The spatiotemporal master equation: Approximation of reaction-diffusion dynamics via Markov state modeling

Stefanie Winkelmann and Christof Schütte

Citation: J. Chem. Phys. **145**, 214107 (2016); doi: 10.1063/1.4971163

View online: http://dx.doi.org/10.1063/1.4971163

View Table of Contents: http://aip.scitation.org/toc/jcp/145/21

Published by the American Institute of Physics

The spatiotemporal master equation: Approximation of reaction-diffusion dynamics via Markov state modeling

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

(Received 30 August 2016; accepted 15 November 2016; published online 7 December 2016)

Accurate modeling and numerical simulation of reaction kinetics is a topic of steady interest. We consider the spatiotemporal chemical master equation (ST-CME) as a model for stochastic reaction-diffusion systems that exhibit properties of metastability. The space of motion is decomposed into metastable compartments, and diffusive motion is approximated by jumps between these compartments. Treating these jumps as first-order reactions, simulation of the resulting stochastic system is possible by the Gillespie method. We present the theory of Markov state models as a theoretical foundation of this intuitive approach. By means of Markov state modeling, both the number and shape of compartments and the transition rates between them can be determined. We consider the ST-CME for two reaction-diffusion systems and compare it to more detailed models. Moreover, a rigorous formal justification of the ST-CME by Galerkin projection methods is presented. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4971163]

I. INTRODUCTION

A reaction network is a system involving several chemical species undergoing multiple reactions. Depending on the particle concentration and mobility, different mathematical models are appropriate. In the case of rapid diffusion, the spatial position of the particles becomes negligible and the system may be considered as *well-mixed*. Then, the state of the system is defined by the number of particles of each species, and the dynamics are modeled by a continuous-time Markov chain with the reactions described by jumps of the chain.[1] A characterization of the process is given by the chemical master equation (CME) which describes the temporal evolution of the probabilities for the system to occupy each different state. Numerical methods for solving the CME are usually based on Monte Carlo simulations of the underlying Markov jump process, such as Gillespie's stochastic simulation algorithm and its variants.[10,12–15] In the limit of high population concentrations (large copy numbers for all species), the dynamics can be approximated by mass action kinetics,[21] leading to a deterministic system described by ordinary differential equations (ODEs).

In biological applications, the well-mixed assumption of the CME is often inadequate due to spatial inhomogeneities or limited speed of the diffusive motion of the particles. In this case, a higher resolution in space is required, leading to microscopic particle-tracking methods. The most detailed standard model is given by particle-based reaction-diffusion dynamics (PBRD), where all individual particle paths and reaction events are explicitly resolved in time and space. Particles are modeled as points (or spheres) in space undergoing Brownian motion. Bimolecular reactions take place with a certain probability per unit of time when two reactive particles meet within a predefined reaction radius of each other.[2,26,32] An overview of the existing simulation tools for such detailed dynamics is given in Ref. 27. Simulations of PBRD-systems are based on time-discretizations. In each time step, the positions of all particles are advanced, followed by a check-up for adjacent reactive particles whose reaction radii overlap such that a reaction could fire. This scheme becomes inefficient when parts of the system are dense because those parts then continuously stop the clock, i.e., require extremely small discretization time steps.

As an alternative there exist finite volume approaches where the state space is discretized into a collection of nonoverlapping cells, and diffusion is approximated by a continuous-time random walk between the cells.[9,17,18,20] A common choice for the discretization is a uniform Cartesian lattice; however, there also exist approaches with other types of meshes, e.g., consisting of triangles.[8] Bimolecular reactions occur with a fixed probability per unit of time between reactive particles situated in the same cell. The state of the system is given by the number of particles of each species in each cell, and the dynamics are described by the so-called reaction-diffusion master equation (RDME) which is formally a CME but with the states having a spatial interpretation. Diffusion is modeled by first-order reactions, allowing each particle to change its position by switching between cells. In this kind of approach, the question of how to choose the mesh size in order to guarantee a small approximation error is of central importance. In fact, a fine discretization of space does not naturally lead to a high approximation quality because encounters of particles within the same (small) cell become unlikely such that bimolecular reactions are suppressed.[16,18] In order to

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

overcome this problem, the *convergent* RDME has been developed which allows particles to react also when being in nearby cells.[19]

In such a space discretization by a mesh with uniform size of the cells, all areas of the space are equally treated, irrespectively of the structural properties of the reaction-diffusion system. However, in many applications the diffusive and reactive properties of the system naturally exhibit certain structural properties which suggest a *more flexible and non-uniform coarse-graining*. Especially situations of metastable diffusion dynamics propose to coarse-grain space into areas of metastability. Such situations will be considered in this article. An evident example is given by the process of gene expression within a eukaryotic cell where some of the involved reactions (e.g., production of messenger RNA) take place only in the nucleus of the cell while others (e.g., production of proteins) are restricted to the cytoplasm, and transitions between these two compartments are relatively rare in time. In this case, a description on the level of a CME (assuming well-mixed behavior in total space) is obviously not appropriate; however, a split-up of space into two compartments (nucleus and cytoplasm) with diffusive transitions described by jumps might be enough to capture the dynamics.

Although such a coarse-graining with only a few compartments which are chosen in consideration of the dynamical properties of the system seems natural, its mathematical background has not yet been examined. In this article, we describe the theoretical foundation of this approach and present practical methods to infer a reasonable splitting of space as well as corresponding transition rates between the compartments out of experimental data or numerical simulations. The central idea is to apply the theory of Markov state models (MSMs) which proposes a scheme to construct coarse-grained representations of conformational molecular kinetics. The existence of metastable sets within the dynamics can be exploited to provide good approximation properties of the reduced model on long time scales.[28] Instead of conformational dynamics of an individual molecule, we consider the total reaction-diffusion system and approximate the diffusion dynamics of each molecule by a jump process between the metastable compartments. With the jumps understood as first-order reactions and the state of the total system given by the number of particles in each of the compartments, the total dynamics are again characterized by a CME with spatial interpretation, which in this context will be called *spatiotemporal CME* (ST-CME). The ST-CME formally conforms to the RDME, but the underlying coarse-graining of space follows a completely different concept, adapting the natural properties of the process and leading to an incomparably lower number of compartments. Typical compartments are not "small" such that the problem of suppressed second-order reactions, which appears for the RDME in the case of a small mesh size, is circumvented in the setting of the ST-CME.

As a prototypical example, we will investigate reaction-diffusion systems within a eukaryote, where the split-up into nucleus and cytoplasm is quite obvious. However, the theory of MSMs also enables to find number and geometry of compartments in cases where they are not known *a priori*. The resulting simplified models are readily understood and easy to interpret. By inferring the spatial clustering and the transition rates from experimental data, particle-based simulations are completely circumvented, and trajectories of the total system can directly be generated by Gillespie simulations of the derived ST-CME. Even if—for lack of experimental data—the MSM-construction requires simulations of the diffusion dynamics, the numerical complexity is reduced to a large extent because only the motion of an individual particle (and not the total population affected by reaction and diffusion) has to be produced.

In this paper, the ST-CME is viewed as an approximation of microscopic particle-based dynamics with the spatial resolution reduced to the minimum possible degree, still maintaining the characteristic properties of the dynamics. Beside the intuitive, heuristic derivation, we give a concrete mathematical justification showing that the ST-CME results from a Galerkin projection of the corresponding particle-based dynamics. Again inspired by the MSM theory, this permits us to derive explicit formulas for the model parameters depending on the microscopic variables.

In Section II the ST-CME is introduced as a reaction-diffusion model with adapted spatial resolution. The construction of the underlying MSM is described, explaining the procedure of space decomposition and the estimation of transition rates. Two examples are presented in order to illustrate the approach. In Section III, we analytically derive the ST-CME for a reduced reaction system by applying the Galerkin projection method to the corresponding particle-based dynamical system. In the first step we consider a simplified model with not more than two particle species; the more general analysis for larger numbers of species and particles is given in the Appendix.

II. MODELING REACTION-DIFFUSION PROCESSES FOR METASTABLE DIFFUSION

In the following we first review some basic approaches to model reaction-diffusion processes on a stochastic level. The spatiotemporal master equation is presented as a model with an *intermediate spatial scaling*, adapted to the natural properties of the system of interest. Then, in Section II B, we give a short introduction to the theory of Markov state models and apply it to determine the spatial coarse-graining and the jump parameters appearing in the ST-CME. Section II C comprises two illustrative examples to compare the approaches.

A. Modeling reaction-diffusion processes

We consider a set of particles moving in a compact domain $\Omega \subset \mathbb{R}^d$ by continuous diffusion. The particles belong to different species S_l, $l = 1, \ldots, L$, and may interact with each other through different reaction channels R_k, $k = 1, \ldots, K$.

1. Particle-based reaction-diffusion (PBRD)

In the most detailed model of *particle-based reaction-diffusion dynamics*, the motion of every individual particle is

resolved in time and space. Positions of particles are represented as points or spheres undergoing Brownian motion, and bimolecular reactions occur with a certain *microscopic reaction rate* as soon as the particles are located within a predefined *reaction radius* of each other (Doi-model).[2,7,32] These systems cannot be explicitly solved but must be sampled by stochastic realizations. Time is discretized, and for each time step, the positions of all individual particles are advanced according to a given rule of motion. Each advance is followed by a check-up for reactive complexes, i.e., sets of reactive particles whose reaction radii overlap. Given such a complex, the reaction can fire; whether this happens or not is decided randomly based on the microscopic reaction rate. Further methods have been developed in order to include interaction potentials between particles.[26] These permit effects such as space exclusion, molecular crowding, and aggregation to be modeled. The particle-based resolution is appropriate in cases of low particle concentrations and slow diffusion. Spatial inhomogeneities can be taken into account up to a very detailed level. The simulations, however, can be extremely costly often requiring several CPU-years in a single simulation for reaching the biological relevant time scales of seconds. If part of the system is dense, the approach becomes numerically inefficient because reactions would fire in every iteration step. In such a case, a description of the population by concentrations is more appropriate, leading to PDE formulations of the dynamics. For low concentration but rapid diffusion over the total space, the high spatial resolution becomes redundant and the system can be modeled by the chemical master equation (CME).

2. The chemical master equation (CME)

In the case of rapid diffusion in a homogeneous environment, the system can be considered as well-mixed such that only the number of particles of each species is relevant while their spatial position is neglected. Let $\boldsymbol{N}(t) = (\boldsymbol{N}^1(t), \ldots, \boldsymbol{N}^L(t))$ denote the state of the system at time $t \geq 0$, with $\boldsymbol{N}^l(t)$ referring to the number of particles of species S_l at time t. Given the state $\boldsymbol{N}(t) = \boldsymbol{n} = (n^l)_{l=1,\ldots,L} \in \mathbb{N}_0^L$ of the system, a reaction is described by transitions of the form $\boldsymbol{n} \to \boldsymbol{n} + \boldsymbol{\nu}_k$ with the integer vector $\boldsymbol{\nu}_k = \left(\nu_k^1, \ldots, \nu_k^L\right)$ giving the change in the number of particles of each species due to reaction R_k. For example, the chemical reaction $2S_l \to S_{l'}$ is described by the vector $\boldsymbol{\nu}_k$ with $\nu_k^l = -2$, $\nu_k^{l'} = 1$, and zeros elsewhere. For each reaction R_k, a propensity function $\alpha^k(\boldsymbol{n})$ defines the probability per unit of time for the reaction to occur given that $\boldsymbol{N}(t) = \boldsymbol{n}$. 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 particles involved in the reaction: For a unimolecular reaction by species l, it holds $\alpha^k(\boldsymbol{n}) = \gamma^k \cdot n^l$; for a second-order reaction by two species l, l', $l \neq l'$, one has $\alpha^k(\boldsymbol{n}) = \gamma^k \cdot n^l \cdot n^{l'}$.[1,22] With $P(\boldsymbol{n}, t) = \mathbb{P}(\boldsymbol{N}(t) = \boldsymbol{n} | \boldsymbol{N}(0) = \boldsymbol{n}_0)$ denoting the probability for the system to be in state $\boldsymbol{n}$ at time t, the dynamics are characterized by the *chemical master equation*

$$\frac{dP(\boldsymbol{n}, t)}{dt} = \sum_{k=1}^{K} \left(\alpha^k(\boldsymbol{n} - \boldsymbol{\nu}_k) P(\boldsymbol{n} - \boldsymbol{\nu}_k, t) - \alpha^k(\boldsymbol{n}) P(\boldsymbol{n}, t)\right).$$

Realizations of the underlying Markov jump process can be created by Gillespie simulations.

3. The spatiotemporal chemical master equation (ST-CME)

There exist many applications where the well-mixed assumption required for the chemical master equation is not fulfilled. Instead, diffusion of particles might be limited by local barriers in space, or the environment Ω offers inhomogeneities with respect to reaction propensities. An example that will be investigated in Section II C describes the setting of diffusion within a eukaryotic cell which naturally decomposes into two compartments, the nucleus and the cytoplasm. By a reduced permeability of the nuclear membrane, the diffusive flow through the cell is restricted such that a description by a chemical master equation would fail. Within each of the two compartment, however, the dynamics can indeed be considered as well-mixed, which motivates to consider the chemical master equation on the level of compartments.

Such situations (in which the space Ω exhibits a particular structure with respect to the diffusion and reaction properties) will be considered here. More precisely, we make the following central assumption.

Assumption 2.1. There is a decomposition of Ω into compartments Ω_r, $r = 1, \ldots, M$, such that

1. transitions between compartments are rare (metastability);
2. within each compartment Ω_r, diffusion is rapid compared to the reaction (well-mixed-property);
3. within each compartment, the reaction rates are constant, i.e., independent of the position (homogeneity).

The first point of Assumption 2.1 delivers the basis for the construction of a Markov state model for the diffusive part of the reaction-diffusion network. With the diffusion processes of all particles being metastable with respect to the same space decomposition, they can be approximated by Markov jump processes on the fixed set of compartments $\{\Omega_r : r = 1, \ldots, M\}$. Within the ST-CME, the jumps are treated as first order reactions. By the second and third points of Assumption 2.1, we make sure that within each compartment, the reaction dynamics can accurately be described by a chemical master equation.

Let $N_r^l(t)$ denote the number of particles of species l in compartment r at time t. We take $\boldsymbol{N}_r(t) = (N_r^1(t), \ldots, N_r^L(t))$ to denote the state of present species in compartment r. The total state of the system at some point in time t is given by the matrix $\boldsymbol{N}(t) = (\boldsymbol{N}_r(t))_{r=1,\ldots,M} \in \mathbb{N}_0^{M,L}$. Changes of the state are induced by diffusive transitions (jumps) between the compartments and by chemical reactions.

A jump of a particle of species l from compartment r to compartment $s \neq r$ is described by a transition of the form

$$\boldsymbol{N} \to \boldsymbol{N} + \boldsymbol{E}_s^l - \boldsymbol{E}_r^l,$$

where $\boldsymbol{E}_r^l$ is a matrix whose elements are all zero except the entry (r, l) which is one. Let λ_{rs}^l denote the jump rate for each

individual particle of species l. Since all particles are assumed to diffuse independently of each other, the total probability per unit of time for a jump of species l from compartment r to s at time t is given by $\lambda_{rs}^l N_r^l(t)$.

For each of the reactions R_k, let again $\boldsymbol{\nu}_k = \left(\nu_k^1, \ldots, \nu_k^L\right)$ describe the change in the number of copies of all species induced by this reaction. Reaction R_k occurring in the rth compartment refers to the transition $\boldsymbol{N}_r(t) \to \boldsymbol{N}_r(t) + \boldsymbol{\nu}_k$. In order to specify where the reaction takes place, the vector $\boldsymbol{\nu}_k$ is multiplied by a column vector $\boldsymbol{e}_r$ with the value 1 at entry r and zeros elsewhere. This gives a matrix $\boldsymbol{e}_r\boldsymbol{\nu}_k$ whose rth row is equal to $\boldsymbol{\nu}_k$ while all other rows contain zeros. With this notation, the change in the total state $\boldsymbol{N}$ due to reaction R_k taking place in compartment r is given by

$$\boldsymbol{N} \to \boldsymbol{N} + \boldsymbol{e}_r\boldsymbol{\nu}_k.$$

The propensity for such a reaction to occur is given by the function $\alpha_r^k(\boldsymbol{n})$ denoting the probability per unit of time for reaction R_k to occur in compartment r given that $\boldsymbol{N}_r(t) = \boldsymbol{n}_r$, i.e., it depends $\alpha_r^k(\boldsymbol{n})$ only on the values of $\boldsymbol{n}$ referring to compartment r. Note that the rate constants which determine the reaction propensities may depend on the compartment such that—in contrast to the CME—spatial inhomogeneities in the reaction propensities can be taken into account.

As in the setting of the CME, let

$$P(\boldsymbol{n}, t) = \mathbb{P}\left(\boldsymbol{N}(t) = \boldsymbol{n} | \boldsymbol{N}(0) = \boldsymbol{n}_0\right)$$

be the probability that the process is in state $\boldsymbol{n}$ at time t given an initial state $\boldsymbol{N}(0) = \boldsymbol{n}_0$. The spatiotemporal chemical master equation (ST-CME) is then given by

$$\begin{aligned}\frac{dP(\boldsymbol{n}, t)}{dt} = &\sum_{r=1}^{M} \sum_{s \neq r} \sum_{l=1}^{L} \left(\lambda_{sr}^l (n_s^l + 1) P(\boldsymbol{n} + \boldsymbol{E}_s^l - \boldsymbol{E}_r^l, t) - \lambda_{rs}^l n_r^l P(\boldsymbol{n}, t)\right) \\ &+ \sum_{r=1}^{M} \sum_{k=1}^{K} \left(\alpha_r^k(\boldsymbol{n} - \boldsymbol{e}_r\boldsymbol{\nu}_k) P(\boldsymbol{n} - \boldsymbol{e}_r\boldsymbol{\nu}_k, t) - \alpha_r^k(\boldsymbol{n}) P(\boldsymbol{n}, t)\right), \end{aligned} \tag{1}$$

where the first line refers to the diffusive part (the transitions between the compartments), while the second line describes the chemical reactions within the compartments.

Numerical realizations of the ST-CME can again be created using the Gillespie method. Compared to particle-based simulations in which the trajectory of each single particle has to be calculated, the simulations of the ST-CME are computationally less expensive. However, due to the spatial interpretation, the state space of the ST-CME is larger than the one of the CME.

Remark 2.2. A reaction-diffusion master equation (RDME) is formally identical to (1) but different regarding the spatial discretization. In a RDME the underlying space Ω is discretized by a structured or unstructured mesh leading to a large number of mesh cells so that the first line of (1) in case of a RDME model is based on a discretization of the Laplacian operator underlying the diffusion process.[8,11,20]

To determine a convenient space discretization as well as suitable jump rates is the fundamental issue for the representation of a reaction-diffusion system in terms of a ST-CME. It can be handled by constructing a Markov state model for the space of motion.

B. Spatial discretization via Markov state modeling

The theory of Markov state models (MSMs) is a well established tool to approximate complex molecular kinetics by Markov chains on a discrete partition of the configuration space.[4,23–25,28,30] The underlying process is assumed to exhibit a number of metastable sets in which the system remains for a comparatively long period of time before it switches to another metastable set. A MSM represents the effective dynamics of such a process by a Markov chain that jumps between the metastable sets with the transition rates of the original process. Two main steps are essential for this approach: the identification of metastable compartments and the determination of transition rates between them. Both steps are extensively studied in the literature, providing concrete practical instructions.[3–5] Although usually interpreted in terms of conformational dynamics, the approaches can directly be transferred to the kinetic dynamics of individual particles within a reaction-diffusion system. Within the ST-CME, jumps and reactions are uncoupled from each other and memoryless in the sense that a reaction taking place does not influence the jump propensity of any species, and the reaction rates within each compartment are independent of the previous evolution of the system. Thereby, a separate analysis of the diffusion properties is justified. In the following, we apply the MSM-theory to the pure diffusion process of an individual particle, fading out its interaction with other particles by reactions.

1. Spatial coarse-graining

Assuming that the space of motion consists of metastable areas where the reaction-diffusion dynamics can be considered to be well-mixed, the first task is to identify the location and shape of these areas. While in certain applications the spatial clustering might be obvious (like, e.g., for the process of gene expression), others will require an analysis of available data (either from simulations or experiment) to define the number and geometry of compartments.

The MSM approach to perform this step of space decomposition is based on a two-stage process exploiting both geometry and kinetics.[4,5,29] Starting with a geometric clustering of space into small volume elements, the kinetic

relevance of this clustering is checked for the given simulation data, followed by a merging of elements which are kinetically strongly connected. As a concrete example we shortly sketch the *automatic state decomposition algorithm* introduced in Ref. 5: The iteration alternates between a *split* step (splitting the space into microstates) and a *lump* step (lumping microstates together to get macrostates with maximum possible metastability). Each macrostate is subsequently again fragmented into microstates, which are lumped again to redefine the macrostates. By this alternation, the boundaries between the macrostates are iteratively refined. The splitting is based on geometric criteria, while the lumping considers kinetic aspects, taking metastability as a measure.

To illustrate the relevance of both geometric and kinetic aspects for the space discretization, one can consider the example of nuclear membrane transport: By the reduced permeability of the nuclear membrane, the cell decomposes into two metastable regions (nucleus and cytoplasm). A small geometric distance between two points is the first indication for belonging to the same compartment; however, this would also group together close points which are located on different sides of the nuclear membrane and are thereby kinetically distant. While other clustering methods might fail to capture the kinetic properties of the system under consideration, the MSM approach incorporates the relevant kinetic information and allocates such points into different compartments.[3]

The theoretical foundation of Markov state modeling comprises the spectral analysis of the transfer operator which describes how the process propagates functions in state space over time. The dominant eigenvalues refer to long implied time scales, with the associated eigenvectors giving information about the metastable areas in space and the aggregate transition dynamics between them.

The algorithms used for spatial clustering by MSMs require the availability of large amount of data sets. Transitions between metastable sets are infrequent events which even in long (but finite) trajectories might be small in number. Fortunately, for MSMs there exist methods to validate the model and to quantify the statistical error, allowing to rerun simulations in case of high uncertainty (so called adaptive sampling).[4,23,24] The development of standards for the number and length of simulations needed for a statistically reliable modeling is in progress.[5]

2. Estimation of transition rates

Once a suitable space partitioning for the problem under consideration is given (known in advance or determined by Markov state modeling as described before), the rates for transitions between the compartments have to be estimated. The theory of MSM also delivers concrete practical techniques to determine the transition rates out of trajectory data. Applied to the diffusion dynamics of a particle in space, the rate estimation can shortly be summarized as follows.

Let the space Ω of motion be decomposed into compartments Ω_r, $r = 1, \dots, M$, with $\bigcup_{r=1}^{M} \Omega_r = \Omega$ and $(\mathrm{int}_{\Omega_r}) \cap (\mathrm{int}_{\Omega_s}) = \emptyset$ for all $r \neq s$ where int_{Ω_r} denotes the interior of Ω_r. The diffusion of a single particle in Ω is given by a homogeneous, time-continuous Markov process $(X_t)_{t\geq 0}$ with X_t denoting the position of the particle in Ω at time $t \geq 0$. We assume that (X_t) has a unique, positive invariant probability measure μ on Ω. Fixing a lag time $\tau > 0$, the MSM process is a discrete-time Markov chain $(\tilde{X}_n)_{n\in\mathbb{N}}$ on $\{1, \dots, M\}$ with transition probabilities

$$P_{rs} = \mathbb{P}_\mu(X_\tau \in \Omega_s | X_0 \in \Omega_r), \tag{2}$$

where the subscript μ indicates that $X_0 \sim \mu$. The Markov process $(\tilde{X}_n)$ serves as an approximation of the process $(\hat{X}_n)$ defined by $\hat{X}_n = r \Leftrightarrow X_{n\cdot\tau} \in \Omega_r$ which, in general, is itself not Markovian but has a memory. For details about the approximation quality see Ref. 24.

Given a trajectory $(x_0, x_1, \dots, x_N)$ of the diffusion process $(X_t)_{t\geq 0}$ with $x_n = X_{n\cdot\Delta t}$ for a fixed time step $\Delta t > 0$ (given from experimental data or separated simulation), the transition matrix P is estimated by counting the transitions between the compartments: Let the lag time $\tau = l\Delta t$ ($l \in \mathbb{N}$) be a multiple of Δt and set

$$\hat{P}_{rs} = \frac{C_{rs}}{\sum_{s'} C_{rs'}}, \quad C_{rs} = \sum_{n=0}^{N-l} \chi_r(x_n)\chi_s(x_{n+l}), \tag{3}$$

with χ_r denoting the indicator function of Ω_r. Then $\hat{P}$ is a maximum-likelihood estimator for the transition matrix P.[23,25]

In order to turn from discrete time to continuous time within the coarse-grained setting, the matrix estimation is repeated for a range of lag times τ and the resulting transition matrices P_τ are used to determine appropriate transition rates $\lambda_{rs} > 0$ between the compartments in the setting of a continuous-time Markov jump process. More precisely, we aim for a rate matrix $\Lambda = (\lambda_{rs})_{r,s=1,\dots,M}$ with $\lambda_{rs} > 0$ for $r \neq s$ and $\lambda_{rr} = -\sum_{s\neq r} \lambda_{rs}$ which describes the time-evolution of a memoryless system by the master equation

$$\frac{dp_t^T}{dt} = p_t^T \Lambda, \tag{4}$$

where $p_t(r)$ denotes the probability for the diffusion process to be in compartment r at time t.[31] With p_0 denoting the initial distribution, the solution to (4) is given by $p_t^T = p_0^T e^{\Lambda t}$ which suggests the relation $P_\tau \approx e^{\Lambda\tau}$. On short time scales, the "true" process $(X_t)_{t\geq 0}$ cannot be accurately approximated by such a memoryless system because recrossings between the compartments induce memory effects.[6] For τ not too small, however, it indeed holds[4,28]

$$\frac{1}{\tau}\log(P_\tau) \approx \text{constant}$$

such that we can set $\Lambda = \frac{1}{\tau}\log(P_\tau)$. The entries of the matrix Λ are the jump rates appearing in the ST-CME. The discretization error arising with the coarse-graining can even be quantitatively bounded.[4]

Remark 2.3. The estimation of the rate matrix can be made more accurate by forcing it to fulfill certain side constraints like detailed balance or the keep of an equilibrium distribution which is known *a priori*.[4]

C. Illustrative examples

For an illustration we consider two different reaction-diffusion systems within the cell of a eukaryote. The cell naturally decomposes into two compartments, the nucleus and the

cytoplasm, with the nuclear membrane exhibiting a reduced permeability which induces a metastability of the diffusion dynamics. The cellular membrane is assumed to be impermeable such that the cell's exterior is out of interest. Particles are assumed to move independently and not to influence each other in their crossing behavior. Reactions between particles can only take place if the particles are in the same compartment. The cell is modeled in two dimensions as two concentric circles representing the membranes of nucleus and cytoplasm.

The first system under consideration is a quite artificial binding-unbinding network which is used to compare the dynamics characterized by the ST-CME to the underlying particle-based dynamics. The congruence of the *quantitative* evolution of the systems (in terms of average population numbers) is revealed. Then, the more applied setting of gene expression is investigated, comparing the *qualitative* properties of the ST-CME system to those induced by a RDME with Cartesian grid as given in Ref. 20. Both examples illustrate the ability of the ST-CME to reproduce the dynamics of more detailed and complex model descriptions.

1. Binding and unbinding

As a first example we formulate a very general reaction-diffusion system consisting of three species $S_1 = A$, $S_2 = B$, $S_3 = C$ which diffuse within the cell and undergo the binding and unbinding reactions

$$R_1 : A + B \to C, \qquad R_2 : C \to A + B. \tag{5}$$

We assume that the binding reaction takes place only in the cytoplasm while unbinding is restricted to the nucleus. This separation of reactions, combined with the metastability of diffusion due to a reduced permeability of the nuclear membrane, obviously keeps the process from being well-mixed in total space, such that a description by a CME is not appropriate. Minimum spatial information—as it is given by a ST-CME—is needed to capture the characteristics of the dynamics. In the following, we compare the ST-CME dynamics to appropriate particle-based dynamics where the diffusion of each particle is modeled by a process of Brownian motion with reflecting boundary conditions at the cellular membrane and limited flux through the nuclear membrane; see Figure 1 for an illustration.

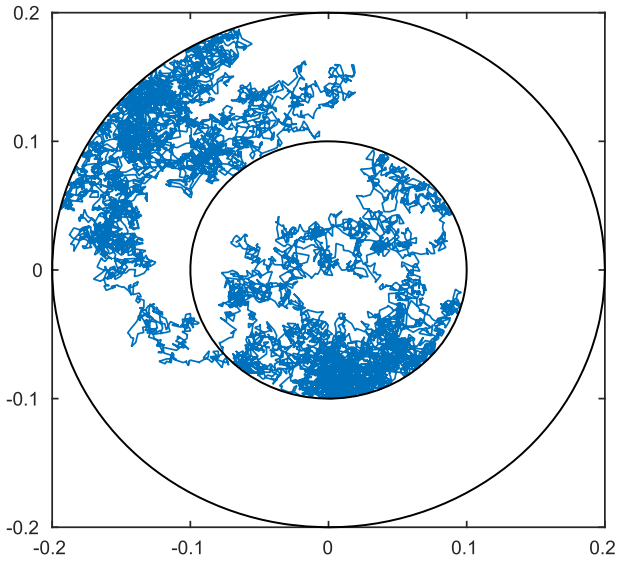

FIG. 1. *Brownian motion on disk with metastable compartments*. Brownian particle moving on a disk (cell) containing a nucleus. The nuclear membrane exhibits a reduced permeability, rejecting most of the crossing trials of the trajectory.

Given the microscopic parameters of the system (i.e., diffusion constants, microscopic reaction rates, and reaction radius for the Doi-model) as well as an initial population (we choose 5 A-particles uniformly distributed in the nucleus and 5 B-particles uniformly distributed in the cytoplasm), numerical realizations of the particle-based reaction-diffusion system are produced by the Euler-Maruyama method. For all time steps of a realization, the number of particles of each species in each compartment is recorded. Then, a MSM for the diffusion process of each species is constructed as described in Section II B, yielding the jump rates between the two compartments necessary for the ST-CME. The average binding time for two particles A and B within the cytoplasm is estimated numerically in order to take its inverse for the macroscopic binding rate. Gillespie simulations of the resulting ST-CME give realizations of the total system on the spatially discretized level.

For both approaches (particle-based dynamics and ST-CME), the average evolution of the population over time is estimated by Monte Carlo samples. In Figure 2 the results are shown, with an apparent consistency of both models.

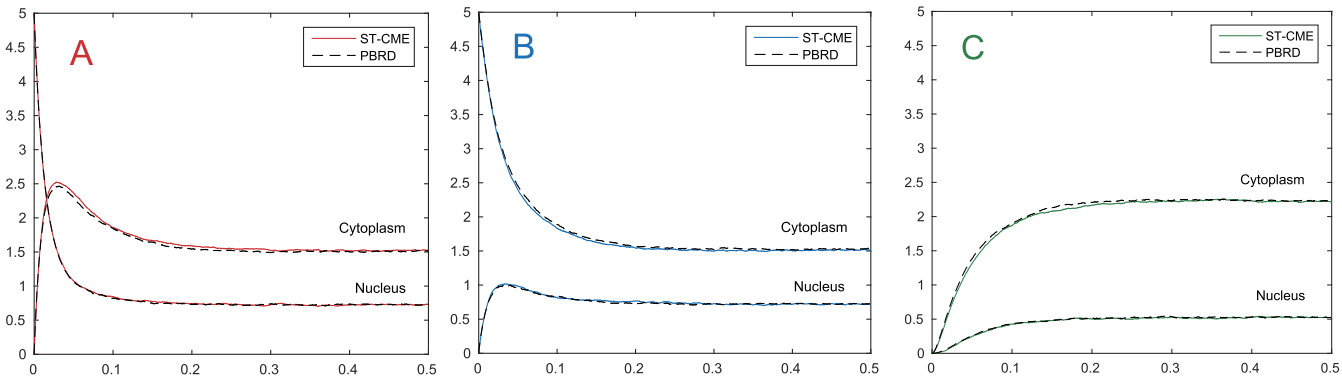

FIG. 2. Mean number of particles in conformation A (left), B (middle), and C (right) in the nucleus and the cytoplasm, respectively, for separated reactions.

This binding-unbinding network will again be investigated in Section III where a formal derivation of the ST-CME is derived.

2. Gene expression

A more applied setting is given by the process of gene expression by which information encoded in DNA is used for the production of functional proteins. This process is a complex network of chemical reactions. In eukaryotes some of the involved reactions take place in the nucleus while others are restricted to the cytoplasm. On a very coarse level, the process can be described as follows. In a first step, the gene is *transcribed* into messenger RNA (mRNA); this takes place in the nucleus. The mRNA molecule is then exported to the cytoplasm where it is *translated* into proteins. The proteins are imported into the nucleus where they can repress the gene, meaning that its functionality in transcription is interrupted. Both mRNA and protein molecules are degradable in both compartments.

Some of these steps contain themselves a complex series of chemical events. For example, the crossing of the nuclear membrane by mRNA and protein molecules requires the assistance of import and export receptors because the pores of the nuclear membrane are too small for the molecules to pass through by diffusion. For the ST-CME, such details are omitted and the crossing of the nuclear membrane is described by a single jump event with the jump rate chosen in a way to take account of all the involved steps. For many questions of interest, a modeling on such a coarse level is appropriate.

In Ref. 20 the process of gene expression is analyzed in terms of the reaction-diffusion master equation. With the cell and its nucleus also modeled by concentric circles, the space is discretized by a regular 37×37 Cartesian mesh. Transitions of mRNA and proteins between nucleus and cytoplasm are modeled by binding and unbinding to export (or import) receptors which are assumed to be uniformly distributed throughout nucleus and cytoplasm at certain steady-state concentrations. After binding to the receptor, the molecule can freely diffuse across the membrane. After this transition, the molecule has to unbind from the receptor before it can undergo further reactions. For the ST-CME the three steps (binding to receptor, transition, and unbinding from receptor) are merged and replaced by a jump event. In Ref. 20, simulations of the total system show a periodicity in the total number of nuclear proteins. This periodicity is reproducible by the ST-CME with only two compartments as will be demonstrated now.

The process of gene expression perfectly complies with the requirements of the ST-CME: Two of the reactions (transcription and translation) are restricted to certain areas of space, namely, nucleus or cytoplasm. Transitions between these two compartments are of central relevance for the total dynamics. Due to the complexity of membrane transition, the diffusion dynamics can be considered as metastable. Within each compartment, local information is redundant and the dynamics can be considered as well-mixed.

More precisely, the following model is plausible. The *transcription* in the nucleus is given by the reaction

$$DNA \xrightarrow{\alpha_r^1} DNA + mRNA,$$

with the index r referring to compartment Ω_r and $\alpha_1^1 > 0$ in the nucleus ($=\Omega_1$) while $\alpha_2^1 = 0$ in the cytoplasm ($=\Omega_2$). *Translation* of mRNA into proteins P is given by

$$mRNA \xrightarrow{\alpha_r^2} mRNA + P,$$

with $\alpha_1^2 = 0$ in the nucleus and $\alpha_2^2 > 0$ in the cytoplasm. *Repressing* of DNA by a protein molecule refers to the reaction

$$DNA + P \xrightarrow{\alpha_r^3} DNA^*,$$

where DNA^* denotes the repressed DNA. The repressing can be reversed by

$$DNA^* \xrightarrow{\alpha_r^4} DNA + P.$$

Both mRNA and protein are degradable which is described by

$$mRNA \xrightarrow{\alpha_r^5} \emptyset$$

and

$$P \xrightarrow{\alpha_r^6} \emptyset,$$

respectively. Transitions between nucleus and cytoplasm (possible for mRNA and protein molecules) are described by jumps with jump rates $\lambda_{12}^{mRNA} > 0$ (export of mRNA) and $\lambda_{21}^{P} > 0$ (import of protein).

In the RDME-formulation given in Ref. 20 most of these reactions are decomposed into several steps. Then, the protein production is shown to occur in bursts, which is due to its repressing function: Given a large number of proteins, the repressing of the gene becomes very likely. If the gene is repressed, no new mRNA is transcribed, and the existing mRNA-molecules quickly vanish because of the large decay rate α_r^5. The translation into proteins is interrupted, and the protein population simply decays over time. With a smaller protein population, the repressing of the gene becomes less likely and transcription can start again. Such a periodic evolution of the number of proteins exists in our coarse-grained model; see Figure 3 which looks much like Figure 8 in Ref. 20. The

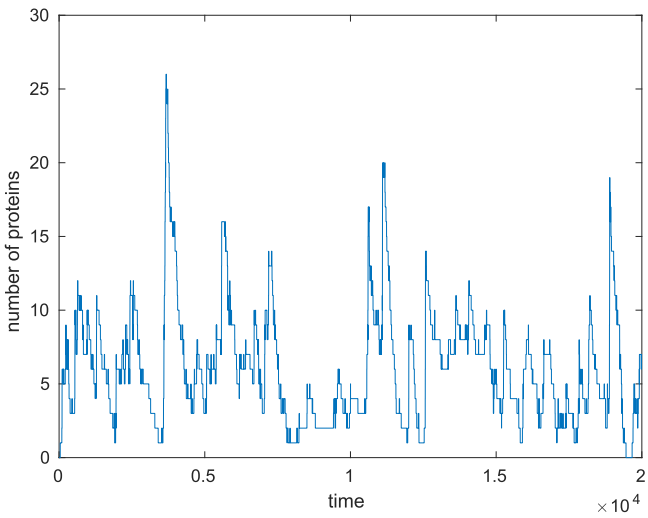

FIG. 3. Number of nuclear proteins over time in one realization of the ST-CME. The protein production occurs in bursts.

chosen reaction rates roughly follow those given in Ref. 20; however, without any claim to reproduce reality, we simply intend to maintain the fundamental qualitative behavior of the system. Concretely, we set $\alpha_1^1 = 0.05$ (transcription in nucleus), $\alpha_2^2 = 0.5$ (translation in cytoplasm), $\alpha_1^3 = 0.01$ (repressing in nucleus), $\alpha_1^4 = 0.01$ (reversal of repressing in nucleus), $\alpha_2^5 = 0.3$ (degradation of *mRNA* in the cytoplasm), $\alpha_r^6 = 0.0025$ (degradation of protein in both compartments), and $\lambda_{12}^{mRNA} = \lambda_{21}^{P} = 0.4$ (jump rates).

III. FORMAL DERIVATION OF THE ST-CME FOR A REDUCED MODEL

In Section II B we presented the theory of Markov state models as a practical tool to significantly coarsen the diffusion dynamics by finding suitable space decompositions as well as rates for transitions between the compartments. Beside the practical and algorithmic details, MSMs also provide a well understood mathematical background, with the coarse-grained dynamics arising from Galerkin projections of the original process.[28]

In this section, we apply the technical concepts and formally derive the ST-CME from a given particle-based model by means of a Galerkin approximation of the associated evolution equations. To this end, we study a reduced model consisting of two particles, one of species A and one of species B, that diffuse independently in space Ω and undergo the binding reaction $A + B \to C$ with a position-independent rate γ^1_{micro} when getting closer than $\varepsilon_1 > 0$. The resulting C-particle diffuses in space and can unbind again with a fixed rate γ^2_{micro}, see the reaction system given in (5). In this setting, the state space of the ST-CME reduces to states $\boldsymbol{n} \in \mathbb{N}_0^{M,3}$ of either the form $\boldsymbol{n} = \boldsymbol{E}_r^1 + \boldsymbol{E}_s^2$—which refers to finding an A-particle in compartment Ω_r and a B-particle in compartment Ω_s while species C is absent—or $\boldsymbol{n} = \boldsymbol{E}_r^3$ referring to the existence of a C-particle in compartment Ω_r, while A and B are absent.

This is a totally over-simplified case that nevertheless allows to perform the formal derivation of the ST-CME in all details. As one can see in the Appendix, the approach can be transferred to larger populations (particle and species numbers).

A. The particle-based model

As for the particle-based dynamics we make the following assumptions. Given A and B undergoing a reaction, the resulting C-particle is placed at the position $x \in \Omega$ of the A-particle. After an unbinding reaction, the particles A and B are placed within a ball $B_{\varepsilon_2}(z)$ of radius $\varepsilon_2 > 0$ around the position $z \in \Omega$ of the preceding C-particle, with the positions x and y of A and B is chosen independently of each other, both uniformly distributed on $B_{\varepsilon_2}(z)$.

We denote by $p(x,y,t)$ the probability density for the particles A and B to be unbound and located at x and y, respectively, at time $t \geq 0$, and by $\tilde{p}(z,t)$ the probability density for the particles to be bound with the C-particle located at z at time $t \geq 0$. With this notation, it holds

$$\int_{\Omega^2} p(x,y,t)\,dxdy + \int_{\Omega} \tilde{p}(z,t)\,dz = 1$$

for all times $t \geq 0$.

The time evolution of the two functions p and $\tilde{p}$ is given by a coupled system of differential equations

$$\begin{aligned}\partial_t p(x,y,t) &= (L_1 p)(x,y,t) + (L_2 p)(x,y,t)\\ &\quad - \gamma^1_{\text{micro}}\,\phi_{\varepsilon_1}(x,y)p(x,y,t)\\ &\quad + \frac{\gamma^2_{\text{micro}}}{|B_{\varepsilon_2}|^2}\int \phi_{\varepsilon_2}(x,z)\phi_{\varepsilon_2}(y,z)\,\tilde{p}\,(z,t)\;dz, \quad (6)\end{aligned}$$

$$\begin{aligned}\partial_t \tilde{p}(z,t) &= (L_3\tilde{p})(z,t) + \gamma^1_{\text{micro}}\int \phi_{\varepsilon_1}(z,y)\,p\,(z,y,t)\,dy\\ &\quad - \gamma^2_{\text{micro}}\,\tilde{p}(z,t), \quad (7)\end{aligned}$$

where L_1, L_2, L_3 denote the generators of the diffusion processes of species A, B, C (in flat spaces without energy barriers these are Laplacian operators), respectively, and ϕ_{ε_k} $(k = 1, 2)$ is given by

$$\phi_{\varepsilon_k}(x,y) = \Phi\left(\frac{|x-y|}{\varepsilon_k}\right),$$

where Φ denotes the indicator function of the ball $B_1(0)$. In (6), outflow is induced by the binding reaction in cases where the positions x and y are ε_1-close to each other (2nd line), while inflow is induced in case of an existing C-particle with its location being able to produce the given positions x and y after unbinding (3rd line). In (7), binding induces inflow (2nd line) where the position of the former A-particle has to be at the given position z of the resulting C-particle, while unbinding induces outflow (3rd line), see the given placement assumptions.

B. Galerkin projection

The dynamics described by the coupled Equations (6) and (7) have a full spatial resolution, breaking down the exact positions of the particles in space. For the ST-CME the information must be coarsened to the level of compartments. This is done by projecting the dynamics onto a suitable low-dimensional ansatz space.

We consider a partition of Ω into subsets Ω_r, $r = 1,\dots,M$, and denote by χ_r the indicator function of Ω_r. That is, χ_r is 1 inside of Ω_r and 0 outside. Let $\langle f, g\rangle$ be the usual L^2-scalar product of functions f and g that depend on Ω, i.e., $\langle f, g\rangle = \int_\Omega f(x)g(x)dx$. If the generators L_1, L_2, L_3 do not denote flat diffusion but diffusion in an energy landscape, then the scalar product must be weighted with the respective product invariant measure, see Ref. 28. Define

$$\mu_r = \langle \mathbf{1}, \chi_r\rangle,$$

where $\mathbf{1}$ is the constant 1-function on Ω. The Galerkin projection $Q : L^2(\Omega) \to D$ onto the finite-dimensional ansatz space $D = \text{span}\{\chi_r : r = 1,\dots,M\}$ is given by

$$Qv = \sum_{r=1}^{M} \frac{1}{\mu_r}\,\langle \chi_r, v\rangle\,\chi_r, \quad v \in L^2(\Omega).$$

We further use the functions

$$\chi_{rs}(x,y) = \chi_r(x)\chi_s(y), \quad r,s = 1,\dots,M,$$

as a partition of unity of $\Omega \times \Omega$. Parallel to the definitions before, we set

$$\mu_{rs} = \langle \mathbf{1}, \chi_{rs} \rangle,$$

where in this context $\langle \cdot, \cdot \rangle$ refers to the L^2-scalar product on $\Omega \times \Omega$, and

$$Qv = \sum_{r,s=1}^{M} \frac{1}{\mu_{rs}} \langle \chi_{rs}, v \rangle \chi_{rs}, \quad v \in L^2\left(\Omega^2\right),$$

as a projection onto $D = \operatorname{span}\{\chi_{rs} : r,s = 1,\dots,M\}$. By definition, it holds

$$\mu_{rs} = \mu_r \mu_s.$$

Applying the projection to the density functions p and $\tilde{p}$ gives the ansatz

$$(Qp)(x,y,t) = \sum_{r,s=1}^{M} p_{rs}(t)\chi_{rs}(x,y),$$

$$(Q\tilde{p})(z,t) = \sum_{r=1}^{M} \tilde{p}_r(t)\chi_r(z),$$

with time dependent coefficients $p_{rs}(t)$ and $\tilde{p}_r(t)$. Inserting into (6) and (7) yields

$$\begin{aligned}\sum_{r,s=1}^{M} \partial_t p_{rs}(t)\chi_{rs}(x,y) = &\sum_{r,s=1}^{M} p_{rs}(t)\left(L_1\chi_{rs}(x,y) + L_2\chi_{rs}(x,y)\right) \\ &- \gamma^1_{\text{micro}} \sum_{r,s=1}^{M} p_{rs}(t)\phi_{\varepsilon_1}(x,y)\chi_{rs}(x,y) \\ &+ \frac{\gamma^2_{\text{micro}}}{|B_{\varepsilon_2}|^2} \sum_i \tilde{p}_i(t) \int \phi_{\varepsilon_2}(x,z)\phi_{\varepsilon_2}(y,z) \\ &\cdot \chi_r(z)\,dz, \end{aligned} \tag{8}$$

$$\begin{aligned}\sum_{r=1}^{M} \partial_t \tilde{p}_r(t)\chi_r(z) = &\sum_{r=1}^{M} \tilde{p}_r(t) L_3 \chi_r(z) \\ &+ \gamma^1_{\text{micro}} \sum_{r,s=1}^{M} p_{rs}(t) \int \phi_{\varepsilon_1}(z,y)\chi_{rs}(z,y)\,dy \\ &- \gamma^2_{\text{micro}} \sum_{r=1}^{M} \tilde{p}_r(t)\chi_r(z). \end{aligned} \tag{9}$$

The local coordinates x, y, z of the particles still appear as arguments. In order to eliminate them, the coupled Equations (8) and (9) are themselves projected onto the given ansatz space, which refers to taking local averages of the terms for each combination of compartments. This automatically delivers the formal relation between the microscopic and the macroscopic model parameters. In fact, we define the macroscopic reaction rate constant for $r, s, r' = 1,\dots,M$ by the averages

$$\gamma^1_{rs} := \gamma^1_{\text{micro}} \cdot \frac{1}{\mu_{rs}} \int \chi_r(x)\chi_s(y)\phi_{\varepsilon_1}(x,y)\,dx\,dy,$$

$$\gamma^2_{rsr'} := \frac{\gamma^2_{\text{micro}}}{|B_{\varepsilon_2}|^2} \cdot \frac{1}{\mu_r} \int \chi_r(x)\chi_s(y)\chi_{r'}(z)\phi_{\varepsilon_2}(x,z)\phi_{\varepsilon_2}(y,z) \cdot dx\,dy\,dz,$$

which depend on the microscopic reaction rate and the reaction radius. Similarly, the jump rates are given by

$$\lambda^l_{rs} := \frac{1}{\mu_r} \int \chi_s(x) L_l \chi_r(x)\,dx = \frac{1}{\mu_r} \langle \chi_s, L_l \chi_r \rangle, \quad l = 1,2,3.$$

Since L_l, l = 1,2, only acts on x_l, we have that

$$\langle \chi_{rs}, L_1 \chi_{r's'} \rangle = \begin{cases} \mu_s \langle \chi_r, L_1 \chi_{r'} \rangle = \mu_{r's}\lambda^1_{r'r} & \text{for } s = s', \\ 0 & \text{otherwise,} \end{cases}$$

$$\langle \chi_{rs}, L_2 \chi_{r's'} \rangle = \begin{cases} \mu_r \langle \chi_s, L_2 \chi_{s'} \rangle = \mu_{rs'}\lambda^2_{s's} & \text{for } r = r', \\ 0 & \text{otherwise.} \end{cases}$$

With this observation and the notations from above we get, by multiplication of (8) with $\langle \chi_{rs}, \cdot \rangle$ and by multiplication of (9) with $\langle \chi_r, \cdot \rangle$, the system

$$\begin{aligned}\mu_{rs}\partial_t p_{rs}(t) = &\sum_{r'=1}^{M} \lambda^1_{r'r} \cdot \mu_{r's} p_{r's}(t) + \sum_{s'=1}^{M} \lambda^2_{s's} \cdot \mu_{rs'} p_{rs'}(t) \\ &- \gamma^1_{rs} \cdot \mu_{rs} p_{rs}(t) + \sum_{r'=1}^{M} \gamma^2_{rsr'} \cdot \mu_{r'} \tilde{p}_{r'}(t), \end{aligned} \tag{10}$$

$$\begin{aligned}\mu_r \partial_t \tilde{p}_r(t) = &\sum_{r'=1}^{M} \lambda^3_{r'r} \cdot \mu_{r'} \tilde{p}_{r'}(t) \\ &+ \sum_{s=1}^{M} \gamma^1_{rs} \cdot \mu_{rs} p_{rs}(t) - \gamma^2_{\text{micro}} \cdot \mu_r \tilde{p}_r(t), \end{aligned} \tag{11}$$

which does not contain any local coordinates, anymore. The system of Equations (10) and (11) actually can be understood as a spatiotemporal master equation, given the assumption that reactions are also possible between particles suited in different compartments. With respect to the notation given in (1), we have the relation

$$P(\boldsymbol{n},t) = \begin{cases} \mu_{rs} \cdot p_{rs}(t) & \text{for } \boldsymbol{n} = \boldsymbol{E}^1_r + \boldsymbol{E}^2_s, \\ \mu_r \cdot \tilde{p}_r(t) & \text{for } \boldsymbol{n} = \boldsymbol{E}^3_r, \\ 0 & \text{otherwise,} \end{cases}$$

where $P(\boldsymbol{n},t)$ is the probability to find the system in state $\boldsymbol{n}$ at time t, see Section II A.

In order to reveal the parallelism to the ST-CME given in (1), we assume that such "mixed" reactions are suppressed, i.e., reactions take place only between particles located within the same compartment, and the product of a reaction is placed in the compartment of the reagents. This refers to setting $\gamma^1_{rs} = 0$, $\gamma^2_{rsr'} = 0$ for all mixed (non constant) combinations of indices. For comparatively large compartments and small $\varepsilon_1, \varepsilon_2$, all these values are close to zero for mixed indices by definition, anyway. For the reaction rate functions appearing in (1), we then obtain

$$\alpha^1_r(\boldsymbol{n}) = \begin{cases} \gamma^1_{\text{micro}} \frac{|B_{\varepsilon_1}|}{\mu_r} & \text{for } \boldsymbol{n} = \boldsymbol{E}^1_r + \boldsymbol{E}^2_r, \\ 0 & \text{otherwise,} \end{cases}$$

$$\alpha^2_r(\boldsymbol{n}) = \begin{cases} \gamma^2_{\text{micro}} & \text{for } \boldsymbol{n} = \boldsymbol{E}^3_r, \\ 0 & \text{otherwise.} \end{cases}$$

Now, the equations simplify to

$$\mu_{rs}\partial_t p_{rs}(t) = \sum_{r'=1}^{M} \lambda^1_{r'r} \cdot \mu_{r's} p_{r's}(t) + \sum_{s'=1}^{M} \lambda^2_{s's} \cdot \mu_{rs'} p_{rs'}(t) - \delta_{rs}\gamma^1_{\text{micro}} \frac{|B_{\varepsilon_1}|}{\mu_r} \cdot \mu_{rs} p_{rs}(t) + \delta_{rs}\gamma^2_{\text{micro}} \cdot \mu_r \tilde{p}_r(t),$$

$$\mu_r \partial_t \tilde{p}_r(t) = \sum_{r'=1}^{M} \lambda^3_{r'r} \cdot \mu_{r'} \tilde{p}_{r'}(t) + \gamma^1_{\text{micro}} \frac{|B_{\varepsilon_1}|}{\mu_r} \cdot \mu_{rr} p_{rr}(t) - \gamma^2_{\text{micro}} \cdot \mu_r \tilde{p}_r(t),$$

where δ_{rs} is the Kronecker delta. Noting that by definition it holds $\lambda^l_{rr} = -\sum_{r' \neq r} \lambda^l_{rr'}$ for all $l = 1, 2, 3$ this is just a reformulation of the ST-CME (1).

Remark 3.1. The available theory of Markov state models shows that one may also use ansatz functions χ_r that are different from indicator functions as long as they form a partition of unity of Ω.[28,30] That is, along these lines we can also derive versions of the ST-CME where spatial discretization is *not* based on sets/compartments but on smooth functions, e.g., committor functions for core sets as in Ref. 30 or radial basis functions as in meshless discretization.[33] This increases the flexibility of our approach significantly and is known to allow for superior approximation quality if recrossing scenarios are important.[28]

IV. CONCLUSION

We introduced the spatiotemporal master equation as a framework for modeling reaction-diffusion dynamics in situations where the diffusion process exhibits metastability. By the construction of a Markov state model for the diffusive part, the structural properties of the dynamics are preserved and diffusion is replaced by a jump process on a comparatively small set of compartments. Although in the literature usually related to conformational dynamics of an individual molecule, the theory of MSM is a powerful tool for finding adequate space decompositions and estimating the related transition parameters in a reaction-diffusion system of several molecules. In application, the method can directly process experimental data such that particle-based simulations are completely circumvented.

For the ST-CME approach to by valid, the well-mixed assumption does not have to apply for the total dynamics (as it does for the CME) but only on the level of compartments. As for the spatial resolution the ST-CME lies between the CME which excludes all spatial information and more detailed models like the RDME (where space is discretized by meshes) or particle-based reaction-diffusion systems (where the exact positions of all particles are retraced). The same is true for the numerical effort induced by simulations of the respective systems. We compared the models for two reaction-diffusion systems within a eukaryotic cell and showed that the ST-CME is able to replicate the more detailed models regarding quantitative and qualitative properties.

Describing a reaction-diffusion system which behaves well-mixed in spatial subareas by a ST-CME is an intuitive approach. We showed that this approach has a clear mathematical foundation, with the ST-CME arising from a Galerkin projection of the corresponding particle-based dynamical system onto the subspace spanned by appropriate ansatz functions. In the standard case, these ansatz functions are the indicator functions of the compartments, and the ST-CME is a combination of jump processes between compartments and the reaction part. However, other families of ansatz functions may provide superior approximation properties. The analysis given in the Appendix shows that, in principle, the ansatz functions used for the Galerkin approach could not only discretize the spatial domain but also the particle number space. A detailed analysis of how to choose the ansatz functions is not given here but will have to be a part of future research.

Another topic of interest is the combination of the different models in situations where parts of a system are well-mixed, while others require a particle-based resolution. Then, the ST-CME has to be coupled to the stochastic dynamics of some particles that are individually tracked in certain areas of space. In contrast, also parts of the system could show high density allowing for a description by ODE's which leads to a stochastic-deterministic evolution equation comprising a ST-CME. The development of such hybrid approaches to handle multi-scale reaction-diffusion dynamics is the subject of ongoing work.

ACKNOWLEDGMENTS

This research has been partially funded by Deutsche Forschungsgemeinschaft (DFG) through Grant No. CRC 1114.

APPENDIX: GENERALIZED FORMAL DERIVATION OF THE ST-CME

In Section III we formally derived the ST-CME from a given particle-based model of binding and unbinding with not more than one particle of each species. We will here generalize the analysis to systems with larger populations: Several particles of species A and B diffuse independently in space Ω and undergo the binding reaction $A + B \to C$ with a position-independent rate γ^1_{micro} when getting closer than $\varepsilon_1 > 0$. The resulting C-particles diffuse in space and can unbind again with a fixed rate γ^2_{micro}.

1. Particle numbers

Subsequently we assume that we never have more than n particles of every individual type in our system. That is, if a, b, c denote the numbers of molecules of types A, B, C in the system, then

$$(a, b, c) \in \mathcal{N}, \quad \mathcal{N} = \{(a, b, c) : a, b, c \leq n, a + b + 2c \leq 2n\}.$$

2. Notation

In addition to the given space of motion Ω, we introduce the artificial *void* state V indicating that the particle is absent. For every individual particle in the system, the position then comes from $\mathcal{S} = V \times \Omega$. The positions of the n particles of type A are denoted by

$$X = (x_1, \ldots, x_n) \in \mathcal{S}^n,$$

and those of B and C by

$$Y = (y_1, \ldots, y_n) \in \mathcal{S}^n, \qquad Z = (z_1, \ldots, z_n) \in \mathcal{S}^n,$$

such that the state of the entire system is given by

$$(X, Y, Z) \in \mathcal{S}^{3n}.$$

In the following we consider only those (X,Y,Z) with

$$\left(\sum_i \chi_\Omega(x_i), \sum_i \chi_\Omega(y_i), \sum_i \chi_\Omega(z_i) \right) \in \mathcal{N}.$$

3. Placement assumption

Given the positions $x_i \in \Omega$ and $y_j \in \Omega$ of an A- and a B-particle undergoing a reaction, the resulting C-particle is placed at the position x_i of the A-particle. After an unbinding reaction, the particles A and B are placed within a ball $B_{\varepsilon_2}(z)$ of radius ε_2 around the position $z_k \in \Omega$ of the preceding C-particle, with the positions x_i and y_j of A and B chosen independently of each other, both uniformly distributed on $B_{\varepsilon_2}(z_k)$.

The index of a C-particle resulting from a binding reaction is chosen to be the first index of a void state in Z. Equivalently, the positions of the A- and B-particles resulting from unbinding are inserted into X and Y at the respective first void entry. The insertion of a position $x \in \Omega$ at index k into state $X \in \mathcal{S}^n$ is denoted by

$$X_k^+(x) = (x_1, \ldots, x_{k-1}, x, x_{k+1}, \ldots, x_n),$$

while setting position x_k in state $X \in \mathcal{S}^n$ to void is denoted by

$$X_k^- = (x_1, \ldots, x_{k-1}, V, x_{k+1}, \ldots, x_n).$$

Equivalent notations are chosen for Y and Z.

4. The particle-based dynamics

By $p(X, Y, Z, t)$ we denote the probability density function of finding the system in state $(X, Y, Z) \in \mathcal{S}^{3n}$ at time t. For integration we use the notation, for example,

$$\int_{\mathcal{S}} p(x_1, \ldots, x_n, Y, Z, t) dx_1 = p(V, x_2, \ldots, x_n, Y, Z, t) + \int_\Omega p(x_1, \ldots, x_n, Y, Z, t) dx_1.$$

Let L_{xi}, L_{yj}, L_{zk} denote the generators of the diffusion processes of species A, B, C acting on coordinate x_j, y_j, and z_k on Ω, respectively. Then we define

$$L_A p(x_1, \ldots, x_n, Y, Z, t) = \sum_{i=1}^n \chi_\Omega(x_i) L_{x_i} p(x_1, \ldots, x_n, Y, Z, t).$$

The time evolution of the function p is given by the following differential equation:

$$\begin{aligned}
\partial_t p(X, Y, Z, t) = {} & (L_A + L_B + L_C) p(X, Y, Z, t) \\
& - \gamma^1_{\text{micro}} \sum_{i,j=1}^n \chi_\Omega(x_i) \chi_\Omega(y_j) \phi_{\varepsilon_1}(x_i, y_j) p(X, Y, Z, t) \\
& + \gamma^1_{\text{micro}} \sum_{k=1}^n \prod_{l=1}^k \chi_\Omega(z_l) \sum_{i,j=1}^n \chi_V(x_i) \chi_V(y_j) \int_\Omega \phi_{\varepsilon_1}(z_k, y) p\left(X_i^+(z_k), Y_j^+(y), Z_k^-, t\right) dy \\
& - \gamma^2_{\text{micro}} \sum_{k=1}^n \chi_\Omega(z_k) p(X, Y, Z, t) \\
& + \frac{\gamma^2_{\text{micro}}}{|B_{\varepsilon_2}|^2} \sum_{k=1}^n \chi_V(z_k) \sum_{i,j=1}^n \prod_{l=1}^i \chi_\Omega(x_l) \prod_{l=1}^j \chi_\Omega(y_l) \cdot \int_\Omega \phi_{\varepsilon_2}(x_i, z) \phi_{\varepsilon_2}(y_j, z) p\left(X_i^-, Y_j^-, Z_k^+(z), t\right) dz, \qquad \text{(A1)}
\end{aligned}$$

where ϕ_ε is given by

$$\phi_\varepsilon(x_1, x_2) = \Phi\left(\frac{|x_1 - x_2|}{\varepsilon} \right),$$

with Φ denoting the indicator function of the ball $B_1(0)$.

Explanation:

- The first line in Equation (A1) refers to the independent diffusion of all particles.
- The second line refers to the outflow from state (X, Y, Z) induced by the binding reaction.
- The third line refers to the inflow induced by the binding reaction of the form

$$\left(X_i^+(z_k), Y_j^+(y), Z_k^-\right) \quad \to \quad (X, Y, Z).$$

For the state (X,Y,Z) to result from this reaction, the entry z_k of Z has to be located in Ω and must fulfill $z_l \in \Omega$ for all $l < k$ because the new positions are assumed to be located at the first index of a void state in Z. The actual positions x_i and y_j of X and Y have to be void, whereas before they were given by z_k and y with y located in an ε_1-environment around z_k.

- The 4th line refers to the outflow from (X,Y,Z) by unbinding.
- The 5th line refers to the inflow induced by unbinding of the form

$$\left(X_i^-, Y_j^-, Z_k^+(z)\right) \to (X, Y, Z).$$

Here, the actual position z_k must be void, while the positions x_i and y_j both have to be in Ω with no preceding void-entries (again because the new positions resulting from unbinding are placed at the respective first void-entries). Both x_i and y_j must be ε_2-close to the z-particle they result from. As both positions are chosen uniformly distributed in the B_{ε_2}-ball around z, the unbinding rate is divided by its size squared.

5. Galerkin projection

a. Ansatz functions

We consider a partition of Ω into subsets $\Omega_1, \ldots, \Omega_m$ and define the indicator function for the system to have exactly $a_r \in \{0, \ldots, n\}$ molecules of type A in set $\Omega_r \subset \Omega$,

$$\chi_{a_r}(X|\Omega_r) = \begin{cases} 1 & \text{if } \sum_{i=1}^n \chi_r(x_i) = a_r, \\ 0 & \text{otherwise,} \end{cases}$$

where χ_r is the indicator function of Ω_r. Then, the indicator function for having exactly a_r molecules of type A and b_r molecules of type B and c_r molecules of type C in set Ω_r ($a_r, b_r, c_r \in \{0, \ldots, n\}$) is given by

$$\chi_{a_r,b_r,c_r}(X, Y, Z|\Omega_r) = \chi_{a_r}(X|\Omega_r)\chi_{b_r}(Y|\Omega_r)\chi_{c_r}(Z|\Omega_r).$$

For the total system we use the ansatz functions

$$\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(X, Y, Z) = \prod_{r=1}^m \chi_{a_r,b_r,c_r}(X, Y, Z|\Omega_r),$$

where $\boldsymbol{a} = (a_r)_{r=1,\ldots,m}$, $\mathbf{b} = (b_r)_{r=1,\ldots,m}$, $\mathbf{c} = (c_r)_{r=1,\ldots,m}$ with $a_r, b_r, c_r \in \{0, \ldots, n\}$ and

$$\left(\sum_{r=1}^m a_r, \sum_{r=1}^m b_r, \sum_{r=1}^m c_r\right) \in \mathcal{N}.$$

The allowed index set of the three vector-indices of our ansatz functions will be denoted $\mathcal{I}$. It is important to notice that these ansatz functions do not overlap,

$$\langle \chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}, \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} \rangle = \mu_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}\, \delta_{\boldsymbol{a}\tilde{\boldsymbol{a}}}\delta_{\boldsymbol{b}\tilde{\boldsymbol{b}}}\delta_{\boldsymbol{c}\tilde{\boldsymbol{c}}}, \qquad \mu_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} > 0,$$

but form a partition of unity of the accessible state space.

b. Galerkin ansatz

We consider the Galerkin projection $Q : L^2\left(\mathcal{S}^{3n}\right) \to D$ onto the finite-dimensional ansatz space $D = \text{span}\{\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} : (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) \in \mathcal{I}\}$ given by

$$Qv = \sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\in\mathcal{I}} \frac{1}{\mu_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}} \langle \chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}, v\rangle\, \chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}, \quad v \in L^2\left(\mathcal{S}^{3n}\right),$$

with

$$\mu_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} = \langle \mathbf{1}, \chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}\rangle$$

and

$$\langle u, v\rangle = \int_{\mathcal{S}} \ldots \int_{\mathcal{S}} u(X, Y, Z)v(X, Y, Z)dx_1 \ldots dx_n dy_1 \ldots dy_n dz_1 \ldots dz_n.$$

The ansatz for $p(X,Y,Z,t)$ now is

$$(Qp)(X, Y, Z, t) = \sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\in\mathcal{I}} p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t)\, \chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(X, Y, Z),$$

with

$$p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t) = \frac{1}{\mu_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}} \langle \chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}, p(\cdot, t)\rangle.$$

Inserting into (A1) gives

$$\begin{aligned}
\sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\in\mathcal{I}} \partial_t p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(X, Y, Z) = &\sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\in\mathcal{I}} p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t)(L_A + L_B + L_C)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(X, Y, Z) \\
&- \gamma^1_{\text{micro}} \sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\in\mathcal{I}} p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t) \sum_{i,j=1}^n \chi_\Omega(x_i)\chi_\Omega(y_j)\phi_{\varepsilon_1}(x_i, y_j)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(X, Y, Z) \\
&+ \gamma^1_{\text{micro}} \sum_{(\boldsymbol{a},\mathbf{b},\mathbf{c})\in\mathcal{I}} p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t) \sum_{i,j,k=1}^n \prod_{l=1}^k \chi_\Omega(z_l)\chi_V(x_i)\chi_V(y_j) \\
&\cdot \int_\Omega \phi_{\varepsilon_1}(z_k, y)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}\left(X_i^+(z_k), Y_j^+(y), Z_k^-\right) dy \\
&- \gamma^2_{\text{micro}} \sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\in\mathcal{I}} p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t) \sum_{k=1}^n \chi_\Omega(z_k)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(X, Y, Z) \\
&+ \frac{\gamma^2_{\text{micro}}}{|B_{\varepsilon_2}|^2} \sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\in\mathcal{I}} p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t) \sum_{i,j,k=1}^n \chi_V(z_k) \prod_{l=1}^i \chi_\Omega(x_l) \prod_{l=1}^j \chi_\Omega(y_l) \\
&\cdot \int_\Omega \phi_{\varepsilon_2}(x_i, z)\phi_{\varepsilon_2}(y_j, z)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}\left(X_i^-, Y_j^-, Z_k^+(z)\right) dz.
\end{aligned}$$

Multiplication by $\left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, \cdot \right\rangle$ gives

$$\begin{aligned}
\mu_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} \cdot \partial_t p_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}(t) = & \sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\in\mathcal{I}} \left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, (L_A + L_B + L_C)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \right\rangle \cdot p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t) \\
& - \gamma^1_{\text{micro}} \sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\in\mathcal{I}} \left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, \sum_{i,j=1}^{n} \chi_\Omega(x_i)\chi_\Omega(y_j)\phi_{\varepsilon_1}(x_i,y_j)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \right\rangle \cdot p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t) \\
& + \gamma^1_{\text{micro}} \sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\in\mathcal{I}} \left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, \sum_{i,j,k=1}^{n} \prod_{l=1}^{k} \chi_\Omega(z_l)\chi_V(x_i)\chi_V(y_j) \right. \\
& \left. \cdot \int_\Omega \phi_{\varepsilon_1}(z_k,y)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}\left(X_i^+(z_k), Y_j^+(y), Z_k^-\right) dy \right\rangle \cdot p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t) \\
& - \gamma^2_{\text{micro}} \sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\in\mathcal{I}} \left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, \sum_{k=1}^{n} \chi_\Omega(z_k)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \right\rangle \cdot p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t) \\
& + \frac{\gamma^2_{\text{micro}}}{|B_{\varepsilon_2}|^2} \sum_{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})\in\mathcal{I}} \left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, \sum_{i,j,k=1}^{n} \chi_V(z_k) \prod_{l=1}^{i} \chi_\Omega(x_l) \prod_{l=1}^{j} \chi_\Omega(y_l) \right. \\
& \left. \cdot \int_\Omega \phi_{\varepsilon_2}(x_i,z)\phi_{\varepsilon_2}(y_j,z)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}\left(X_i^-, Y_j^-, Z_k^+(z)\right) dz \right\rangle \cdot p_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(t).
\end{aligned}$$

Due to the scalar product structure, most of the summands vanish. For the remaining summands, we choose the following notations:

$$\alpha^1(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) := \gamma^1_{\text{micro}} \cdot \frac{1}{\mu_{a,b,c}} \int \sum_{i,j} \chi_\Omega(x_i)\chi_\Omega(y_j)\phi_{\varepsilon_1}(x_i,y_j)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(X,Y,Z)\, dX\, dY\, dZ$$

$$\begin{aligned}
\alpha^1_{rr'}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) &:= \gamma^1_{\text{micro}} \cdot \frac{1}{\mu_{a,b,c}} \int \sum_{i,j,k=1}^{n} \prod_{l=1}^{k-1} \chi_\Omega(z_l)\chi_r(z_k)\chi_V(x_i)\chi_V(y_j) \int_{\Omega_{r'}} \phi_{\varepsilon_1}(z_k,y)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}\left(X_i^+(z_k), Y_j^+(y), Z_k^-\right) dy\, dX\, dY\, dZ \\
&= \gamma^1_{\text{micro}} \cdot \frac{1}{\mu_{a,b,c}} \int \sum_{i,j,k=1}^{n} \prod_{l=1}^{k-1} \chi_\Omega(z_l)\chi_V(z_k)\chi_r(x_i)\chi_{r'}(y_j)\phi_{\varepsilon_1}(x_i,y_j)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(X,Y,Z)\, dX\, dY\, dZ \\
&= \gamma^1_{\text{micro}} \cdot \frac{1}{\mu_{a,b,c}} \int \sum_{i,j=1}^{n} \chi_r(x_i)\chi_{r'}(y_j)\phi_{r'}(y_j)\phi_{\varepsilon_1}(x_i,y_j)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(X,Y,Z)\, dX\, dY\, dZ,
\end{aligned}$$

which implies

$$\sum_{r,r'} \alpha^1_{rr'}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) = \alpha^1(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$$

for all $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c} \in \mathcal{I}$, and

$$\alpha^2(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) := \gamma^2_{\text{micro}} \cdot \frac{1}{\mu_{a,b,c}} \int \sum_{k} \chi_\Omega(z_k)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(X,Y,Z)\, dX\, dY\, dZ$$

$$\begin{aligned}
\alpha^2_{rr'}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) &:= \frac{\gamma^2_{\text{micro}}}{|B_{\varepsilon_2}|^2} \cdot \frac{1}{\mu_{a,b,c}} \int \sum_{i,j,k=1}^{n} \chi_V(z_k) \prod_{l=1}^{i-1} \chi_\Omega(x_l)\chi_r(x_i) \prod_{l=1}^{j-1} \chi_\Omega(y_l)\chi_{r'}(y_j) \\
&\quad \cdot \int_{\Omega_r} \phi_{\varepsilon_2}(x_i,z)\phi_{\varepsilon_2}(y_j,z)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}\left(X_i^-, Y_j^-, Z_k^+(z)\right) dz\, dX\, dY\, dZ \\
&= \frac{\gamma^2_{\text{micro}}}{|B_{\varepsilon_2}|^2} \cdot \frac{1}{\mu_{a,b,c}} \int \sum_{i,j,k=1}^{n} \chi_r(z_k) \prod_{l=1}^{i-1} \chi_\Omega(x_l)\chi_r(x_i) \prod_{l=1}^{j-1} \chi_\Omega(y_l)\chi_{r'}(y_j) \\
&\quad \cdot \phi_{\varepsilon_2}(x_i,z_k)\phi_{\varepsilon_2}(y_j,z_k)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}\left(X_i^-, Y_j^-, Z\right) dX\, dY\, dZ \\
&= \frac{\gamma^2_{\text{micro}}}{|B_{\varepsilon_2}|^2} \cdot \frac{1}{\mu_{a,b,c}} \int \sum_{i,j,k=1}^{n} \chi_r(z_k) \prod_{l=1}^{i-1} \chi_\Omega(x_l)\chi_V(x_i) \prod_{l=1}^{j-1} \chi_\Omega(y_l)\chi_V(y_j) \\
&\quad \cdot \int_{\Omega_r}\int_{\Omega_{r'}} \phi_{\varepsilon_2}(x,z_k)\phi_{\varepsilon_2}(y,z_k)\, dx\, dy \cdot \chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(X,Y,Z)\, dX\, dY\, dZ \\
&= \frac{\gamma^2_{\text{micro}}}{|B_{\varepsilon_2}|^2} \cdot \frac{1}{\mu_{a,b,c}} \int \sum_{k=1}^{n} \chi_r(z_k) \cdot |B^r_{\varepsilon_2}(z_k)| \cdot |B^{r'}_{\varepsilon_2}(z_k)| \cdot \chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}}(X,Y,Z)\, dX\, dY\, dZ,
\end{aligned}$$

where

$$|B^r_{\varepsilon_2}(z)| := \int_{\Omega_r} \phi_{\varepsilon_2}(x,z)\,dx$$

such that $\sum_{r=1}^m |B^r_{\varepsilon_2}(z)| = |B_{\varepsilon_2}|$ for all $z \in \Omega$ and therefore

$$\sum_{r,r'} \alpha^2_{rr'}(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) = \alpha^2(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$$

for all $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c} \in \mathcal{I}$.

For the diffusion part we have

$$\left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, (L_A + L_B + L_C)\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \right\rangle = \left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, L_A\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \right\rangle + \left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, L_B\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \right\rangle + \left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, L_C\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \right\rangle$$

and

$$\left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, L_A\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \right\rangle = \left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, \sum_{i=1}^n \chi_\Omega(x_i) L_{x_i}\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \right\rangle. \tag{A2}$$

In each of these summands, the generator only acts on the coordinates X_i of one individual particle of species A which means that the related scalar product can be nonzero only for those $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ with $\boldsymbol{a} = \tilde{\boldsymbol{a}} + 1_{r'} - 1_r$ for some $r,r' \in \{1,\dots,m\}$, while $\tilde{\boldsymbol{b}} = \boldsymbol{b}$ and $\tilde{\boldsymbol{c}} = \boldsymbol{c}$. (In an equivalent notation, this statement holds for species B and C.)

We assume that all particles of one species follow the same diffusion dynamics (i.e., $L_{x_i} = L_1$ for all i; equivalently for y_i,z_i). Then, for $r,r' = 1,\dots,m$ and $l = 1,2,3$, we define the jump rate from $\Omega_{r'}$ to Ω_r by

$$\lambda^l_{r'r} := \frac{1}{\mu_{r'}} \langle \chi_r, L_l \chi_{r'} \rangle$$

which—by the properties of the propagator L_l—fulfills $\lambda^l_{r'r} \geq 0$ for $r' \neq r$ and $\lambda^l_{rr} = -\sum_{r'\neq r} \lambda^l_{rr'}$ for all r.

Fixing one x_i (for simplicity $i = 1$), we have

$$\begin{aligned}
\left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, \chi_\Omega(x_1)L_{x_1}\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \right\rangle &= \sum_r \left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, \chi_r(x_1)L_{x_1}\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}} \right\rangle \\
&= \sum_r \int \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}(X,Y,Z)\chi_r(x_1)(L_{x_1}\chi_{\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}})(X,Y,Z)\,dX\,dY\,dZ \\
&= \sum_{\substack{r:\tilde{\boldsymbol{a}}_r>0 \\ r':\boldsymbol{a}_{r'}>0}} \int \chi_r(x_1)(L_{x_1}\chi_{r'})(x_1)\chi_{\tilde{\boldsymbol{a}}-1_r}(x_2,\dots,x_n)\chi_{\boldsymbol{a}-1_{r'}}(x_2,\dots,x_n) \\
&\quad \cdot \chi_{\tilde{\boldsymbol{a}}}(Y)\chi_{\boldsymbol{b}}(Y)\chi_{\tilde{\boldsymbol{a}}}(Z)\chi_{\boldsymbol{c}}(Z)\,dX\,dY\,dZ \\
&= \sum_{\substack{r:\tilde{\boldsymbol{a}}_r>0 \\ r':\boldsymbol{a}_{r'}>0}} \mu_{r'}\lambda^1_{r'r} \cdot \delta_{\boldsymbol{a}-1_{r'},\tilde{\boldsymbol{a}}-1_r} \int \chi_{\tilde{\boldsymbol{a}}-1_r}(x_2,\dots,x_n)\,dx_2\,\dots\,dx_n \cdot \delta_{\tilde{\boldsymbol{b}},\boldsymbol{b}}\,\mu_{\tilde{\boldsymbol{b}}} \cdot \delta_{\tilde{\boldsymbol{c}},\boldsymbol{c}}\,\mu_{\tilde{\boldsymbol{c}}} \\
&= \sum_{\substack{r:\tilde{\boldsymbol{a}}_r>0 \\ r':\boldsymbol{a}_{r'}>0}} \delta_{\boldsymbol{a}-1_{r'},\tilde{\boldsymbol{a}}-1_r} \cdot \delta_{\tilde{\boldsymbol{b}},\boldsymbol{b}} \cdot \delta_{\tilde{\boldsymbol{c}},\boldsymbol{c}} \cdot \lambda^1_{r'r}\mu_{\tilde{\boldsymbol{a}}+1_{r'}-1_r,\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}.
\end{aligned}$$

In (A2) we sum up over all $i = 1,\dots,n$, which leads to a multiplication of $\lambda^1_{r'r}$ by $\boldsymbol{a}_{r'} = (\tilde{\boldsymbol{a}} + 1_{r'} - 1_r)_{r'}$ giving

$$\begin{aligned}
\left\langle \chi_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}, L_A\chi_{\tilde{\boldsymbol{a}}+1_{r'}-1_r,\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} \right\rangle &= \sum_{r,r'} (\tilde{\boldsymbol{a}} + 1_{r'} - 1_r)_{r'} \cdot \lambda^1_{r'r} \cdot \mu_{\tilde{\boldsymbol{a}}+1_{r'}-1_r,\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} \\
&= \sum_r \left(\tilde{\boldsymbol{a}}_r \cdot \lambda^1_{rr} \cdot \mu_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} + \sum_{r'\neq r} (\tilde{\boldsymbol{a}}_{r'} + 1) \cdot \lambda^1_{r'r} \cdot \mu_{\tilde{\boldsymbol{a}}+1_{r'}-1_r,\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} \right).
\end{aligned}$$

We finally get

$$\begin{aligned}
\mu_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} \cdot \partial_t p_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}(t) &= \sum_r \sum_{r'\neq r} \left(\lambda^1_{r'r}(\tilde{a}_{r'} + 1)\mu_{\tilde{\boldsymbol{a}}+1_{r'}-1_r,\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} p_{\tilde{\boldsymbol{a}}+1_{r'}-1_r,\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} - \lambda^1_{rr'}\tilde{a}_r\mu_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} p_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} \right) \\
&\quad + \sum_r \sum_{r'\neq r} \left(\lambda^2_{r'r}(\tilde{b}_{r'} + 1)\mu_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}}+1_{r'}-1_r,\tilde{\boldsymbol{c}}} p_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}}+1_{r'}-1_r,\tilde{\boldsymbol{c}}} - \lambda^2_{rr'}\tilde{b}_r\mu_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} p_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} \right) \\
&\quad + \sum_r \sum_{r'\neq r} \left(\lambda^3_{r'r}(\tilde{c}_{r'} + 1)\mu_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}+1_{r'}-1_r} p_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}+1_{r'}-1_r} - \lambda^3_{rr'}\tilde{c}_r\mu_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} p_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} \right)
\end{aligned}$$

$$
\begin{aligned}
&- \alpha^1(\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}, \tilde{\boldsymbol{c}}) \cdot \mu_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} \cdot p_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}(t) + \sum_{r,r':\tilde{c}_r>0} \alpha^1_{rr'}(\tilde{\boldsymbol{a}} + 1_r, \tilde{\boldsymbol{b}} + 1_{r'}, \tilde{\boldsymbol{c}} - 1_r) \\
&\cdot \mu_{\tilde{\boldsymbol{a}}+1_r,\tilde{\boldsymbol{b}}+1_{r'},\tilde{\boldsymbol{c}}-1_r} \cdot p_{\tilde{\boldsymbol{a}}+1_r,\tilde{\boldsymbol{b}}+1_{r'},\tilde{\boldsymbol{c}}-1_r}(t) - \alpha^2(\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}, \tilde{\boldsymbol{c}}) \cdot \mu_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} \cdot p_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}(t) \\
&+ \sum_{r,r':\tilde{a}_r>0,b_{r'}>0} \alpha^2_{rr'}(\tilde{\boldsymbol{a}} - 1_r, \tilde{\boldsymbol{b}} - 1_{r'}, \tilde{\boldsymbol{c}} + 1_r) \cdot \mu_{\tilde{\boldsymbol{a}}-1_r,\tilde{\boldsymbol{b}}-1_{r'},\tilde{\boldsymbol{c}}+1_r} \cdot p_{\tilde{\boldsymbol{a}}-1_r,\tilde{\boldsymbol{b}}-1_{r'},\tilde{\boldsymbol{c}}+1_r}(t).
\end{aligned}
$$

Setting $\alpha^1_{rr'} = 0$, $\alpha^2_{rr'} = 0$ for $r \neq r'$, this is the ST-CME given in (1), where $\boldsymbol{n}$ is given by $(\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}, \tilde{\boldsymbol{c}})$ with $n^1_r = \tilde{a}_r$, $n^2_r = \tilde{b}_r$, $n^3_r = \tilde{c}_r$, and

$$P(\boldsymbol{n}, t) = \mu_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}} \cdot p_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\tilde{\boldsymbol{c}}}(t)$$

for $\boldsymbol{n} = (\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}}, \tilde{\boldsymbol{c}})$.

[1] 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.

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

[3] G. R. Bowman, X. Huang, and V. S. Pande, "Using generalized ensemble simulations and Markov state models to identify conformational states," Methods **49**(2), 197–201 (2009).

[4] G. R. Bowman, V. S. Pande, and F. Noé, "An introduction to Markov state models and their application to long timescale molecular simulation," *Advances in Experimental Medicine and Biology* (Springer, 2014).

[5] J. D. Chodera, K. A. Dill, N. Singhal, V. S. Pande, W. C. Swope, and J. W. Pitera, "Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics," J. Chem. Phys. **126**, 155101 (2007).

[6] J. D. Chodera, P. J. Elms, W. C. Swope, J.-H. Prinz, S. Marqusee, C. Bustamante, F. Noé, and V. S. Pande, "A robust approach to estimating rates from time-correlation functions," e-print arXiv:1108.2304 (2011).

[7] M. Doi, "Stochastic theory of diffusion-controlled reaction," J. Phys. A: Math. Gen. **9**(9), 1479 (1976).

[8] S. Engblom, L. Ferm, A. Hellander, and P. Lötstedt, "Simulation of stochastic reaction-diffusion processes on unstructured meshes," SIAM J. Sci. Comput. **31**(3), 1774–1797 (2009).

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

[10] R. Erban, J. Chapman, and P. Maini, "A practical guide to stochastic simulations of reaction-diffusion processes," preprint arXiv:0704.1908 (2007).

[11] C. W. Gardiner, K. J. McNeil, D. F. Walls, and I. S. Matheson, "Correlations in stochastic theories of chemical reactions," J. Stat. Phys. **14**(4), 307–331 (1976).

[12] 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).

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

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

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

[16] S. Hellander, A. Hellander, and L. Petzold, "Reaction-diffusion master equation in the microscopic limit," Phys. Rev. E **85**(4), 042901 (2012).

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

[18] 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).

[19] S. A. Isaacson, "A convergent reaction-diffusion master equation," J. Chem. Phys. **139**(5), 054101 (2013).

[20] 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).

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

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

[23] J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Fischbach, M. Held, C. Schütte, J. D. Chodera, and F. Noé, "Markov models of molecular kinetics: Generation and validation," J. Chem. Phys. **134**(17), 174105 (2011).

[24] M. Sarich, F. Noé, and C. Schütte, "On the approximation quality of Markov state models," Multiscale Model. Simul. **8**(4), 1154–1177 (2010).

[25] M. Sarich, R. Banisch, C. Hartmann, and C. Schütte, "Markov state models for rare events in molecular dynamics," Entropy **16**(1), 258 (2013).

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

[27] J. Schöneberg, A. Ullrich, and F. Noé, "Simulation tools for particle-based reaction-diffusion dynamics in continuous space," BMC Biophys. **7**(1), 1 (2014).

[28] C. Schütte and M. Sarich, *Metastability and Markov State Models in Molecular Dynamics: Modeling, Analysis, Algorithmic Approaches*, Courant Lecture Notes Vol. 24 (American Mathematical Society, 2013).

[29] C. Schütte, A. Fischer, W. Huisinga, and P. Deuflhard, "A direct approach to conformational dynamics based on hybrid Monte Carlo," J. Comput. Phys. **151**, 146–168 (1999).

[30] C. Schütte, F. Noé, J. Lu, M. Sarich, and E. Vanden-Eijnden, "Markov state models based on milestoning," J. Chem. Phys. **134**(20), 204105 (2011).

[31] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 4th ed. (Elsevier, Amsterdam, 2006).

[32] 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).

[33] M. Weber, "Meshless methods in conformation dynamics," Ph.D. thesis, Freie University of Berlin, 2006.