions $s_0(\cdot|Z;t)$ are identical for every $Z \in \mathcal{Z}_k$. Given equal initial conditions $X = \bar{\boldsymbol{X}}_0(Z)$ for every $Z \in \mathcal{Z}_k$, they would therefore propagate in time along the same characteristic $\bar{\boldsymbol{X}}(t|\mathcal{Z}_k)$. Consequently, we assign the same expectation $\mathbb{E}^\varepsilon_{\mathcal{Z}_k}[\boldsymbol{X}]$ to each subset $\mathcal{Z}_k$ of discrete states:

$$\mathbb{E}^\varepsilon_{\mathcal{Z}_k}[\boldsymbol{X}] := \sum_{Z\in\mathcal{Z}_k}\mathbb{E}^\varepsilon_Z[\boldsymbol{X}] = \bar{\boldsymbol{X}}(t|\mathcal{Z}_k)\sum_{Z\in\mathcal{Z}_k}P_0(Z;t) + \mathcal{O}(\varepsilon). \quad (3.49)$$

The time evolution of $\mathbb{E}^0_{\mathcal{Z}_k}[\boldsymbol{X}] := \bar{\boldsymbol{X}}(t|\mathcal{Z}_k)\sum_{Z\in\mathcal{Z}_k}P_0(Z;t)$ is than given by summation

of eq. (3.47) over all discrete states $Z \in \mathcal{Z}_k$, i.e.,

$$\frac{\partial}{\partial t}\mathbb{E}^0_{\mathcal{Z}_k}[\boldsymbol{X}] = \sum_{Z\in\mathcal{Z}_k} P_0(Z;t) \sum_{\mu\in M^c} \nu_\mu a_\mu\left(\bar{\boldsymbol{X}}(t\,|\,\mathcal{Z}_k), \mathcal{Z}_k\right) + \sum_{Z\in\mathcal{Z}_k}\sum_{\mu\in M^d}\left[a_\mu\left(\bar{\boldsymbol{X}}\left(t\,|\,Z-\zeta_\mu\right), Z-\zeta_\mu\right)\mathbb{E}^0_{Z-\zeta_\mu}[\boldsymbol{X}] - a_\mu\left(\bar{\boldsymbol{X}}(t\,|\,Z), Z\right)\mathbb{E}^0_Z[\boldsymbol{X}]\right], \tag{3.50}$$

where $a_\mu\left(\bar{\boldsymbol{X}}(t\,|\,\mathcal{Z}_k), \mathcal{Z}_k\right)$ in the first sum denotes the propensity of a continuous reaction on $\mathcal{Z}_k$ evaluated for $\bar{\boldsymbol{X}}(t\,|\,\mathcal{Z}_k)$. Again, the second summand in the right hand side of eq. (3.50) describes the impact of the discrete reactions on the partial expectation of the continuous species. As these are coarse grained, all terms describing the exchange on the same subset $\mathcal{Z}_k$ cancel out, and only those terms related to an in- or outflow of probability to be in $\mathcal{Z}_k$ remain in the second summand of eq. (3.50).

Any conditional value $\bar{\boldsymbol{X}}(t\,|\,\mathcal{Z}_k)$ necessary for the evaluation of eq. (3.50) can be computed from $\mathbb{E}^0_{\mathcal{Z}_k}[\boldsymbol{X}] := \bar{\boldsymbol{X}}(t\,|\,\mathcal{Z}_k)\sum_{Z\in\mathcal{Z}_k} P_0(Z;t)$ and the PDFs $P_0(Z;t)$ as described in the previous subsection. Since $\bar{\boldsymbol{X}}(t\,|\,Z) = \bar{\boldsymbol{X}}(t\,|\,\mathcal{Z}_k)$ for all $Z \in \mathcal{Z}_k$, this also allows to compute the leading order approximation $\mathbb{E}^0_Z[\boldsymbol{X}] = \bar{\boldsymbol{X}}(t\,|\,Z)\,P_0(Z;t)$ of the partial expectations at a specific state Z, which are used in the second summand of eq. (3.50). The evolution equations of the PDFs $P_0(Z;t)$ are not effected by the suggested coarse graining.

3.9. Algorithmic Flow. The main steps of our hybrid approach may be summarized as follows:

Step 1 (Partition the Species). Define a partition of the N species into discrete and continuous species, i.e., $S^d_1,\ldots,S^d_{N^d}$ and $S^c_1,\ldots,S^c_{N^c}$, respectively, with $N^d + N^c = N$. Such partitioning can be based on the expected levels and conservation properties of the species, different time-scales of the reactions or other prior knowledge on the system dynamics.

Step 2 (Partition the Reactions). According to conditions (3.3) and (3.4), treat every reaction R_μ that changes a discrete species S^d_i as a discrete process. Set all remaining reactions as continuous processes.

Step 3 (Formulate Hybrid CME–ODE Equations). If the dynamics of the N^d discrete species is *a priori* known to be restricted to a specific subset of states, formulate the corresponding system of (linear) ODEs for the time evolution of the PDF as described by eq. (3.46). Otherwise, choose a reasonable initial subset of discrete states. The approximation error can be bounded by the finite state projection (FSP) algorithm [5]. Further check the dependence of the propensities of all continuous reactions on the discrete state space in order to identify subsets $\mathcal{Z}_k$ of discrete states where their values are constant. If coarse graining can be applied, assign a system of ODEs for all N^c continuous species to every subset $\mathcal{Z}_k$ that satisfies eq. (3.50). Otherwise assign an ODE system for the continuous species to every discrete state as described by eq. (3.47). Furthermore, our above scaling approach does hold for more general situation than just some species with large copy numbers. We just require equations (3.10) and (3.11) to be valid in the specific subspaces. Therefore, we can also integrate rapidly firing reactions (with rate constants scaling like ε^{-1}). Vice versa, observation of a critically small propensity (order ε) in one of the reactions in

M^c may spoil the approximation property, and thus has to be considered in choosing the discrete states and reactions.

Step 4 (Compute Numerical Solution). Numerical integration of the final hybrid model requires appropriate methods for efficient solution of differential algebraic equations (DAEs). Since a variety of such methods are available and choosing the optimal one for the case at hand is a problem on its own, we refrain from addressing this topic in depth. Instead we apply a straight-forward approach to the simulation that exploits the threshold value δ introduced above: Choose an appropriate δ for the evaluation of the continuous levels from $\mathbb{E}^0_Z[\boldsymbol{X}]$ or $\mathbb{E}^0_{\mathcal{Z}_k}[\boldsymbol{X}]$, respectively. Start numerical integration of the system. If applicable, monitor the loss in probability mass based on the FSP algorithm and dynamically update the support of the system by expanding the discrete state space. Update the corresponding ODE system of the PDF and the associated expectations of the continuous species accordingly. For state space expansion, strategies elaborated in [5, 13] can be used.

To highlight the benefits of our hybrid approach, we compare numerical costs in terms of number of equations that have to be integrated. Assume that the population level of every discrete species is bounded by the same maximal value m, and the level of every continuous species can maximally reach h, with $h \gg m$ by assumption. Hence, the state space of the system would include $(m+1)^{N^d} \cdot (h+1)^{N^c}$ states that have to be considered in the spatial discretization of the CME. In contrast, in our hybrid approach the system dynamics would be described by $(m+1)^{N^d} \cdot (1+N^c)$ equations only. These are given by the $(m+1)^{N^d}$ states necessary for the support of the PDF of discrete species, and the ODEs for the expectation levels of the N^c different continuous species associated with each of these discrete states. The reduction would be even higher if the continuous processes can be further coarse grained to subsets of the discrete state space. In that case we would have only $(m+1)^{N^d} + K \cdot N^c$ equations, where $K \ll (m+1)^{N^d}$ denotes the number of subsets used for coarse graining. As the numerical costs of our hybrid approach basically scale with the number of discrete states, we expect it to be especially efficient for systems that include a few species that have to be modeled discretely, e.g., gene regulatory or signaling networks.

4. Numerical Experiments. The SSA and an explicit Runge-Kutta method of order 4 with error and step size control were implemented in C++ . All numerical experiments were performed on an Intel® Core™2 Duo processor with 2 GHz and 2 GB RAM. In each example, the numerical solution of the proposed hybrid approach was compared to the corresponding predictions obtained from ten thousand SSA simulations of the full CME. Numerical integration of the hybrid equations was performed with an absolute tolerance of 10^{-6} and a relative tolerance of 10^{-3}. The δ-threshold was set to 10^{-6}, in accordance with the selected absolute error. For state space truncation, appropriate levels for each discrete species were chosen initially, accounting for higher values than observed in the SSA simulations. No additional boundary conditions were applied to the truncated systems. Truncation error was monitored by loss of probability mass, which was at final times of all experiments much lower than the allowed absolute error.

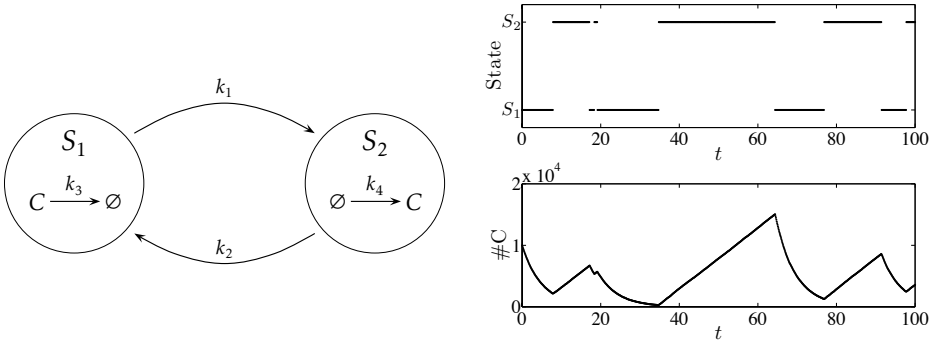

FIGURE 4.1. *Left: Illustration of system* (4.1)*, used as a first test example. The system can switch between two (discrete) states S_1 and S_2. The degradation and production of the (continuous) species C depends on the actual state of the system: In S_1 only the degradation of C is active, whereas in S_2 only the production of C is active. Right: Typical results of a single SSA-realization of system* (4.1) *for the parameter values listed in Table 4.1.*

4.1. A Simple Switch-Model. Metastability is an important property of biological systems, necessitating in general a stochastic modeling approach for its *in silico* analysis. As a first test example for our hybrid approach, we designed a simple bistable model of three species and four elementary reactions:

$$\begin{aligned} R_1\colon\ & S_1 \xrightarrow{k_1} S_2, & R_2\colon\ & S_2 \xrightarrow{k_2} S_1, \\ R_3\colon\ & C + S_1 \xrightarrow{k_3} S_1, & R_4\colon\ & S_2 \xrightarrow{k_4} C + S_2. \end{aligned} \tag{4.1}$$

The concept behind this network is that S_1 and S_2 represent two metastable states of a system, as illustrated in Fig. 4.1, left panel. These states are considered as discrete species in our hybrid approach with a possible value of either zero ore one. Transitions between S_1 and S_2 are modeled by reactions R_1 and R_2 that are treated as discrete processes in the following. Depending on the actual state of the system, the third, continuously treated species C gets either degraded or produced through reactions R_3 and R_4, respectively. We included this state-dependence of R_3 and R_4 in their stoichiometries, see eqs. (4.1). The discrete species S_1 and S_2 are not affected by a firing of channel R_3 or R_4, and hence, both reactions are treated as continuous, deterministic processes (see also Fig. 4.1, right panel, for a typical SSA-realization of the system).

The full set of hybrid equations of network (4.1) is given by

$$\begin{aligned} \frac{\partial}{\partial t} P_0(S_1;t) &= -k_1 P_0(S_1;t) + k_2 P_0(S_2;t) = -\frac{\partial}{\partial t} P_0(S_2;t), \\ \frac{\partial}{\partial t} \mathbb{E}^0_{S_1}[\boldsymbol{X}] &= -k_1 \mathbb{E}^0_{S_1}[\boldsymbol{X}] + k_2 \mathbb{E}^0_{S_2}[\boldsymbol{X}] - k_3 \mathbb{E}^0_{S_1}[\boldsymbol{X}] \\ \frac{\partial}{\partial t} \mathbb{E}^0_{S_2}[\boldsymbol{X}] &= +k_1 \mathbb{E}^0_{S_1}[\boldsymbol{X}] - k_2 \mathbb{E}^0_{S_2}[\boldsymbol{X}] + k_4 P_0(S_2;t), \end{aligned} \tag{4.2}$$

where $P_0(S_1;t)$ and $P_0(S_2;t)$ denote the probabilities that the system is either in state S_1 or S_2 at time t, respectively. The approximated partial expectations of species C

with respect to S_1 and S_2 are denoted by $\mathbb{E}^0_{S_1}[\boldsymbol{X}]$ and $\mathbb{E}^0_{S_2}[\boldsymbol{X}]$. It should be noted that the corresponding conditional levels $\bar{\boldsymbol{X}}(t|S_1)$ and $\bar{\boldsymbol{X}}(t|S_2)$ of C do not have to be computed for the evaluation of the hybrid equations (4.2). This is generally the case if all propensities in a reaction network are linear functions with respect to the continuous species.

In Figures 4.2 and 4.3, the numerical solution of eqs. (4.2) is compared to predictions obtained by SSA simulations. In the upper panels of Figure 4.2, the time evolution of the probabilities $P_0(S_1;t)$ and $P_0(S_2;t)$ (left), as well as the evolution of the partial expectations $\mathbb{E}^0_{S_1}[\boldsymbol{X}]$ and $\mathbb{E}^0_{S_2}[\boldsymbol{X}]$ (right) are shown. Additionally, Figure 4.2 depicts the corresponding conditional levels $\bar{\boldsymbol{X}}(t|S_1)$ and $\bar{\boldsymbol{X}}(t|S_2)$ (lower left panel) and the approximation of the total expectation $\mathbb{E}^0[\boldsymbol{X}]$ of C (lower right panel). The relative error between the results obtained by both methods is plotted against time in Figure 4.3. As can be seen, the solution of our hybrid equations is in excellent agreement with the SSA results. For the model system (4.1), we may demonstrate that our hybrid approach corresponds to the exact evolution equations. This is guaranteed by two properties of the network: Independence of R_1 and R_2 from C, and linearity of R_3 and R_4 with respect to C.

In general, the derived hybrid equations (3.46) and (3.47) are exact if:

(i) The stochastic part is decoupled from the deterministic processes, i.e., the propensities of all discrete reactions M^d are independent functions of the continuous species X

$$a_\mu(X,Z) = a_\mu(Z) \quad \text{for all } \mu \in M^d. \tag{4.3}$$

(ii) All continuous reactions M^c are linear functions with respect to the continuous species X, which guarantees

$$\mathbb{E}[a_\mu(\boldsymbol{X},Z)\,|Z] = a_\mu(\mathbb{E}[\boldsymbol{X}\,|Z],Z) \quad \text{for all } \mu \in M^c, \tag{4.4}$$

with $\mathbb{E}[f(\boldsymbol{X})\,|Z] := \sum_X f(X)P(X|Z;t)$ denoting the conditional expectation of any function f of $\boldsymbol{X}$ with respect to $P(\cdot\,|Z;t)$.

Starting from the full, unscaled CME (3.6) one can simply sum over the continuous species X to obtain the marginal probability for the discrete species Z. Assuming (4.3), the equation for $P(Z,t)$ becomes exactly equation (3.46) for $P_0(Z,t)$ obtained in the previous sections. Similarly, by multiplying (3.12) by X and summing over all possible states, the linearity assumption (4.4) yields equation (3.47) for the evolution of $\mathbb{E}_Z[\boldsymbol{X}]$ by identifying the conditional expectations $\mathbb{E}[\boldsymbol{X}\,|Z]$ of the continuous species with the conditional levels $\bar{\boldsymbol{X}}(t|Z)$ in our setting. For that reason, any discrepancies to SSA results, as for instance illustrated in Figure 4.3, is in this case associated to the sampling error of the MC method.

TABLE 4.1
Parameter values used for system (4.1), *conversion factor* $\Omega = 1$. *The system is started in state* S_1 *with ten thousand entities of species* C *and solved for a time period of* 100.

k_1	k_2	k_3	k_4
0.06	0.04	0.2	500.0

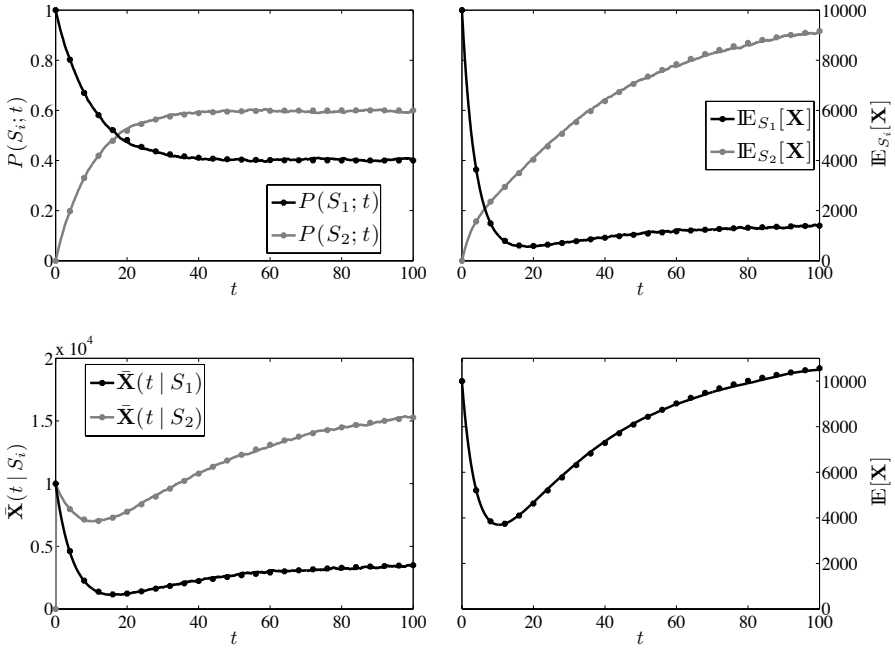

FIGURE 4.2. *Time evolution of system* (4.1) *as predicted by ten thousand SSA runs (solid lines) and the numerical solution (marked by stars) of the corresponding hybrid equations* (4.2) *for the parameter values listed in Table 4.1.*

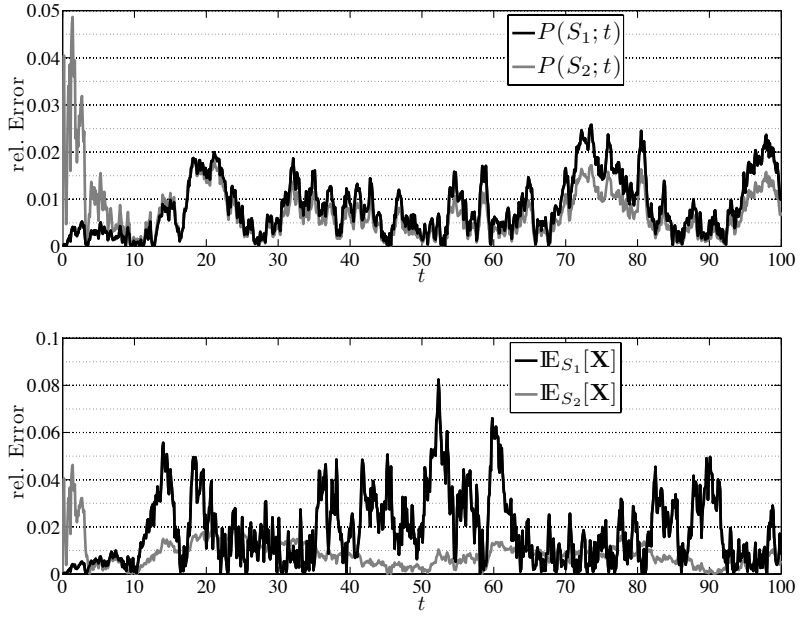

FIGURE 4.3. *Relative error of the numerical solution of the hybrid equations* (4.2) *with respect to the predictions obtained by ten thousand SSA runs of system* (4.1)*, as shown in Fig. 4.2.*

4.2. Viral Infection Kinetics. As a second example, we consider the infection model of a non-lytic virus proposed by Srivastava et al. [36]. We also use this model to illustrate the suggested coarse graining of the continuous processes. The model includes three viral components: viral nucleic acids, classified as either genomic (*gen*) or template (*tem*), and viral structural protein (*struct*), governed by six elementary reactions:

$$\begin{aligned} R_1&: \quad gen \xrightarrow{k_1} tem, & R_2&: \quad tem \xrightarrow{k_2} \varnothing, \\ R_3&: \quad tem \xrightarrow{k_3} tem + gen, & R_4&: \quad gen + struct \xrightarrow{k_4} \varnothing, \\ R_5&: \quad tem \xrightarrow{k_5} tem + struct, & R_6&: \quad struct \xrightarrow{k_6} \varnothing. \end{aligned} \tag{4.5}$$

The viral infection of a host-cell is initiated with a single molecule of *tem*, where the viral template *tem* denotes the 'active' form of nucleic acids that is involved in the catalytic synthesis of the viral components *gen* and *struct* (reactions R_3 and R_5, respectively). The *gen* component refers to nucleic acids that transport the viral genetic information, e.g., DNA or RNA, which is either processed into the active form *tem* (reaction R_1), or used together with structural proteins *struct* to compose a new viral cell that gets released from the host (reaction R_4). The number of *tem* and *struct* molecules is further regulated by degradation (reactions R_2 and R_6, respectively). The considered values of all rate constants are given in Table 4.2.

Even though the network (4.5) has a relative simple structure, it is capable of resembling realistic scenarios of a viral infection by rendering two steady states; the first representing a successful infection with approximately 20 *gen*, 200 *tem* and 10.000 *struct* molecules, and the second representing a successful rejection with no molecule of any viral component left. However, a pure deterministic model of system (4.5) will never account for these different scenarios, as the first steady state is deterministically stable, whereas the second, absorbing state is deterministically unstable. To study the system behavior correctly, the discrete–stochastic formulation has to be used instead. Unfortunately, the system dynamics lives on a much too large state space to solve the corresponding CME directly (e.g., there are more than 42 million states that 'directly' connect the two steady states), and usually MC realizations are considered instead.

As the level of *struct* is at least two orders of magnitude larger in a successful infection than the levels of *tem* and *gen* (except for some transient phase), we treat *struct* as a continuous species; and regard *tem* and *gen* as discrete species. Accordingly, reactions R_1–R_4 are treated as discrete, stochastic processes; reactions R_5 and R_6 are approximated as continuous, deterministic processes. Since the propensities

TABLE 4.2
Parameter values used for the viral infection kinetics model (4.5), *conversion factor* $\Omega = 1$. *Initially, the system is with probability one in the state with* 1 *tem, and* 0 *gen and struct molecules. The system is simulated for a period of* 200 *days.*

k_1 [1/day]	k_2 [1/day]	k_3 [1/day]	k_4 [1/day]	k_5 [1/day]	k_6 [1/day]
0.025	0.25	1.0	7.5×10^{-6}	1000	2.0

of both continuous reactions, R_5 and R_6, have constant values for a fixed number of *tem* molecules, we further coarse grain the number of continuous processes, i.e., we only associate a continuous process to each subset of discrete states with the same level of *tem*. We use each of these processes to obtain a deterministic approximation of the expected level of *struct* conditioned on the number of *tem* molecules. This further reduces the number of hybrid equations drastically.

In Figs. 4.4 and 4.5 the predicted PDFs of the discrete species *tem* and *gen*, respectively, are shown at four different time points. The corresponding expectations of *struct* for a given number of *tem* are plotted in Fig. 4.6. The results of the hybrid model are in excellent agreement with the approximations obtained by SSA realizations. Furthermore, it can be seen that the conditional expectations $\mathbb{E}_t[\#struct|\#tem]$ remain almost constant during later phase (compare, for instance, the results for $t = 100, 150$ and 200 days in Fig. 4.6), indicating that the continuously approximated processes are already in equilibrium. Instead it is the probability to be in a specific subset of the discrete state space that changes (see Figs. 4.4 and 4.5), and hence the impact of the discrete reactions that results in changes of the partial expectations of *struct* (data not shown). As illustrated in Fig. 4.7, this results in a very accurate prediction of the total expectation of *struct* by the hybrid model, whereas the 'pure' deterministic model fails.

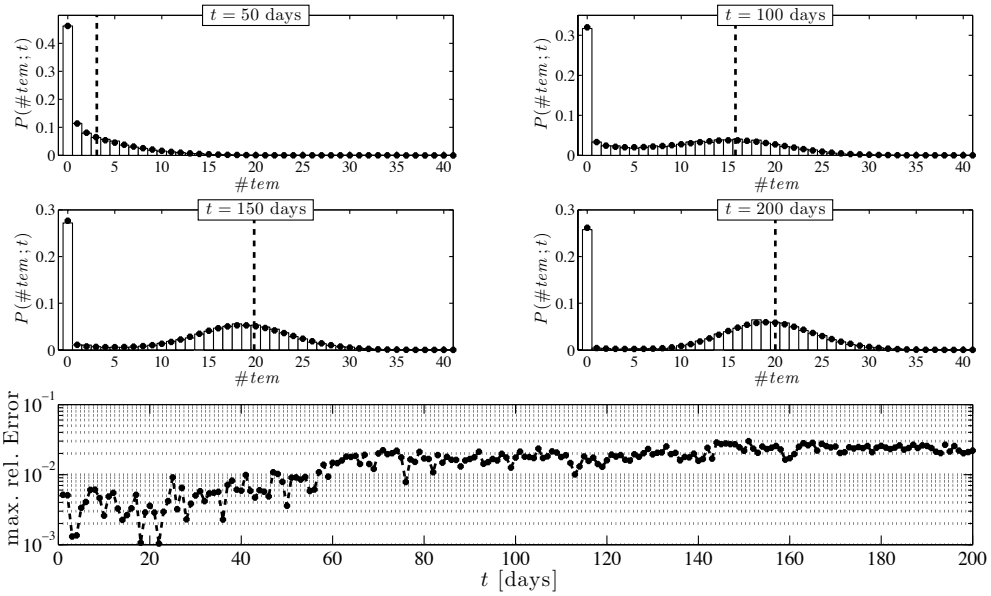

FIGURE 4.4. *Probability distribution of tem at* $t = 50, 100, 150$ *and* 200 *days in the viral infection kinetics model* (4.5) *for parameter values as listed in Table 4.2. The approximations obtained by ten thousand SSA runs are indicated by bars, the numerical solution of the suggested hybrid model is marked by stars, pure deterministic predictions are highlighted by dashed lines. Bottom: The maximal relative (with respect to the SSA results) error of the hybrid solution against time.*

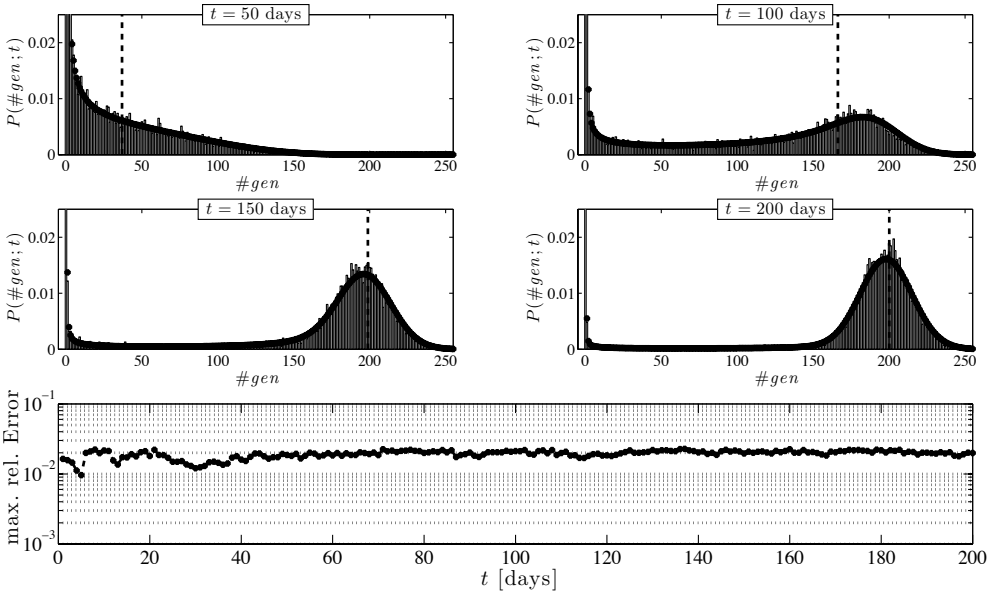

FIGURE 4.5. *Probability distribution of gen at* $t = 50, 100, 150$ *and* 200 *days in the viral infection kinetics model* (4.5) *for parameter values as listed in Table 4.2. The approximations obtained by ten thousand SSA runs are indicated by bars, the numerical solution of the suggested hybrid model is marked by stars, pure deterministic predictions are highlighted by dashed lines. Bottom: The maximal relative (with respect to the SSA results) error of the hybrid solution against time.*

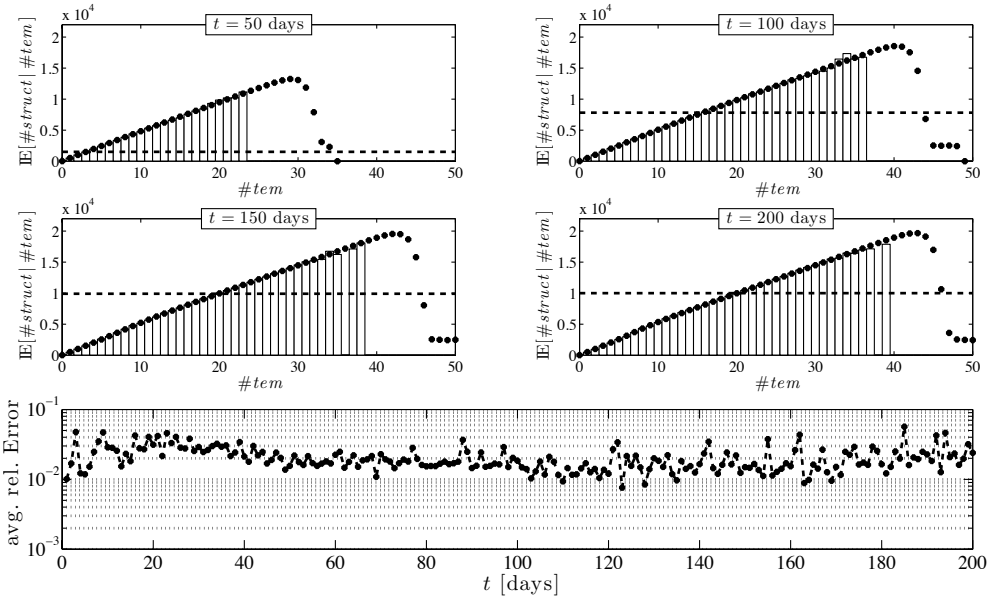

FIGURE 4.6. *Expected values of struct conditioned on the number of tem molecules in system* (4.5) *at* $t = 50, 100, 150$ *and* 200 *days after begin of infection (parameters values listed in Table 4.2). The approximations obtained by ten thousand SSA runs are indicated by bars, the numerical solution of the suggested hybrid model is marked by stars, pure deterministic predictions are highlighted by dashed lines. Bottom: The average relative (with respect to the SSA results) error of the hybrid solution against time.*

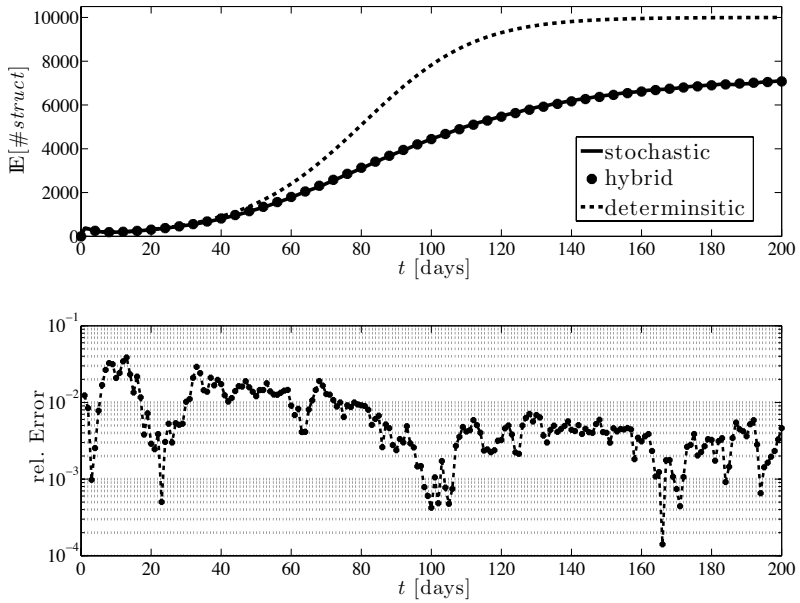

FIGURE 4.7. *Expected value of struct in system* (4.5)*, plotted against the simulated period of* 200 *days (parameter values listed in Table 4.2). The approximation obtained by ten thousand SSA runs is plotted as a straight line, the predicted expectation in the hybrid approach is marked by stars, the pure deterministic prediction is illustrated by a dashed line. Below: The relative error of the hybrid solution with respect to the SSA results against time.*

4.3. Transcription Regulation. To further illustrate the proposed hybrid approach on a more complex system and to study the error resulting from a deterministic approximation of the continuous species, we consider the transcriptional regulatory system published in [35]. The system includes six species that interact through the following ten elementary reactions:

$$
\begin{aligned}
R_1&: \quad mRNA \xrightarrow{k_1} mRNA + M, & R_2&: \quad M \xrightarrow{k_2} \varnothing,\\
R_3&: \quad DNA{\cdot}D \xrightarrow{k_3} DNA{\cdot}D + mRNA, & R_4&: \quad mRNA \xrightarrow{k_4} \varnothing,\\
R_5&: \quad DNA + D \xrightarrow{k_5} DNA{\cdot}D, & R_6&: \quad DNA{\cdot}D \xrightarrow{k_6} DNA + D,\\
R_7&: \quad DNA{\cdot}D + D \xrightarrow{k_7} DNA{\cdot}2D, & R_8&: \quad DNA{\cdot}2D \xrightarrow{k_8} DNA{\cdot}D + D,\\
R_9&: \quad 2M \xrightarrow{k_9} D, & R_{10}&: \quad D \xrightarrow{k_{10}} 2M,
\end{aligned}
\tag{4.6}
$$

where M is a protein (monomer) the can reversibly dimerise (reactions R_9 and R_{10}) to form the transcription factor D (dimer). The DNA template has two different binding sites for D, where DNA denotes the state with both sites free, $DNA{\cdot}D$ the state with D bound at the first site, and $DNA{\cdot}2D$ the state with dimers bound at both

sites. It is assumed that D can occupy the second binding site (reactions R_7 and R_8) only when the first site is already occupied (R_5 and R_6). Transcription only occurs from the state $DNA{\cdot}D$ (reaction R_3), where the dimer is bound to the first binding site of the DNA. The produced messenger RNA, *mRNA*, is translated into protein M (reaction R_1). Both, M and *mRNA*, are subject of degradation (reactions R_2 and R_4).

Based on common knowledge in cell biology, we expect the level of proteins to be much larger than the levels of mRNA and DNA. Therefore, we regard the monomer and dimer forms of the protein, M and D, as continuous species in our hybrid approach, whereas *mRNA* and all DNA forms (*DNA*, $DNA{\cdot}D$ and $DNA{\cdot}2D$) are treated as discrete species. As reactions R_3–R_8 change the levels of discrete species, we treat these reactions as discrete, stochastic processes. Reactions R_1, R_2, R_9 and R_{10} have only impact on the levels of the continuous species (M and D), and are hence approximated as continuous, deterministic processes. Given this partition of the species and reactions, we can formulate and numerically solve the corresponding hybrid equations of the system.

In Fig. 4.8 the evolution of the system as predicted by the hybrid model is compared to the results obtained from ten thousand SSA realizations. The predictions are in excellent agreement, particularly those for the discrete species, as further illustrated in Fig. 4.9. We again emphasize that the results computed with SSA necessarily include an unknown sampling error. While the error in the numerical integration of the proposed hybrid equations is controlled by standard methods, the error made with an indirect MC method is hard to estimate. In the hybrid model, however, an additional error arises from approximating one part of the network continuously and deterministically.

To study the error that results from modeling the two protein forms M and D as continuous species, we decrease the rate constant of *mRNA* synthesis, k_3. A lower level of *mRNA* results in an overall weaker protein synthesis (reaction R_1) and, consequently, lower levels of monomer M and dimer D. As can be seen from the results shown in Figs. 4.10 and 4.11, for instance, the error made with the hybrid model increases in this scenario, which is also intuitively clear. Especially in the later transient phase (after ≈ 10 min), the predictions for the discrete species show a noticeable mismatch to the SSA results.

In Fig. 4.12, we further investigate the dependence of the hybrid solution on the rate of *mRNA* synthesis k_3. As indicated by these results, for higher levels of M and D (i.e., $\varepsilon \to 0$ in the partial scaling of the network) the error resulting from their continuous–deterministic approximation decreases. At the same time, the CPU-time needed to numerically solve the corresponding hybrid equations increases. This is due to the fact that the *mRNA* species is part of the discrete partition. Hence, for higher *mRNA* synthesis rates the state space expands and the number of ODEs increases in the hybrid model, resulting in higher numerical cost to compute its solution. However, a similar effect is also present in the computation of the SSA realizations (Fig. 4.12, upper plot), and, compared to that, the hybrid solution is always computed approximately two orders of magnitude faster (or as fast as one hundred SSA realizations, respectively).

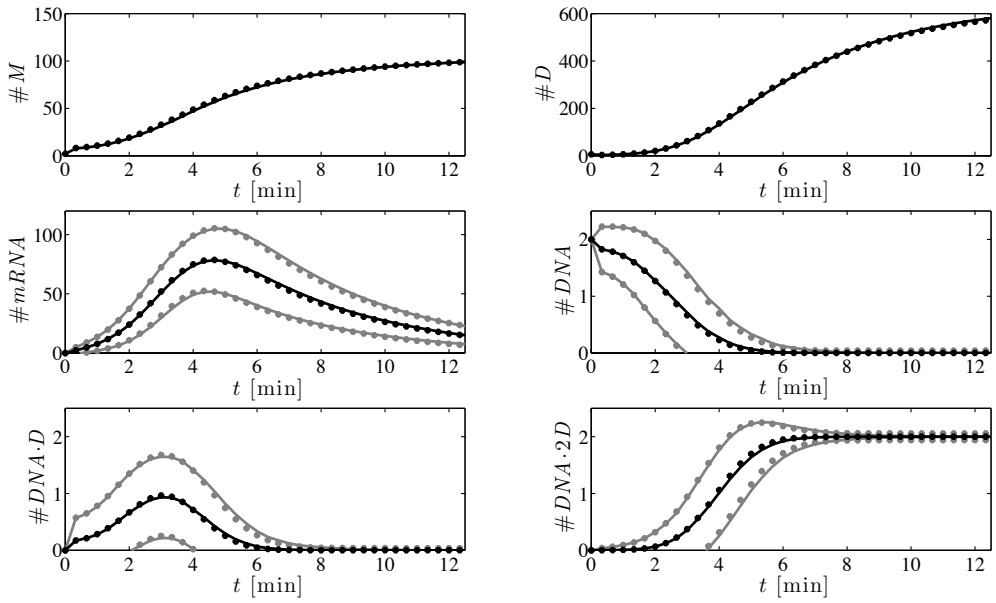

FIGURE 4.8. *Time evolutions of the expected values (black) and standard deviations (gray) in the transcription regulation system* (4.6) *for the parameter values as listed in Table 4.3. The approximations obtained by ten thousand SSA runs are plotted as lines, the numerical solutions of the suggested hybrid model are marked by stars.*

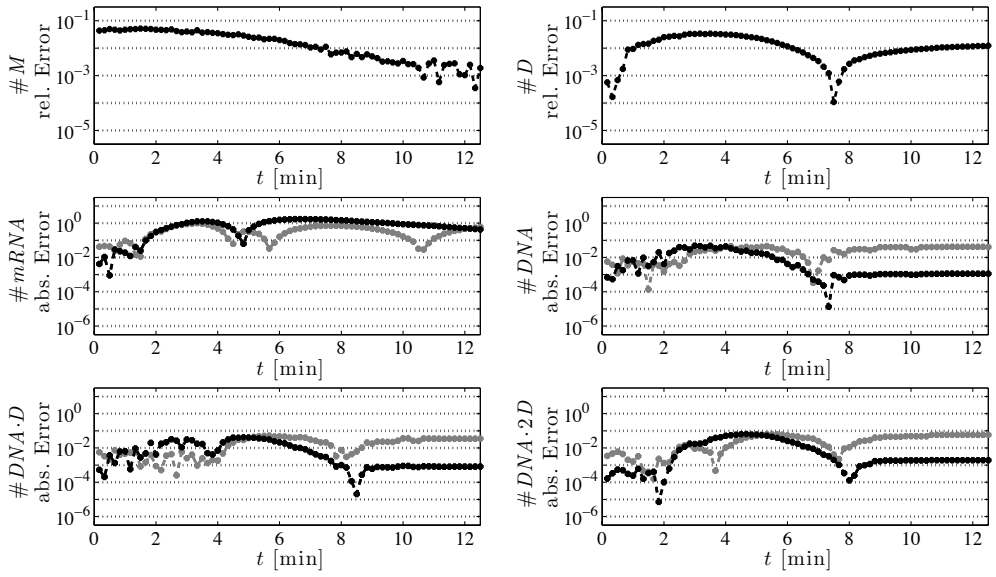

FIGURE 4.9. *Absolute and relative (with respect to the SSA predictions) errors, respectively, of the results presented in Fig. 4.8. Black lines refer to the errors in expected values, gray lines to errors in standard deviations.*

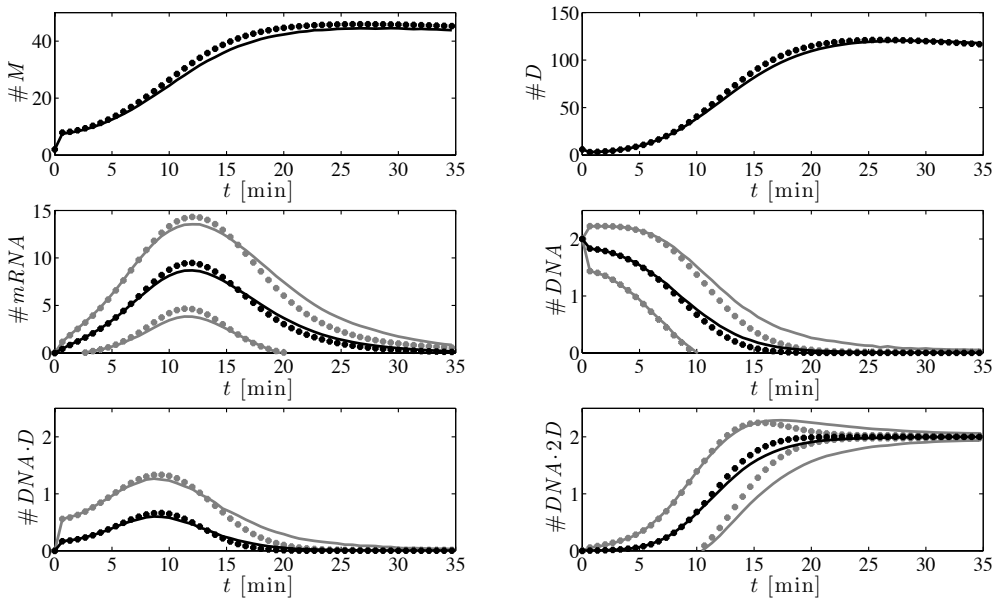

FIGURE 4.10. *Time evolutions of the expected values (black) and standard deviations (gray) in the transcription regulation system* (4.6) *for the parameter values as listed in Table 4.3, but with a ten-fold slower mRNA production (k_3). The approximations obtained by ten thousand SSA runs are plotted as lines, the numerical solutions of the suggested hybrid model are marked by stars.*

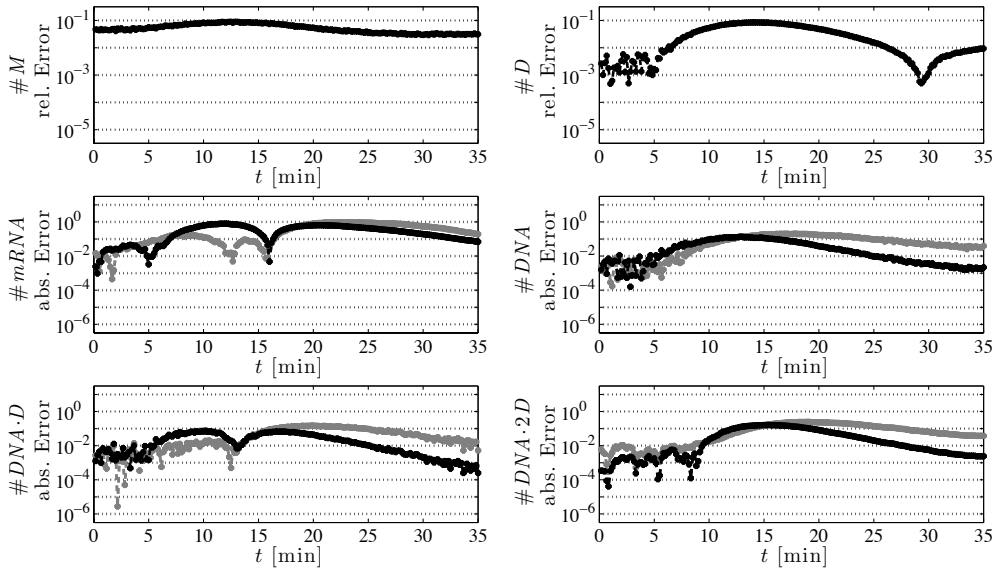

FIGURE 4.11. *Absolute and relative (with respect to the SSA predictions) errors, respectively, of the results presented in Fig. 4.10. Black lines refer to errors in expected values, gray lines to errors in standard deviations.*

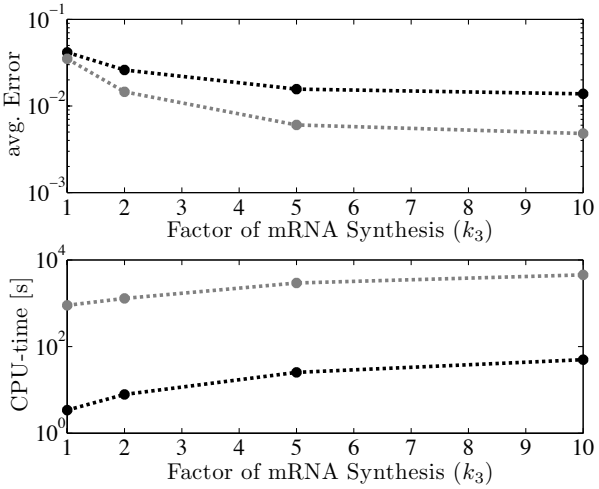

FIGURE 4.12. *Impact of the mRNA synthesis rate (k_3) on the approximation error and the computational costs in the numerical solution of the proposed hybrid model of system* (4.6). *The individual parameter settings (marked by dots) correspond to changes of the original value of k_3 by factors* 1, 2, 5 *and* 10. *For all other parameter the same values as listed in Table 4.3 were used. Top: The average relative error for the predicted expectations of the continuous species M and D (black), and the average absolute error for the predicted expectations of the discrete species DNA, DNA·D and DNA·2D (gray) with respect to the approximations obtained by ten thousand SSA runs. Bottom: Comparison of the required CPU-times to numerically compute the hybrid solution (black) and ten thousand SSA simulations (gray) for each parameter setting.*

TABLE 4.3

Parameter values used for the transcription regulation system (4.6). *These are the original parameter values for the average cell volume $V \approx 1.44 \times 10^{-15}$ l (conversion factor $\Omega = N_A \cdot V \approx 8.64 \times 10^8$ l/mol) as published in [35], except for a* 10*-fold faster mRNA production (k_3). The system is started with probability one in the state with* 2 *M*, 6 *D*, 2 *DNA, and* 0 *mRNA, DNA·D and DNA·2D molecules and simulated for a period of* 35 *min.*

k_1 [1/s]	k_2 [1/s]	k_3 [1/s]	k_4 [1/s]	k_5/Ω [1/s]
0.043	7.0×10^{-4}	0.72[§]	3.9×10^{-3}	0.014
k_6 [1/s]	k_7/Ω [1/s]	k_8 [1/s]	k_9/Ω [1/s]	k_{10} [1/s]
0.48	1.4×10^{-4}	8.8×10^{-12}	0.029	0.5

[§]The results presented in Figs. 4.10 and 4.11 are computed for the original rate of *mRNA* production, corresponding to $k_3 = 0.072\ \mathrm{s}^{-1}$.

5. Conclusion. We used multiscale analysis techniques to derive a novel hybrid model for approximation of the PDF solution of the CME. To this end, we singled out a subspace associated with species of low copy numbers and assumed that its complement subspace is associated with large copy numbers which can be well approximated by a continuous distribution. We exploited the natural Bayesian decomposition $P(X,Z;t) = P(X|Z;t)P(Z;t)$ of the joint PDF into the PDF $P(Z;t)$ on the low copy number space and the conditional distribution $P(X|Z;t)$ on the large copy number space, where the condition is on the discrete states of the low copy number subspace. The hybrid model resulting from multiscale asymptotics based on this Bayesian decomposition couples a CME for $P(Z;t)$ to differential algebraic equations (DAE) for the first moments of $P(X|Z;t)$. Therefore, there has to be a DAE for every discrete state Z of the low copy number CME. Although this may look like a huge complication, the hybrid model solution is expected to be particularly suitable for networks including a few 'discrete' species only since the numerical costs directly scale with the dimension of the CME subspace which is much smaller for the hybrid model than for the original CME. Hence, a direct solution of the proposed hybrid model becomes feasible, which was demonstrated in applications to viral infection kinetics and a transcription regulatory network.

It should be emphasized that the DAEs that govern the evolution of the PDF in the large copy number subspace are *not* identical with the usually expected equations of chemical reaction kinetics; instead they include additional coupling terms resulting from changes in the population of the discrete subspace. Since such kinds of hybrid models have not been investigated before there are some important open questions that will be subject of further research. Let us just mention three of these: Where are the limits of our hybrid model, i.e., if there are species with moderate copy numbers in between low and large copy numbers, when will their fluctuations destroy the approximation quality of the model? In order to decide whether the asymptotic assumptions underlying our hybrid model are valid, how can we estimate the value of ε for a given chemical reaction network? How can one construct an efficient and robust numerical scheme that allows to adaptively change the low copy number subspace on the fly during numerical integration based on some accuracy requirements?

Acknowledgments. Contributions by SM were made possible by DFG funding provided through the Dahlem Research School of Freie Universität Berlin. This work is supported by the DFG Research Center Matheon "Mathematics for Key Technologies" (FZT86) in Berlin.

REFERENCES

[1] H. H. McAdams and A. Arkin, *Stochastic mechanisms in gene expression*, Proc. Natl. Acad. Sci. USA, 94 (1997), pp. 814–819.
[2] H. H. McAdams and A. Arkin, *It's a noisy business! Genetic regulation at the nanomolar scale*, Trends Genet., 15 (1999), pp. 65–69.
[3] M. B. Elowitz, A. J. Levine, E. D. Siggia and P. S. Swain, *Stochastic Gene Expression in a Single Cell*, Science, 297 (2002), pp. 1183–1186.
[4] J. M. Raser and E. K. O'Shea, *Control of Stochasticity in Eukaryotic Gene Expression*, Science, 304 (2004), pp. 1811–1814.
[5] B. Munsky and M. Khammash, *The finite state projection algorithm for the solution of the chemical master equation*, J. Chem. Phys., 124 (2006), 044104.
[6] P. Deuflhard, W. Huisinga, T. Jahnke and M. Wulkow, *Adaptive Discrete Galerkin Methods Applied to the Chemical Master Equation*, SIAM J. Sci. Comput., 30 (2008), pp. 2990–3011.

[7] S. Engblom, *Galerkin Spectral Method applied to the Chemical Master Equation*, Commun. Comput. Phys., 5 (2009), pp. 871–896.
[8] T. Jahnke, *An adaptive wavelet method for the chemical master equation*, SIAM J. Sci. Comput., 31 (2010), pp. 4373–4394.
[9] D. T. Gillespie, *Exact Stochastic Simulation of Coupled Chemical Reactions*, J. Phys. Chem., 81 (1977), pp. 2340–2361.
[10] M. A. Gibson and J. Bruck, *Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels*, J. Phys. Chem. A, 104 (2000), pp. 1876–1889.
[11] K. Burrage, M. Hegland, S. MacNamara and R. B. Sidje, *A Krylov-based Finite State Projection algorithm for solving the chemical master equation arising in the discrete modelling of biological systems*, in MAM 2006: Markov Anniversary Meeting, A. N. Langville and W. J. Stewart, eds., Boson Books, Raleigh, North Carolina, USA, 2006, pp. 21–38.
[12] M. Hegland, C. Burden, L. Santoso, S. MacNamara and H. Booth, *A solver for the stochastic master equation applied to gene regulatory networks*, J. Comput. Appl. Math., 205 (2007), pp. 708–724.
[13] V. Wolf, R. Goel, M. Mateescu and T. A. Henzinger, *Solving the chemical master equation using sliding windows*, BMC Systems Biology, 4 (2010).
[14] T. E. Turner, S. Schnell and K. Burrage, *Stochastic approaches for modelling in vivo reactions*, Comput. Biol. Chem., 28 (2004), pp. 165–178.
[15] C. Schütte, J. Walter, C. Hartmann and W. Huisinga, *An averaging principle for fast degrees of freedom exhibiting long-term correlations*, Multiscale Model. Simul., 2 (2004), pp. 501–526.
[16] D. T. Gillespie, *Approximate accelerated stochastic simulation of chemically reacting systems*, J. Chem. Phys., 115 (2001), pp. 1716–1733.
[17] Y. Cao, D. T. Gillespie and L. R. Petzold, *Adaptive explicit-implicit tau-leaping method with automatic tau selection*, J. Chem. Phys., 126 (2007), 224101.
[18] E. L. Haseltine and J. B. Rawlings, *Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics*, J. Chem. Phys., 117 (2002), pp. 6959–6969.
[19] K. Takahashi, K. Kaizu, B. Hu and M. Tomita, *A multi-algorithm, multi-timescale method for cell simulation*, Bioinformatics, 20 (2004), pp. 538–546.
[20] A. Alfonsi, E. Cancès, G. Turinici, B. D. Ventura and W. Huisinga, *Adaptive simulation of hybrid stochastic and deterministic models for biochemical systems*, ESAIM Proc., 14 (2005), pp. 1–13.
[21] C. V. Rao and A. P. Arkin, *Stochastic chemical kinetics and the quasi-steady state assumption: Application to the Gillespie algorithm*, J. Chem. Phys., 118 (2003), pp. 4999–5010.
[22] Y. Cao, D. T. Gillespie and L. R. Petzold, *The slow-scale stochastic simulation algorithm*, J. Chem. Phys., 122 (2005), 014116.
[23] W. E, D. Liu and E. Vanden-Eijnden, *Nested stochastic simulation algorithms for chemical kinetic systems with multiple time scales*, J. Comput. Phys., 221 (2007), pp. 158–180.
[24] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory*, Springer, New York, 1999.
[25] M. S. Samoilov and A. P. Arkin, *Deviant effects in molecular reaction pathways*, Nat. Biotechnol., 24 (2006), pp. 1235–1240.
[26] T. A. Henzinger, M. Mateescu, L. Mikeev and V. Wolf, *Hybrid Numerical Solution of the Chemical Master Equation*, Proc. 8th Int. Conf. Comput. Methods Syst. Biol. (CMSB'10), (2010), pp. 55–65.
[27] D. T. Gillespie, *A rigorous derivation of the chemical master equation*, Phys. A, 188 (1992), pp. 404–425.
[28] K. Ball, T. G. Kurtz, L. Popovic and G. Rempala, *Asymptotic Analysis of Multiscale Approximations to Reaction Networks*, Ann. Appl. Probab., 16 (2006), pp. 1925–1961.
[29] T. G. Kurtz, *The Relationship between Stochastic and Deterministic Models of Chemical Reactions*, J. Chem. Phys., 57 (1972), pp. 2976–2978.
[30] R. Kubo, K. Matsuo and K. Kitahara, *Fluctuation and Relaxation of Macrovariables*, J. Stat. Phys., 9 (1973), pp. 51–96.
[31] M. I. Dykman, E. Mori, J. Ross and P. M. Hunt, *Large fluctuations and optimal paths in chemical kinetics*, J. Chem. Phys., 100 (1994), pp. 5735–5750.
[32] F. John, *Partial Differential Equations*, 4th edition, Springer, New York, 1981.
[33] E. Vanden-Eijnden, *Fast Communications: Numerical Techniques for Multi-Scale Dynamical Systems with Stochastic Effects*, Comm. Math. Sci., 1 (2003), pp. 385–391.
[34] N. G. de Bruijn, *Asymptotic Methods in Analysis*, Dover, New York, 1981.
[35] J. Goutsias, *Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems*, J. Chem. Phys., 122 (2005), 184102.
[36] R. Srivastava, L. You, J. Summers and J. Yin, *Stochastic vs. Deterministic Modeling of Intracellular Viral Kinetics*, J. Theoret. Biol., 218 (2002), pp. 309–321.