Journal of Mathematical Physics | ARTICLE | scitation.org/journal/jmp

Chemical diffusion master equation: Formulations of reaction–diffusion processes on the molecular level

Cite as: J. Math. Phys. 64, 013304 (2023); doi: 10.1063/5.0129620
Submitted: 6 October 2022 • Accepted: 2 January 2023 •
Published Online: 20 January 2023

View Online Export Citation CrossMark

Mauricio J. del Razo,[1,2,a)] Stefanie Winkelmann,[3,a)] Rupert Klein,[1] and Felix Höfling[1,3]

AFFILIATIONS

[1] Department of Mathematics and Computer Science, Freie Universität Berlin, Berlin, Germany
[2] Dutch Institute for Emergent Phenomena, 1090GL Amsterdam, The Netherlands
[3] Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany

Note: Paper published as part of the Special Topic on Cascades of Scales: Applications and Mathematical Methodologies.
[a)] **Authors to whom correspondence should be addressed:** m.delrazo@fu-berlin.de and winkelmann@zib.de

ABSTRACT

The chemical diffusion master equation (CDME) describes the probabilistic dynamics of reaction–diffusion systems at the molecular level [del Razo *et al.*, Lett. Math. Phys. **112**, 49 (2022)]; it can be considered as the master equation for reaction–diffusion processes. The CDME consists of an infinite ordered family of Fokker–Planck equations, where each level of the ordered family corresponds to a certain number of particles and each particle represents a molecule. The equations at each level describe the spatial diffusion of the corresponding set of particles, and they are coupled to each other via reaction operators—linear operators representing chemical reactions. These operators change the number of particles in the system and, thus, transport probability between different levels in the family. In this work, we present three approaches to formulate the CDME and show the relations between them. We further deduce the non-trivial combinatorial factors contained in the reaction operators, and we elucidate the relation to the original formulation of the CDME, which is based on creation and annihilation operators acting on many-particle probability density functions. Finally, we discuss applications to multiscale simulations of biochemical systems among other future prospects.

I. INTRODUCTION

It is a well-established paradigm to consider biochemical dynamics as an interplay between the spatial transport (diffusion) of molecules and their chemical kinetics (reaction), both of which are inherently stochastic. There exist different approaches for modeling and mathematically formalizing such reaction–diffusion processes, ranging from reaction–diffusion master equations,[1–5] where spatial transport is modeled by diffusive jumps between local compartments, to concentration-based approaches, such as deterministic[6–9] or stochastic partial differential equations[10] and partial integro-differential equations.[11] The preceding modeling approaches may be regarded as approximations or limiting cases of particle-based reaction–diffusion (PBRD) models, which explicitly resolve the diffusive trajectories of individual particles in space and time, as well as reactions between them. In the standard PBRD models, particles move freely in space following Brownian motion, or any other form of diffusion process,[12,13] and can undergo chemical reactions, which involve one, two, or more reactants in such a way that the reaction rate can depend on the positions or relative positions between the reactants.[14,15] Because of their high complexity, PBRD systems are mostly studied numerically by means of Monte Carlo simulations of the underlying stochastic reaction–diffusion process.

The mathematical formalization and analysis of PBRD models, however, are difficult because reactions constantly change the number of particles of each species, changing the dimension and composition of the system. Recent work presents a probabilistic framework and the characteristic evolution equation for PBRD termed chemical diffusion master equation (CDME).[16] The CDME consists of an infinite ordered family of Fokker–Planck equations (i.e., an enumerated collection), where each equation corresponds to a certain number of particles

$n = 0, 1, 2, \ldots$. The equations, for each fixed n, describe the spatial diffusion for the corresponding n-particle probability distribution, and they are coupled via reaction operators that express the changes in the system's state due to chemical reactions. These operators change the number of particles in the system, and thus, they can be conveniently expressed in terms of creation and annihilation operators,[16] following a classical analog of the quantum mechanical Fock space concept.[17,18] First steps toward solving the CDME analytically by means of the Malliavin calculus were taken recently.[19] A more comprehensive introduction on the topic can be found in Ref. 16.

In this work, we explore the CDME from several perspectives and present three approaches to motivate and formulate it. This work not only improves our understanding of how to formulate the CDME but also provides a more illustrative and accessible approach to practitioners than the original work.[16] In general, the CDME is composed of a diffusion operator and several reaction operators (one for each included reaction), all of them acting on a symmetric many-particle distribution function. In analogy to the well-known chemical master equation,[20–23] which characterizes spatially well-mixed stochastic reaction kinetics, each reaction operator consists of a *loss* term describing the probabilistic outflow from a given configuration state by the reaction and a *gain* term that captures the probabilistic inflow from other configuration states due to the reaction. The crucial part is to determine these loss and gain operators for different types of reactions in the absence of a spatially well-mixed setting; examples are binding and unbinding, creation and degradation, and mutual annihilation. Here, non-trivial combinatorial factors enter for preserving symmetry and normalization of the many-particle distribution functions under time evolution. The local rate function, which defines the probability per unit of time for a reaction to take place depending on the spatial positions of its reactants and products, has to be transformed into an expression that takes the whole system state into account. This issue is addressed via the following three approaches:

1. We use the local rate functions to specify also the loss and gain operators on a local scale (acting on subsets of reactants and products) and then combine them into *global* operators taking all combinations of reacting subgroups into account. The combinatorial factors included in the operators are motivated by an inductive argument. The CDME may then directly be expressed in terms of these *global* loss and gain operators (Sec. II).
2. The *global* loss and gain operators are expressed in terms of *many-particle propensity functions*, which define the probability per unit of time for a reaction to occur as a function of the whole system state. We explicitly derive these many-particle propensity functions from the given local rate functions using permutations and Dirac δ-distributions. For the exemplary settings of decay and binding, it will be shown that the resulting CDME agrees with the one of the first approach (Sec. III).
3. The operators in the CDME are expressed as expansions in terms of creation and annihilation operators as in Ref. 16. These expansions can be condensed in a compact notation that allows us to write the CDME, for a given system of reactions, in a simple, fast, and straightforward manner. The combinatorial factors do not appear explicitly; instead, they are naturally encoded in the creation and annihilation operators (Sec. IV). A dictionary specifying the relation between the compact notation for the expansions and the concrete algebraic expressions in the classical representation is provided in the Appendix.

In all three approaches, we start with a simplified setting containing only one molecular species, which drastically simplifies the notation, and then generalize to reactions involving several species, such as complex formation and general association reactions.

II. THE CHEMICAL DIFFUSION MASTER EQUATION: AN INTUITIVE FORMULATION

We consider an open system of a varying number of diffusing particles of the same chemical species in a finite space domain $\mathbb{X} \subset \mathbb{R}^d$. The diffusion process changes the spatial configuration of the particles, while the reaction process can change the number of particles in the system. The configuration of the system is thus given by the numbers of particles and their positions. The probability distribution of such a system is given as an ordered family of probability density functions,

$$\rho = (\rho_0, \rho_1, \rho_2, \ldots, \rho_n, \ldots), \tag{1}$$

where $\rho_n(x^{(n)})$ is the probability density of finding n particles at the positions $x^{(n)} = (x_1^{(n)}, \ldots, x_n^{(n)})$ for $n \geqslant 1$, while ρ_0 is the probability for no particles being present. As the particles are statistically indistinguishable from each other, the densities must be symmetric with respect to permutations of particle labels, e.g., $\rho_2(y, z) = \rho_2(z, y)$ for all $y, z \in \mathbb{X}$, and more generally,

$$\rho_n(x^{(n)}) = \rho_n(Px^{(n)}) \qquad \text{for all } P \in \mathcal{P}_n, \tag{2}$$

where $\mathcal{P}_n$ is the set of all permutations of an n-tuple. The normalization condition is

$$\rho_0 + \sum_{n=1}^{\infty} \int_{\mathbb{X}^n} \rho_n(x^{(n)}) dx^{(n)} = 1. \tag{3}$$

In general, ρ will also depend on time, $\rho_n = \rho_n(t, x^{(n)})$, but we will omit t for simplicity. As a remark, the distribution ρ is an element of a linear function space similar to the Fock space of quantum mechanics; see Refs. 16–18 and Sec. IV.

Given that there are $M \in \mathbb{N}$ reactions, the CDME has the general form

$$\frac{\partial \rho}{\partial t} = \left(\mathcal{D} + \sum_{r=1}^{M} \mathcal{R}^{(r)}\right)\rho \tag{4}$$

for a diffusion operator $\mathcal{D}$ and reaction operators $\mathcal{R}^{(r)}$. Each of the reaction operators $\mathcal{R}^{(r)}$ corresponds to one possible reaction, and it is conveniently split into loss and gain operators,

$$\mathcal{R}^{(r)} = \mathcal{G}^{(r)} - \mathcal{L}^{(r)}, \tag{5}$$

similarly; the reaction operator in Ref. 16 was split into a particle conserving part (the loss operator) and a non-conserving part (the gain operator), In the following, we will construct these loss and gain operators at first for reactions of a single species and then for a multi-species scenario. In each case, we consider a system with only one reaction such that the index r can be skipped. For systems with several reactions, the results may simply be combined by summing up these operators as in Eq. (4).

A. One species

To start with, we assume that there is only one chemical species A. The most general reaction in this case is of the form

$$kA \to lA \tag{6}$$

for $k, l \in \mathbb{N}_0$. The rate at which a reaction event occurs is given by $\lambda(y^{(l)};\ x^{(k)}) > 0$, and it depends on the positions $x^{(k)} \in \mathbb{X}^k$ of the reactants and the positions $y^{(l)} \in \mathbb{X}^l$ of the products. Note that the rate function λ should be symmetric with respect to pair exchanges in both of its arguments.

We can now write the nth component of the CDME as

$$\frac{\partial \rho_n}{\partial t} = \mathcal{D}_n \rho_n + \mathcal{G}_n \rho_{n+k-l} - \mathcal{L}_n \rho_n \tag{7}$$

for appropriate operators $\mathcal{D}_n$, $\mathcal{G}_n$, and $\mathcal{L}_n$ referring to diffusion, gain, and loss, respectively. Note that in Ref. 16, the loss operator is denoted by $\mathcal{R}^{(k)}$ and the gain operator by $\mathcal{R}^{(k,l)}$; however, we find the new notation less cumbersome. Reactions at the n-particle state produce a transition to the $(n-k+l)$-particle state. Thus, the loss of probability for the n-particle state ρ_n depends only on itself. Similarly, reactions at the $(n+k-l)$-particle state produce a transition to the ρ_n state. Thus, the gain of probability for the n-particle state depends on ρ_{n+k-l}.

For physically non-interacting particles, the diffusion operator $\mathcal{D}_n$ can be expressed in terms of the one-particle diffusion D_ν applied to the νth particle,

$$\mathcal{D}_n = \sum_{\nu=1}^{n} D_\nu, \tag{8}$$

where D_ν is the infinitesimal generator of the one-particle Fokker–Planck equation. For example, one may think of D_ν as something as simple as the d-dimensional Laplacian, $D_\nu = \nabla^2_{x_\nu}$. Ignoring the reaction operators and assuming that there is no exchange of particles with a reservoir outside of $\mathbb{X}$,[24] all the resulting equations are uncoupled and one obtains a family of uncoupled Fokker–Planck equations unless there is an exchange of particles with the world outside of $\mathbb{X}$, in which case one ends up again with a similar family of many-particle densities, albeit with a different structure of the coupling between its levels.[25] For simplicity of the exposition, we assume reflecting boundaries for $\mathbb{X}$ from here on, i.e., a confinement by rigid walls.

The loss operator acting on the n-particle density will output the total rate of probability loss of ρ_n due to all possible combinations of reactants. It is given in terms of the loss per reaction $L_{\nu_1,\ldots,\nu_k}$ (*local* loss), which acts on k particles at a time, with $(\nu_1,\ldots,\nu_k)$ denoting the indices of the particles that it acts on. The loss per reaction quantifies how much probability is lost to the current state due to one reaction; it is thus the integral over the density and the rate function λ over all the possible positions of the products,

$$(L_{\nu_1,\ldots,\nu_k}\rho_n)(x^{(n)}) = \rho_n(x^{(n)}) \int_{\mathbb{X}^l} \lambda(y^{(l)}; x^{(n)}_{\nu_1,\ldots,\nu_k})\, dy^{(l)}, \tag{9}$$

where $x^{(n)}_{\nu_1,\ldots,\nu_k} := (x^{(n)}_{\nu_1},\ldots,x^{(n)}_{\nu_k})$. The total loss is then the sum of the loss per reaction over all possible reactions,

$$\mathcal{L}_n = \sum_{1\leqslant \nu_1 < \cdots < \nu_k \leqslant n} L_{\nu_1,\ldots,\nu_k}. \tag{10}$$

The form of the ordered sum guarantees that we count all the possible ways of picking up k particles without double counting; see Fig. 1 for a diagram of the calculation. For the special case of $k = 0$, we have

$$(\mathcal{L}_n \rho_n)(x^{(n)}) = \rho_n(x^{(n)}) \int_{\mathbb{X}^l} \lambda(y^{(l)})\, dy^{(l)}. \tag{11}$$

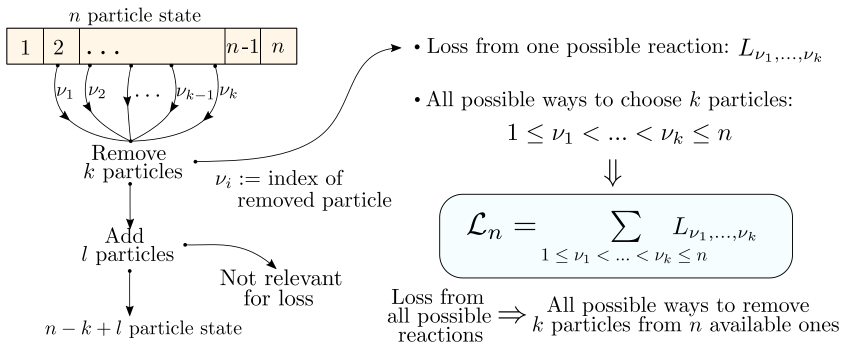

FIG. 1. Diagram representing the loss of probability from the n-particle state due to the reaction $kA \to lA$ [Eq. (10)]. The particle states are represented by a set of boxes, where each box corresponds to the index of a particle.

Similarly, the gain operator acting on the n-particle density will output the total rate of probability gain of ρ_n. It can be expressed in terms of the gain per reaction (*local* gain) resulting from k reacting particles with indices $(\nu_1,\ldots,\nu_k)$ producing l products with indices $(\mu_1,\ldots,\mu_l)$, termed $G_{\mu_1,\ldots,\mu_l}$. The gain per reaction quantifies how much probability is gained by the current state due to one reaction; it is thus the integral over the density and the rate function λ over all the possible positions of the reactants,

$$(G_{\mu_1,\ldots,\mu_l}\rho_{n+k-l})(x^{(n)}) = \int_{\mathbb{X}^k} \lambda(x^{(n)}_{\mu_1,\ldots,\mu_l}; z^{(k)})\rho_{n+k-l}(x^{(n)}_{\setminus\{\mu_1,\ldots,\mu_l\}}, z^{(k)})dz^{(k)}, \tag{12}$$

where the subscript $\setminus\{\mu_1,\ldots,\mu_l\}$ means that the entries with indices $\mu_1,\ldots,\mu_l$ are excluded from the tuple $x^{(n)}$ of particle positions. Note that the indices of the reacting particles $\nu_1,\ldots,\nu_k$ are not relevant for the gain since the reactants' positions are integrated out (and both the density and the rate function are symmetric). The total gain is then the sum of the gain per reaction over all possible reactions,

$$\mathcal{G}_n = \frac{(n-l)!}{n!}\binom{n+k-l}{k} \sum_{\substack{\mu_1\ldots\mu_l=1 \\ \mu_i\neq\mu_j\ \forall i,j}}^{n} G_{\mu_1,\ldots,\mu_l} \tag{13a}$$

$$= \binom{n}{l}^{-1}\binom{n+k-l}{k} \sum_{1\leqslant\mu_1<\cdots<\mu_l\leqslant n} G_{\mu_1,\ldots,\mu_l}, \tag{13b}$$

where we used the symmetry of $G_{\mu_1,\ldots,\mu_l}$ with respect to the indices. The complicated form of the gain operator is due to the fact that it needs to consider all the possible ways to pick up k particles from the $(n+k-l)$-particle state, just as the loss operator, but, in addition, it also needs to consider all the possible ways of incorporating l particles into the current state in a symmetry-preserving manner; see Fig. 2 for a diagram illustrating the calculation. Note that the output of the loss and gain operators is also symmetric.

Let us use the preceding formulas for general reactions involving one species to derive the CDME for some common reactions [for simplicity, we write $\rho_n(x^n,t)$ as $\rho_n(x^n)$]:

- Degradation $A \to \emptyset$: This case is recovered with $k=1, l=0$ using the rate function $\lambda_d(x) = \lambda_d(;\ x)$. The CDME reads

$$\frac{\partial \rho_n}{\partial t}(x^n) = \sum_{\nu=1}^{n} D_\nu \rho_n(x^n) + (n+1)\int_{\mathbb{X}} \lambda_d(z)\rho_{n+1}(x^{(n)}, z)dz - \rho_n(x^{(n)})\sum_{\nu=1}^{n}\lambda_d(x_\nu^{(n)}). \tag{14}$$

- Creation $\emptyset \to A$: Here, we set $k=0, l=1$ using the rate function $\lambda_c(y) = \lambda_c(y;)$; then, the CDME is

$$\frac{\partial \rho_n}{\partial t}(x^n) = \sum_{\nu=1}^{n} D_\nu \rho_n(x^n) + \frac{1}{n}\sum_{\mu=1}^{n}\rho_{n-1}\left(x^{(n)}_{\setminus\{\mu\}}\right)\lambda_c(x^{(n)}_\mu) - \rho_n(x^{(n)})\int_{\mathbb{X}}\lambda_c(y)dy. \tag{15}$$

- Mutual annihilation $A + A \to \emptyset$: In this case, we have $k=2, l=0$ with the rate function $\lambda_a(x_1,x_2) = \lambda_a(;\ x_1,x_2)$. Then,

$$\frac{\partial \rho_n}{\partial t}(x^n) = \sum_{\nu=1}^{n} D_\nu \rho_n(x^n) + \frac{(n+2)(n+1)}{2}\int_{\mathbb{X}^2}\lambda_a(z_1,z_2)\rho_{n+2}(x^{(n)}, z_1, z_2)dz_1dz_2$$
$$-\rho_n(x^{(n)})\sum_{1\leqslant\nu_1<\nu_2\leqslant n}\lambda_a(x^{(n)}_{\nu_1}, x^{(n)}_{\nu_2}). \tag{16}$$

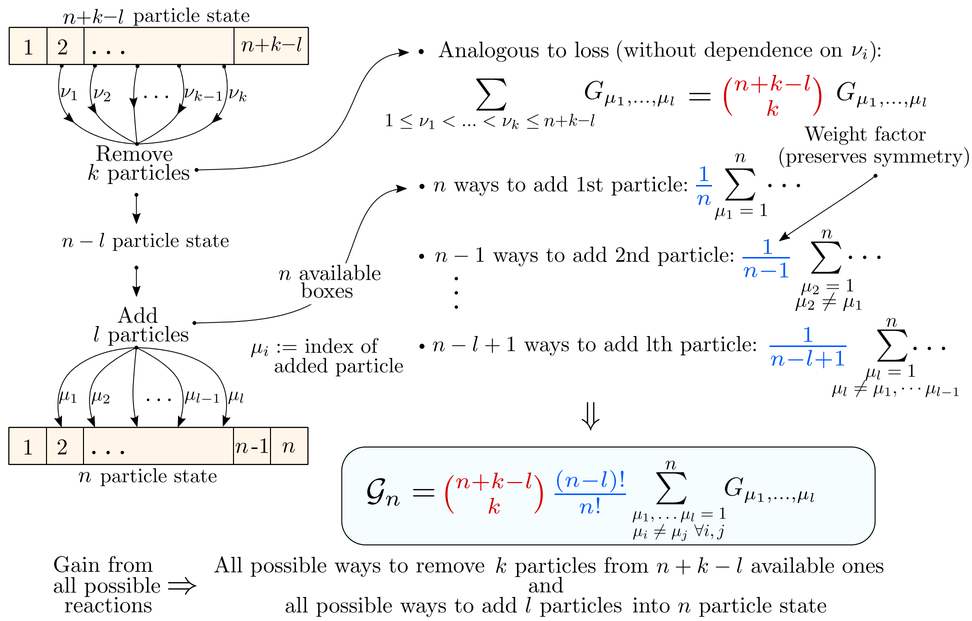

FIG. 2. Diagram representing the gain of probability for the n-particle state for the reaction $kA \to lA$ [Eq. (10)]. The particle states are represented by a set of boxes, where each box corresponds to the index of a particle. The final expression can be further simplified; see Eq. (13b).

- Trimolecular reaction: $3A \to 2A$: Here, $k = 3$, $l = 2$, and the rate function is $\lambda(y^{(2)};\ x^{(3)})$; then,

$$\frac{\partial\rho_n}{\partial t}(x^n) = \sum_{v=1}^{n} D_v\rho_n(x^n) + \frac{n+1}{3}\sum_{1\leqslant\mu_1<\mu_2\leqslant n}\int_{\mathbb{X}^3}\lambda(y^{(2)};z^{(3)})\rho_{n+1}(x^{(n)}_{\setminus\{\mu_1,\mu_2\}},z^{(3)})dz^{(3)}$$
$$-\rho_n(x^{(n)})\sum_{1\leqslant\nu_1<\nu_2<\nu_3\leqslant n}\int_{\mathbb{X}^2}\lambda(y^{(2)};x^{(n)}_{\nu_1},x^{(n)}_{\nu_2},x^{(n)}_{\nu_3})dy^{(2)}. \tag{17}$$

Several reactions: Given a system with several reactions of the form $k_rA \to l_rA$ for different $k_r, l_r \in \mathbb{N}_0$, $r = 1, \ldots, M$, the nth component of the CDME is given by a sum of the form

$$\frac{\partial\rho_n}{\partial t} = \mathcal{D}_n\rho_n + \sum_{r=1}^{M}\left(\mathcal{G}_n^{(r)}\rho_{n+k_r-l_r} - \mathcal{L}_n^{(r)}\rho_n\right) \tag{18}$$

with accordingly defined operators $\mathcal{G}_n^{(r)}$ and $\mathcal{L}_n^{(r)}$. Special cases of this type of equation were proposed earlier in the context of pattern formation[26,27] as models for adsorption/desorption processes and for the dynamics of agents with internal states; these dynamics can be interpreted as reactions of orders 0 and 1 (creation, degradation, and unimolecular reactions, respectively).

As one can see from the expressions above, the explicit formulation of the loss and gain operators can become quite complex due to the combinatorics. This issue worsens when several species are involved. Thus, it appears convenient to have a formalism where the combinatorial factors are intrinsically built-in,[16] and we will present such an approach in Sec. IV. Beforehand, we will explore one example with multiple species and an alternative explicit representation of the CDME.

B. Multiple species

Consider the reaction

$$A + B \to C \tag{19}$$

with rate function $\lambda(y;\ x_A, x_B)$, where x_A and x_B are the locations of one pair of reactants and y is the location of the product. The stochastic dynamics of the system is described in terms of the distributions $\rho_{a,b,c}\left(x^{(a)}, x^{(b)}, x^{(c)}\right)$, where a, b, c indicate the numbers of A, B, and C

particles, respectively, and $x^{(a)}$ indicates the positions of the A particles, $x^{(b)}$ indicates the positions of the B particles, and $x^{(c)}$ indicates the positions of the C particles. The normalization condition [Eq. (3)] generalizes to

$$\sum_{a,b,c=0}^{\infty} \int_{\mathbb{X}^a\times\mathbb{X}^b\times\mathbb{X}^c} \rho_{a,b,c}\Big(x^{(a)},x^{(b)},x^{(c)}\Big)dx^{(a)}\ dx^{(b)}\ dx^{(c)} = 1. \tag{20}$$

The CDME for this reaction has the same structure as before, namely, $\partial\rho/\partial t = \mathcal{D}\rho + \mathcal{R}\rho$ for a diffusion operator $\mathcal{D}$ and a reaction operator $\mathcal{R}$. Writing the equation component-wise and separating the reaction operator into its total loss and gain operators, we obtain

$$\frac{\partial\rho_{a,b,c}}{\partial t} = \mathcal{D}\rho_{a,b,c} + \mathcal{G}_{a,b,c}\rho_{a+1,b+1,c-1} - \mathcal{L}_{a,b}\rho_{a,b,c}. \tag{21}$$

The total loss and gain operators can be written explicitly by defining them per reaction (*locally*) and applying them to all possible combinations of reactors and products in the corresponding state (*globally*). Following the same logic as in Fig. 1, the loss operator is given by

$$\mathcal{L}_{a,b} = \sum_{\mu=1}^{a}\sum_{\nu=1}^{b} L_{\mu,\nu} \tag{22}$$

with

$$(L_{\mu,\nu}\ \rho_{a,b,c})(x^{(a)},x^{(b)},x^{(c)}) = \rho_{a,b,c}\Big(x^{(a)},x^{(b)},x^{(c)}\Big)\int_{\mathbb{X}}\lambda\Big(y;x_\mu^{(a)},x_\nu^{(b)}\Big)dy. \tag{23}$$

Note that for the loss, the positions of the products are not relevant, so $L_{\mu,\nu}$ just depends on the indices μ,ν of the reactants. Moreover, in contrast to Eq. (10), the sum is not ordered since the reaction involves different species. Analogously, we can write the gain, but it is usually more complex since now the location of the products does matter. In analogy to Fig. 2, the gain operator is given by

$$\mathcal{G}_{a,b,c} = \frac{1}{c}\sum_{\xi=1}^{c}\sum_{\mu=1}^{a+1}\sum_{\nu=1}^{b+1} G_\xi = (a+1)(b+1)\frac{1}{c}\sum_{\xi=1}^{c} G_\xi \tag{24}$$

with

$$(G_\xi\rho_{a+1,b+1,c-1})(x^{(a)},x^{(b)},x^{(c)}) = \int_{\mathbb{X}^2}\lambda\Big(x_\xi^{(c)};z,z'\Big)\rho_{a+1,b+1,c-1}\Big((x^{(a)},z),(x^{(b)},z'),x^{(c)}_{\setminus\{\xi\}}\Big)dz\ dz'. \tag{25}$$

Gathering the terms and incorporating the diffusion term in the same way as before for each species, we obtain the CDME,

$$\begin{aligned}\frac{\partial\rho_{a,b,c}}{\partial t} &= \sum_{\mu=1}^{a} D_\mu^A\rho_{a,b,c} + \sum_{\nu=1}^{b} D_\nu^B\rho_{a,b,c} + \sum_{\xi=1}^{c} D_\xi^C\rho_{a,b,c}\\ &\quad + \frac{(a+1)(b+1)}{c}\sum_{\xi=1}^{c}\int_{\mathbb{X}^2}\lambda\Big(x_\xi^{(c)};z,z'\Big)\rho_{a+1,b+1,c-1}\Big((x^{(a)},z),(x^{(b)},z'),x^{(c)}_{\setminus\{\xi\}}\Big)dz\ dz'\\ &\quad - \rho_{a,b,c}\Big(x^{(a)},x^{(b)},x^{(c)}\Big)\sum_{\mu=1}^{a}\sum_{\nu=1}^{b}\int_{\mathbb{X}}\lambda\Big(y;x_\mu^{(a)},x_\nu^{(b)}\Big)dy,\end{aligned} \tag{26}$$

where the dependence of $\rho_{a,b,c}$ on the positions $\Big(x^{(a)},x^{(b)},x^{(c)}\Big)$ and time t has been skipped in the first line to simplify notation.

We see again that the main difficulty in writing down the CDME correctly is to come up with expressions that relate the loss and gain operators acting on a subset of particles to the loss and gain operators acting on the whole system. This is expected as the operators need to account for all possible combinations of particles that can undergo a certain reaction.

III. CDME FORMULATION USING MANY-PARTICLE PROPENSITIES

In this section, we provide another justification of the form of the gain and loss operators (especially of the combinatorial factors) by utilizing permutations and Dirac δ-distributions to mathematically describe the particle selection process and by transforming the local rate function into many-particle propensity functions.

For the simplicity of the notation, we again restrict to the case of only one chemical species as in Sec. II A; a case with multiple species will be discussed in Sec. III B. Given the component-wise formulation [Eq. (18)] of the CDME, we would like to express the gain and loss operators by means of global *many-particle propensities*, which express the likeliness for a reaction to take place depending on the whole system state. More concretely, given a single reaction of the form $kA \to lA$, we consider for each n the propensity functions $\Lambda_n : \mathbb{X}^{n-k+l}\times\mathbb{X}^n \to [0,\infty)$, where $\Lambda_n(y^{(n-k+l)};\, x^{(n)})$ refers to the probability per unit of time that a system with n particles in the ordered positions $x_1^{(n)},\dots,x_n^{(n)}$ gets to be transformed into a system with $n-k+l$ particles in the ordered positions $y_1^{(n-k+l)},\dots,y_{n-k+l}^{(n-k+l)}$. Reflecting the assumption that particles of a single species are modeled as indistinguishable, the many-particle propensities are required to be symmetric with respect to pair exchanges in both of their arguments.

In terms of the many-particle propensities Λ_n, the loss and gain operators are given by

$$(\mathcal{L}_n\rho_n)(x^{(n)}) = \rho_n(x^{(n)})\int_{\mathbb{X}^{n-k+l}}\Lambda_n(y^{(n-k+l)};x^{(n)})dy^{(n-k+l)}, \tag{27a}$$

$$(\mathcal{G}_n\rho_{n+k-l})(y^{(n)}) = \int_{\mathbb{X}^{n+k-l}}\Lambda_{n+k-l}(y^{(n)};x^{(n+k-l)})\rho_{n+k-l}(x^{(n+k-l)})dx^{(n+k-l)}, \tag{27b}$$

in analogy to the operators given in Sec. II A. These expressions are symmetry preserving owing to the symmetry properties of the densities and of the propensities. They are probability preserving, too, because taking into account that $\mathcal{L}_n\rho_n$ is a loss for ρ_n while $\mathcal{G}_{n-k+l}\rho_n$ is a gain for ρ_{n-k+l}, the sum of the changes of the total probability in the n- and $n-k+l$-particle spaces due to the considered reaction is

$$\int_{\mathbb{X}^n}(\mathcal{L}_n\rho_n)(x^{(n)})dx^{(n)} - \int_{\mathbb{X}^{n-k+l}}(\mathcal{G}_{n-k+l}\rho_n)(y^{(n-k+l)})dy^{(n-k+l)} = 0. \tag{28}$$

If the densities ρ_n are symmetric with respect to arbitrary particle permutations initially, the loss and gain operations from Eqs. (27a) and (27b) will preserve this property. Moreover, owing to the way the densities are normalized in Eq. (3), no normalizing combinatorial factors arise in Eqs. (27) and (28); rather, the combinatorics is hidden in the definition of Λ_n. Thus, we conclude that conservation of symmetry and probability is straightforwardly ensured when working with the many-particle propensities Λ_n.

Given a finite number M of reactions of the form $k_rA \to l_rA$, we denote the propensity functions of the rth reaction by $\Lambda_n^{(r)}$ and the corresponding loss and gain operators by $\mathcal{L}_n^{(r)}$ and $\mathcal{G}_n^{(r)}$. Inserting into Eq. (18), we obtain the nth component of the CDME in terms of the many-particle propensities $\Lambda_n^{(r)}$.

The next step is to derive the concrete form of the many-particle propensity Λ_n for specific reactions and relate them to the local rate functions λ. Remember that, in contrast to the propensities Λ_n, the rate functions λ define the rate for a reaction taking place solely depending on the positions of reactants and products. More concretely, $\lambda(y^{(l)},x^{(k)})$ defines the probability per unit of time for k particles located at $x_1^{(k)},\dots,x_k^{(k)}$ to be fully replaced due to the reaction $kA\to lA$ by l particles located at $y_1^{(l)},\dots,y_l^{(l)}$. In contrast, the global many-particle propensities Λ_n depend on the complete system state before and after the reactions and already contain combinatorial factors and symmetrization. As a first scenario, we consider the example of simple decay.

A. Many-particle propensity for simple decay

Here, we develop an explicit formula that relates the reaction rate $\lambda_d(x)$ of the decay process, see Eq. (14), to the associated many-particle propensity $\Lambda_{n+1}:\mathbb{X}^n\times\mathbb{X}^{n+1}\to[0,\infty)$. The following formula captures the essence of the remaining many-particle propensities but does not yet respect the required symmetries and the normalization:

$$\Lambda_{n+1}^{\mathrm{bs}}\Big(y^{(n)},x^{(n+1)}\Big) = \sum_{\nu=1}^{n+1}\lambda_d\Big(x_\nu^{(n+1)}\Big)\delta^n\Big(x_{\setminus\{\nu\}}^{(n+1)} - y^{(n)}\Big), \tag{29}$$

where the superscript "bs" stands for "before symmetrization" and δ^n refers to the Dirac distribution in n dimensions; in particular,

$$\delta^n\Big(x_{\setminus\{\nu\}}^{(n+1)} - y^{(n)}\Big) = \prod_{\mu=1}^{\nu-1}\delta\Big(x_\mu^{(n+1)} - y_\mu^{(n)}\Big)\prod_{\mu=\nu}^{n}\delta\Big(x_{\mu+1}^{(n+1)} - y_\mu^{(n)}\Big). \tag{30}$$

The term under the sum in Eq. (29) describes (i) the probability per unit time that the νth particle disappears from position $x_\nu^{(n+1)}$ and (ii) the fact that the rest of the configuration remains unchanged so that its probability is transferred from $\rho_{n+1}\Big(x^{(n+1)}\Big)$ to $\rho_n\Big(x_1^{(n+1)},\dots,x_{\nu-1}^{(n+1)},x_{\nu+1}^{(n+1)}\dots,x_{n+1}^{(n+1)}\Big)$. The summation over ν accounts for the fact that any of the particles out of configuration $x^{(n+1)}$ might decay.

The properly symmetrized version of Eq. (29) is obtained by averaging over all permutations of the target space configurations $y^{(n)}$, i.e.,

$$\Lambda_{n+1}\Big(y^{(n)},x^{(n+1)}\Big) = \frac{1}{n!}\sum_{P\in\mathcal{P}_n}\Lambda_{n+1}^{\mathrm{bs}}\Big(Py^{(n)},x^{(n+1)}\Big). \tag{31}$$

Owing to the summation over ν in Eq. (29), this formula is already symmetric with respect to permutations of the second argument $x^{(n+1)}$. In turn, averaging over the permutations in $\mathcal{P}_n$ guarantees that the probability associated with a particle disappearing from the $(n+1)$-particle configuration $x^{(n+1)}$ is distributed symmetrically to that of all equivalent n-particle configurations on the receiving end.

Now, the crucial step is to insert the propensities into Eq. (27) and translate the expressions given in Eq. (31) into combinatorial factors. Due to the particle exchange symmetry of ρ_{n+1}, the contribution of any of the terms under the sum in Eq. (31) to $(\mathcal{G}_n\rho_{n+1})\left(y^{(n)}\right)$ from Eq. (27b) obeys [see also Eq. (29)]

$$\begin{aligned}\int_{\mathbb{X}^{n+1}} &\lambda_d\left(x_\nu^{(n+1)}\right)\delta^n\left(x_{\setminus\{\nu\}}^{(n+1)} - Py^{(n)}\right)\rho_{n+1}\left(x^{(n+1)}\right)dx^{(n+1)}\\ &= \int_{\mathbb{X}}\lambda_d(x)\ \rho_{n+1}\left(\left(Py^{(n)}\right)_1,\ldots,\left(Py^{(n)}\right)_{\nu-1},x,\left(Py^{(n)}\right)_\nu,\ldots,\left(Py^{(n)}\right)_n\right)dx\\ &= \int_{\mathbb{X}}\lambda_d(x)\ \rho_{n+1}\left(Py^{(n)},x\right)\ dx = \int_{\mathbb{X}}\lambda_d(x)\ \rho_{n+1}\left(y^{(n)},x\right)\ dx,\end{aligned} \tag{32}$$

i.e., they are all the same. Summation of this expression over ν [see Eq. (29)] yields a factor of $(n+1)$, and summation over the n-particle permutations $P \in \mathcal{P}_n$ together with the division by $n!$ [see Eq. (31)] ensures that the $(n+1)$-particle probability is distributed symmetrically over the n-particle space.

A similar calculation for the loss $(\mathcal{L}_n\rho_n)(x^{(n)})$ reads

$$\rho_n\left(x^{(n)}\right)\lambda_d\left(x_\nu^{(n)}\right)\int_{\mathbb{X}^{n-1}}\delta^{n-1}\left(x_{\setminus\{\nu\}}^{(n)} - Py^{(n-1)}\right)dy^{(n-1)} = \rho_n\left(x^{(n)}\right)\lambda_d\left(x_\nu^{(n)}\right) \tag{33}$$

for each permutation P and each index ν, where we translated formula (31) for Λ_{n+1} to Λ_n by a shift in n. Summation over the n-particle permutations $P \in \mathcal{P}_n$ cancels the factor $1/n!$. By the summation over ν and combining with the result for the gain, we obtain the evolution equation for the n-particle density under a simple decay process,

$$\partial_t\rho_n\left(x^{(n)}\right) = (n+1)\int_{\mathbb{X}}\lambda_d(y)\ \rho_{n+1}\left(x^{(n)},y\right)\ dy - \rho_n\left(x^{(n)}\right)\sum_{\nu=1}^{n}\lambda_d\left(x_\nu^{(n)}\right), \tag{34}$$

and this is in line with the reaction terms in Eq. (14).

B. Many-particle propensity for multiple species

We continue with the scenario of multiple species as described in Sec. II B. Let $\lambda(y;\ x_A, x_B)$ again denote the conditional probability per unit time that the reaction $A + B \to C$ occurs with a product particle of species C appearing in y, given that two reactants A and B reside in x_A and x_B, respectively. Then, we are interested in the associated many-particle propensities,

$$\Lambda\left(y^{(a-1)},y^{(b-1)},y^{(c+1)};x^{(a)},x^{(b)},x^{(c)}\right), \tag{35}$$

which denotes the transfer of probability density per unit time from $\rho_{a,b,c}$ to $\rho_{a-1,b-1,c+1}$ due to the considered reaction. Note that we have here suppressed subscripts a, b, c on Λ to simplify notation.

At first, we define for each tuple of indices $\nu,\ \mu, \xi$ the propensity

$$\Lambda_{\nu,\mu,\xi}\left(y^{(a-1)},y^{(b-1)},y^{(c+1)};x^{(a)},x^{(b)},x^{(c)}\right) = \lambda\left(y_\xi^{(c+1)};x_\mu^{(a)},x_\nu^{(b)}\right)\delta^{a-1}\left(x_{\setminus\{\mu\}}^{(a)} - y^{(a-1)}\right)\delta^{b-1}\left(x_{\setminus\{\nu\}}^{(b)} - y^{(b-1)}\right)\delta^{c}\left(x^{(c)} - y_{\setminus\{\xi\}}^{(c+1)}\right). \tag{36}$$

The interpretation of the expression in Eq. (36) is as follows: Given the reactant and product tuples in the source and target spaces, $\left(x^{(a)},x^{(b)},x^{(c)}\right)$ and $\left(y^{(a-1)},y^{(b-1)},y^{(c+1)}\right)$, respectively, it assigns the (probability) transfer rate $\lambda(y_\xi^{(c+1)};x_\mu^{(a)},x_\nu^{(b)})$ to the reaction occurring between the reactants located at $x_\mu^{(a)},x_\nu^{(b)}$ and producing a product particle in $y_\xi^{(c+1)}$. The products of δ-distributions make sure that in the transfer, all other particle positions remain those from the source space tuples.

In analogy to Eq. (29), we can now write down the many particle propensity before symmetrization as

$$\Lambda^{\mathrm{bs}}\left(y^{(a-1)},y^{(b-1)},y^{(c+1)};x^{(a)},x^{(b)},x^{(c)}\right) = \frac{1}{c+1}\sum_{\mu=1}^{a}\sum_{\nu=1}^{b}\sum_{\xi=1}^{c+1}\Lambda_{\nu,\mu,\xi}\left(y^{(a-1)},y^{(b-1)},y^{(c+1)};x^{(a)},x^{(b)},x^{(c)}\right). \tag{37}$$

The prefactor of $1/(c+1)$ is to be included for the following reason: If $y^{(a-1)},y^{(b-1)}$ are the same as $x^{(a)},x^{(b)}$ after the removal of $x_\mu^{(a)},x_\nu^{(b)}$ and if $y^{(c+1)}$ after the removal of $y_\xi^{(c+1)}$ agrees with $x^{(c)}$, then there are $c+1$ possibilities of augmenting $x^{(c)}$ with the target position $y_\xi^{(c+1)}$ to generate a $(c+1)$-tuple. The probability that out of the reaction of reactants at $x_\mu^{(a)},x_\nu^{(b)}$ emerges a particle in $y_\xi^{(c+1)}$ must be equi-distributed over these equivalent configurations of $c+1$ product particles to retain the required particle exchange symmetry.

Regarding the symmetrization, we observe that there are $(a-1)!(b-1)!(c+1)!$ equivalent configurations in the target space over which the probability of being transferred to has to be distributed. In analogy to Eq. (31), we obtain

$$\Lambda\Big(y^{(a-1)},y^{(b-1)},y^{(c+1)};x^{(a)},x^{(b)},x^{(c)}\Big)=\sum_{P\in\mathcal{P}_{a-1}}\sum_{Q\in\mathcal{P}_{b-1}}\sum_{R\in\mathcal{P}_{c+1}}\frac{\Lambda^{\mathrm{bs}}\Big(Py^{(a-1)},Qy^{(b-1)},Ry^{(c+1)};x^{(a)},x^{(b)},x^{(c)}\Big)}{(a-1)!\,(b-1)!\,(c+1)!}. \tag{38}$$

The formula in Eq. (38) is obviously symmetric with respect to the target space configurations by construction. It is also symmetric with respect to the source space configurations because of the summation over all possible pairs of reactant particles in Eq. (37) and the symmetrization over the target space configurations in Eq. (38).

Let us now derive the structure of the loss and gain expressions analogous to those in Eq. (27) for this representation of the many-particle propensity.

1. The loss term $\mathcal{L}$

Extending the definition in Eq. (27a) to the two-species reaction and dropping the superscript on $\mathcal{L}$ as it is clear from the context, we have

$$(\mathcal{L}\rho_{a,b,c})\Big(x^{(a)},x^{(b)},x^{(c)}\Big)=\rho_{a,b,c}\Big(x^{(a)},x^{(b)},x^{(c)}\Big)\sum_{\mu=1}^{a}\sum_{\nu=1}^{b}\int_{\mathbb{X}}\lambda\Big(y;x_\mu^{(a)},x_\nu^{(b)}\Big)dy. \tag{39}$$

To obtain this result, we have used that integration over just one of the terms in the multiple sum over particle indices in Eq. (37) and permutations in Eq. (38) may be summarized as follows: Dropping the prefactors of $1/(c+1)$ and $\rho_{a,b,c}(x^{(a)},x^{(b)},x^{(c)})/(a-1)!\,(b-1)!\,(c+1)!$ for the moment, we consider only the terms relevant for the integration, i.e.,

$$\begin{aligned}&\int_{\mathbb{X}^{a-1}\times\mathbb{X}^{b-1}\times\mathbb{X}^{c+1}}\Lambda_{\nu,\mu,\xi}\Big(Py^{(a-1)},Qy^{(b-1)},Ry^{(c+1)};x^{(a)},x^{(b)},x^{(c)}\Big)dy^{(a-1)}\,dy^{(b-1)}\,dy^{(c+1)}\\&=\int_{\mathbb{X}^{a-1}\times\mathbb{X}^{b-1}\times\mathbb{X}^{c+1}}\Lambda_{\nu,\mu,\xi}\Big(y^{(a-1)},y^{(b-1)},y^{(c+1)};x^{(a)},x^{(b)},x^{(c)}\Big)dy^{(a-1)}\,dy^{(b-1)}\,dy^{(c+1)}\\&=\int_{\mathbb{X}}\lambda\Big(y;x_\mu^{(a)},x_\nu^{(b)}\Big)\,dy.\end{aligned} \tag{40}$$

Here, the first equality follows by a transformation of the integration variables from the components of $\Big(y^{(a-1)},y^{(b-1)},y^{(c+1)}\Big)$ to the components of $\Big(Py^{(a-1)},Qy^{(b-1)},Ry^{(c+1)}\Big)$ and relabeling. The second equality follows because all the δ-distributions in Eq. (36) will generate unity once upon the integrations over $y_i^{(a-1)}$ $(i=1,\ldots,a-1)$, $y_j^{(b-1)}$ $(j=1,\ldots,b-1)$, and $y_k^{(c+1)}$ $(k=1,\ldots,\xi-1,\xi+1,c+1)$, whereas the integration over $y=y_\xi^{(c+1)}$ remains non-trivial. Thus, we observe that all these terms are identical for any of the $c+1$ terms in the sum over ξ in Eq. (37) and also for any of the permutations in Eq. (38). Carrying out the summation over the permutations yields a factor of $(a-1)!(b-1)!(c+1)!$, which cancels the denominator in Eq. (38), while summing over ξ in Eq. (37) cancels the factor of $1/(c+1)$ in that equation. This establishes Eq. (39).

2. The gain term $\mathcal{G}$

To calculate the gain operator $\mathcal{G}$ for the target space, $\mathbb{X}^a\times\mathbb{X}^b\times\mathbb{X}^c$, of the reaction, we have to compute the expectation of the propensity over the source space, $\mathbb{X}^{a+1}\times\mathbb{X}^{b+1}\times\mathbb{X}^{c-1}$, in analogy with Eq. (27b). The associated density-weighted integration over $\Big(x^{(a+1)},x^{(b+1)},x^{(c-1)}\Big)$ of $\Lambda_{\nu,\mu,\xi}$ in Eq. (36) yields

$$\begin{aligned}&\int_{\mathbb{X}^a\times\mathbb{X}^b\times\mathbb{X}^c}\Lambda_{\nu,\mu,\xi}\Big(y^{(a-1)},y^{(b-1)},y^{(c+1)};x^{(a)},x^{(b)},x^{(c)}\Big)\rho_{a,b,c}\Big(x^{(a)},x^{(b)},x^{(c)}\Big)\,dx^{(a)}\,dx^{(b)}\,dx^{(c)}\\&=\int_{\mathbb{X}^2}\lambda(y_\xi^{(c+1)};z,z',)\;\rho_{a,b,c}\Big(\Big(y^{(a-1)},z\Big),\Big(y^{(b-1)},z'\Big),y_{\setminus\{\xi\}}^{(c+1)}\Big)\,dz\,dz'.\end{aligned} \tag{41}$$

Here, we have already used the symmetry properties of $\rho_{a,b,c}$ to shift the remaining integration variables z and z' to the end of the tuples of its first two arguments. These calculations show that the result is again independent of the summation indices μ,ν so that the summation over these indices in Eq. (37) just generates a prefactor of ab. Summation over ξ guarantees that the configuration $\Big(y^{(a-1)},y^{(b-1)},y^{(c+1)}\Big)$ receives its appropriate share of probability transfer from all reactions that produce a particle in any of the positions collected in the tuple $y^{(c+1)}$.

Any permutation of $y^{(a-1)}$ or $y^{(b-1)}$ will not change the result either owing to the symmetry of $\rho_{a,b,c}$ in its first two arguments. Therefore, the averaging over these permutations will just cancel the prefactor of $1/(a-1)!(b-1)!$ in Eq. (38). After the summation over ξ in Eq. (37), the resulting expression is also invariant under permutations of $y^{(c+1)}$ owing to the symmetry of $\rho_{a,b,c}$ in its last argument. Thus, the summation over these permutations will just generate a factor of $(c+1)!$, canceling the remaining factor in the denominator of Eq. (38). Note, however, that the factor of $1/(c+1)$ from Eq. (37) is retained in the process.

The result for the gain function reads

$$(\mathcal{G}\rho_{a,b,c})\Big(y^{(a-1)},y^{(b-1)},y^{(c+1)}\Big) = \frac{ab}{c+1}\sum_{\xi=1}^{c+1}\int_{\mathbb{X}^2}\lambda\Big(y_\xi^{(c+1)};z,z'\Big)\rho_{a,b,c}\Big(\Big[y^{(a-1)},z\Big],\Big[y^{(b-1)},z'\Big],y^{(c+1)}_{\setminus\{\xi\}}\Big)\; dz\; dz'. \tag{42}$$

After a shift from (a,b,c) to $(a+1,b+1,c-1)$, we obtain an operator that agrees with the mid-term in Eq. (26).

Conservation of the total probability under the loss and gain functions in Eqs. (39) and (42) is guaranteed as we have

$$\int_{\mathbb{X}^a\times\mathbb{X}^b\times\mathbb{X}^c}(\mathcal{L}\rho_{a,b,c})\Big(x^{(a)},x^{(b)},x^{(c)}\Big)\; dx^{(a)}\; dx^{(b)}\; dx^{(c)}$$

$$= \int_{\mathbb{X}^{a-1}\times\mathbb{X}^{b-1}\times\mathbb{X}^{c+1}}(\mathcal{G}\rho_{a,b,c})\Big(y^{(a-1)},y^{(b-1)},y^{(c+1)}\Big)\; dy^{(a-1)}\; dy^{(b-1)}\; dy^{(c+1)} \tag{43}$$

$$= a\; b\int_{\mathbb{X}^3}\lambda(y;z,z')\int_{\mathbb{X}^{a-1}\times\mathbb{X}^{b-1}\times\mathbb{X}^c}\rho_{a,b,c}\big((\xi,z),(\eta,z'),\zeta\big)\; d\xi d\eta d\zeta\; dydzdz'. \tag{44}$$

Collecting the loss and gain terms and adding the diffusion terms, we obtain again the CDME given by Eq. (26).

In total, we end up with the same equation, but the way to get there is different: In Sec. II, we have expressed the loss and gain operators as sums of *local* operators (acting on subsets of particles), while here in Sec. III, we have translated the local rate functions into many-particle propensities. In Sec. IV, the combinatorics will be encoded in the annihilation and creation operators, again ending up in the same CDME.

IV. CDME FORMULATION USING CREATION AND ANNIHILATION OPERATORS

Using creation and annihilation operators as presented in Ref. 16, we can formulate the CDME at once without having to worry about the combinatorial factors. Assuming a system involving only one chemical species, we introduce the creation and annihilation operators acting on an n-particle density ρ_n as[16]

$$\big(a^+\{w\}\rho_n\big)(x^{(n+1)}) = \frac{1}{n+1}\sum_{j=1}^{n+1} w(x_j^{(n+1)})\rho_n(x^{(n+1)}_{\setminus\{j\}}), \tag{45a}$$

$$\big(a^-\{f\}\rho_n\big)(x^{(n-1)}) = n\int_{\mathbb{X}} f(y)\rho_n\Big(x^{(n-1)},y\Big)\; dy. \tag{45b}$$

The creation operator $a^+\{w\}$ adds a particle of species A with distribution w by multiplying the single-particle density w with the density ρ_n. The resulting density is a function of $n+1$ positions, $x^{(n+1)}$, and the sum over j and the prefactor are required to render the result symmetric with respect to permutations of particle labels. The annihilation operator $a^-\{f\}$ removes a particle at x with the rate $f(x)$ by marginalization of the density with the weight function f. As ρ_n is symmetric, we can simply integrate against the last variable. The resulting density is a function of $x^{(n-1)}$. As there are n possible ways to remove a particle, the factor of n appears in front of the integral. In Ref. 16, it was shown that the creation and annihilation operators satisfy some special properties that are useful for calculations, including the commutation relations,

$$\big[a^-\{f\},a^+\{w\}\big] = \langle f,w\rangle, \qquad \big[a^-\{f\},a^-\{g\}\big] = \big[a^+\{w\},a^+\{v\}\big] = 0, \tag{46}$$

where $\langle u,v\rangle := \int_{\mathbb{X}} u(x)v(x)dx$ for suitable functions u,v and $[a,b] := ab-ba$ for operators a,b. Furthermore, the definitions of a^+ and a^- extend naturally to the family of n-particle densities by operating element-wise, e.g., $a^+\{w\}(\rho_0,\rho_1,\ldots) = (a^+\{w\}\rho_0, a^+\{w\}\rho_1,\ldots)$.

The following representation of the CDME will be given in terms of a basis $(u_1,u_2,\ldots)$ of the space of single-particle densities. We emphasize that the obtained results are independent of the specific basis chosen although the expansion coefficients will naturally depend on the choice of the basis. For a concrete application, the basis functions can be adapted to the problem and reflect some physical properties, e.g., possible symmetries. We recall that, in quantum mechanics, the common expansions in terms of spherical harmonics and associated polynomials are motivated by the isotropy of atoms. For keeping the presentation concise, we restrict here to square-integrable probability densities, which form the separable Hilbert space $L^2(\mathbb{X})$, and we require that the basis is orthonormal, i.e., $\langle u_\alpha,u_\beta\rangle = \delta_{\alpha,\beta}$, where $\langle\cdot,\cdot\rangle$ denotes the inner product in $L^2(\mathbb{X})$. More generally, one uses the Banach space $L^1(\mathbb{X})$ of integrable functions as it was done in Ref. 16. However, this adds a number of technical issues, and there are no relevant differences in the final expressions. In both cases, the existence of a basis $(u_1, u_2, \ldots)$ is granted, and in the $L^1(\mathbb{X})$ case, the representations are exact in the sense that every probability density can be expanded in such a basis.

A. One species

Let us consider again a general one-species reaction $kA \to lA$ with rate function $\lambda(y^{(l)};\, x^{(k)})$. Under the assumption of independently and identically diffusing particles, following Eq. (8), the diffusion operator $\mathcal{D}_n$ decomposes into a sum of single-particle diffusions D_ν applied to the νth particle, which was shown to have an expansion in terms of creation and annihilation operators,[16]

$$\begin{aligned}\mathcal{D}_n &= \sum_{\nu=1}^{n} D_\nu \\ &= \sum_{\alpha,\beta} \langle u_\alpha, D_1 u_\beta\rangle a_\alpha^+ a_\beta^-,\end{aligned} \tag{47}$$

where we used the compressed notations $a_\alpha^+ := a^+\{u_\alpha\}$ and $a_\beta^- := a^-\{u_\beta\}$. One observes that the expansion in Eq. (47) does not depend explicitly on the particle number n, and thus, formally, it represents the full diffusion operator acting on the whole family $\rho = (\rho_0, \rho_1, \ldots)$. In the presence of pair interactions between the particles, the diffusion operator is made of two-particle operators, which can be expanded into products $a_\alpha^+ a_\beta^+ a_\gamma^- a_\delta^-$; see Sec. 4.1.2 of Ref. 16. Equation (47) is easily seen for the case $n = 1$. For an arbitrary single-particle density ρ_1, we have the basis expansion $\rho_1 = \sum_\alpha \langle u_\alpha, \rho_1\rangle u_\alpha$, and for the action of $\mathcal{D}_1$, one finds

$$\mathcal{D}_1\rho_1 = \sum_\alpha \langle u_\alpha, D_1\rho_1\rangle\ u_\alpha = \sum_{\alpha\beta} \langle u_\alpha, D_1 u_\beta\rangle \langle u_\beta, \rho_1\rangle\ u_\alpha, \tag{48}$$

which corresponds to Eq. (47) upon verifying that $a_\alpha^+ a_\beta^- \rho_1 = u_\alpha \langle u_\beta, \rho_1\rangle$ from (45). The general case of an n-particle density ρ_n is obtained by noting the representation $\rho_n = \sum_{\alpha_1 \leqslant \cdots \leqslant \alpha_n} c_{\alpha_1,\ldots,\alpha_n} a_{\alpha_1}^+ \cdots a_{\alpha_n}^+ \rho_{\text{vac}}$, where $c_{\alpha_1,\ldots,\alpha_n}$ are the expansion coefficients and $\rho_{\text{vac}} = 1$ denotes the vacuum state $(n = 0)$.

We now need to expand the loss and gain operators in the same manner. First, we consider the loss and gain operators per reaction from Eqs. (9) and (12), which are linear operators and are thus fully specified by their action on products of single-particle basis functions,

$$\big(L(u_{\beta_1} \otimes \cdots \otimes u_{\beta_k})\big)(x^{(k)}) := (u_{\beta_1} \otimes \cdots \otimes u_{\beta_k})(x^{(k)}) \int_{\mathbb{X}^l} \lambda(y^{(l)}; x^{(k)})\, dy^{(l)}, \tag{49}$$

$$\big(G(u_{\beta_1} \otimes \cdots \otimes u_{\beta_k})\big)(y^{(l)}) := \int_{\mathbb{X}^k} \lambda(y^{(l)}; x^{(k)}) (u_{\beta_1} \otimes \cdots \otimes u_{\beta_k})(x^{(k)})\, dx^{(k)}, \tag{50}$$

with the tensor product $v_1 \otimes \cdots \otimes v_n = \bigotimes_{j=1}^n v_j$ defined as $(v_1 \otimes \cdots \otimes v_n)(x^{(n)}) := v_1(x_1^{(n)}) \cdots v_n(x_n^{(n)})$. One can show that $(u_{\alpha_1} \otimes \cdots \otimes u_{\alpha_n})_{\alpha_i \in \mathbb{N}}$ is a basis of the corresponding tensor space of Hilbert spaces, which is itself a Hilbert space, referred to as a Fock space. Analogous to the diffusion operator, the total loss and gain over all possible reactions from Eqs. (10) and (13) also have expansions in terms of creation and annihilation operators,[16]

$$\mathcal{L}_n = \frac{1}{k!} \sum_{\substack{\alpha_1,\ldots,\alpha_k \\ \beta_1,\ldots,\beta_k}} \left\langle \bigotimes_{i=1}^k u_{\alpha_i}, L \bigotimes_{j=1}^k u_{\beta_j} \right\rangle \prod_{i=1}^k a_{\alpha_i}^+ \prod_{j=1}^k a_{\beta_j}^-, \tag{51}$$

$$\mathcal{G}_n = \frac{1}{k!} \sum_{\substack{\alpha_1,\ldots,\alpha_l \\ \beta_1,\ldots,\beta_k}} \left\langle \bigotimes_{i=1}^l u_{\alpha_i}, G \bigotimes_{j=1}^k u_{\beta_j} \right\rangle \prod_{i=1}^l a_{\alpha_i}^+ \prod_{j=1}^k a_{\beta_j}^-. \tag{52}$$

These expansions seem to be rather involved at first sight, yet they is a key element to develop a straightforward formulation of the CDME even for complex reaction–diffusion networks. The structure of the expressions becomes more transparent by introducing the following short-hand notation. Let $\mathrm{a}^+ = (a^+\{u_\alpha\})_{\alpha \in \mathbb{N}}$ denote the family of creation operators for the basis (u_α), and analogously, $\mathrm{a}^- = (a^-\{u_\beta\})_{\beta \in \mathbb{N}}$. For the coefficients of $\mathcal{D}_n$ in Eq. (47), we arrange them as $\mathrm{D} = (\langle u_\alpha, D u_\beta\rangle)_{(\alpha,\beta) \in \mathbb{N}^2}$, which is reminiscent of a tensor of rank 2. The expansion of $\mathcal{D}_n$ then reads

$$\mathcal{D} = \mathrm{a}^+\ \mathrm{D}\mathrm{a}^-, \tag{53}$$

where the products between the symbols in upright font face imply full contractions of "tensor" indices α and β; see Eq. (47); here, we have dropped the subscript n from $\mathcal{D}_n$ noting again that the right-hand side holds for any n. For the loss and gain terms, we make use of multi-indices $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)$ and write $(\mathrm{a}^+)^k = (a^+\{u_{\alpha_1}\} \cdots a^+\{u_{\alpha_k}\})_{\boldsymbol{\alpha} \in \mathbb{N}^k}$ and analogously for $(\mathrm{a}^-)^l$. The coefficients of $\mathcal{L}_n$ in Eq. (51) are denoted as $\mathrm{L} = \left(\langle u_{\alpha_1} \otimes \cdots \otimes u_{\alpha_k}, L u_{\beta_1} \otimes \cdots \otimes u_{\beta_k} \rangle\right)_{\boldsymbol{\alpha\beta}}$. With this compact notation, the expansions of the gain and loss operators in Eqs. (51) and (52) take the form (Fig. 3)

$$\mathcal{L} = (\mathrm{a}^+)^k \ \mathrm{L} \ (\mathrm{a}^-)^k \qquad \text{and} \qquad \mathcal{G} = (\mathrm{a}^+)^l \ \mathrm{G} \ (\mathrm{a}^-)^k, \tag{54}$$

with products implying contractions over multi-indices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$; additionally, we agree that contractions involving several annihilation operators $(\mathrm{a}^-)^k$ introduce a factor of $k!$, corresponding to the length of the multi-index $\boldsymbol{\beta}$. Then, the CDME in its compact form is

$$\frac{\partial \rho}{\partial t} = \left(\mathrm{a}^+ \ \mathrm{D}\mathrm{a}^- + (\mathrm{a}^+)^l \ \mathrm{G} \ (\mathrm{a}^-)^k - (\mathrm{a}^+)^k \ \mathrm{L} \ (\mathrm{a}^-)^k\right)\rho. \tag{55}$$

To recover the explicit form of the equation (as derived in Secs. II and III), one must explicitly evaluate the expressions containing the creation and annihilation operators. To circumvent these often cumbersome calculations, we provide a dictionary of the expansions for common reactions in the Appendix, where we can easily verify that the nth component of this equation matches that of Eq. (7), where the loss operator always acts on ρ_n and the gain operator acts on ρ_{n+k-l}.

The compact notation has a very intuitive logic behind (Fig. 3): Given the reaction $kA \to lA$, the loss acts on the k reactants at once, so it involves k creation and k annihilation operators. As the gain depends on both reactants and products, it consists of k annihilation and l creation operators. The diffusion operator, as it acts on solely one particle at a time, involves only one annihilation and one creation operator. If diffusion incorporated physical pair interactions, it would act on two particles at a time, so it would involve two creation and two annihilation operators.

B. Bimolecular reactions

For reaction systems involving multiple species, it is equally easy to obtain the desired equation. We only need to use different creation and annihilation operators for each species. For examples, for the reaction

$$A + B \to C, \tag{56}$$

with rate function $\lambda(y_C;\ x_A, x_B)$, where x_A and x_B are the locations of the reactants and y_C is the location of the product, we immediately obtain

$$\frac{\partial \rho}{\partial t} = \left(\mathrm{a}^+ \ \mathrm{D}^A \ \mathrm{a}^- + \mathrm{b}^+ \ \mathrm{D}^B \ \mathrm{b}^- + \mathrm{c}^+ \ \mathrm{D}^C \ \mathrm{c}^- + \mathrm{c}^+ \ \mathrm{G} \ \mathrm{a}^-\mathrm{b}^- - \mathrm{a}^+\mathrm{b}^+ \ \mathrm{L} \ \mathrm{a}^-\mathrm{b}^-\right)\rho, \tag{57}$$

where ρ is the family of n-particle densities of the form $\rho_{a,b,c}(x^{(a)}, x^{(b)}, x^{(c)})$ for all possible values of the particle numbers a, b, and c. The creation and annihilation operators for each species are denoted by the corresponding lower case letter. The first three terms describe the diffusion of the different species, the fourth term is the total loss due to reactions, and the last term is the total gain. Note that the loss of probability will only depend on the number of reactants of the current state; thus, it only contains operators for the A and B species. On the other hand, the gain will depend on the number of reactants in another state and the products needed to bring the system to the current state. These terms have the following expansions:[16]

$$\mathrm{a}^+\mathrm{b}^+ \ \mathrm{L} \ \mathrm{a}^-\mathrm{b}^- = \frac{1}{2}\sum_{\substack{\alpha_1,\alpha_2\\ \beta_1,\beta_2}} \langle u_{\alpha_1} \otimes u_{\alpha_2}, L(u_{\beta_1} \otimes u_{\beta_2})\rangle a^+\{u_{\alpha_1}\} b^+\{u_{\alpha_2}\} a^-\{u_{\beta_1}\} b^-\{u_{\beta_2}\}, \tag{58}$$

$$\mathrm{c}^+ \ \mathrm{G} \ \mathrm{a}^-\mathrm{b}^- = \frac{1}{2}\sum_{\substack{\alpha\\ \beta_1,\beta_2}} \langle u_\alpha, G(u_{\beta_1} \otimes u_{\beta_2})\rangle c^+\{u_\alpha\} a^-\{u_{\beta_1}\} b^-\{u_{\beta_2}\}, \tag{59}$$

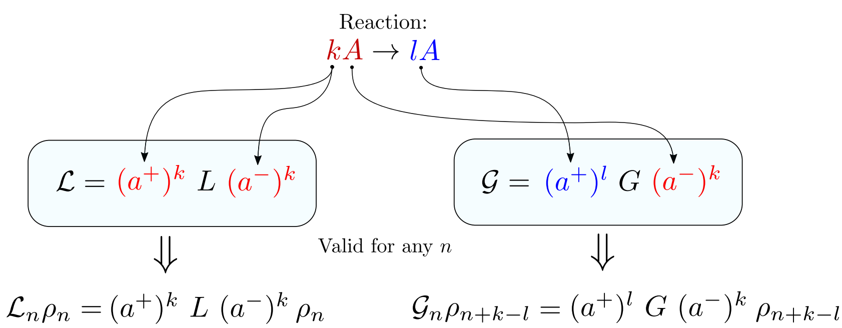

FIG. 3. Diagram representing how to write the loss and gain operators for the reaction $kA \to lA$ in the compact notation using creation and annihilation operators. The operators L and G represent the *local* loss and gain operators; the operators $\mathcal{L}$ and $\mathcal{G}$ represent the *global* loss and gain operators.

recalling the short-hand $a^+_{\alpha_1} = a^+\{u_{\alpha_1}\}$, etc. The *local* loss and gain operators are given in terms of the rate function λ as

$$\big(L(u_{\beta_1}\otimes u_{\beta_2})\big)(x_A,x_B) := (u_{\beta_1}\otimes u_{\beta_2})(x_A,x_B)\int_{\mathbb{X}}\lambda(y_C;x_A,x_B)dy_C, \tag{60}$$

$$\big(G(u_{\beta_1}\otimes u_{\beta_2})\big)(y_C) := \int_{\mathbb{X}^2}(u_{\beta_1}\otimes u_{\beta_2})(x_A,x_B)\lambda(y_C;x_A,x_B)dx_Adx_B, \tag{61}$$

in analogy to Eqs. (49) and (50).

Using the dictionary of the Appendix, it is straightforward to transform Eq. (57) into the explicit integral notation,

$$\begin{aligned}\frac{\partial\rho_{a,b,c}}{\partial t} &= \sum_{\mu=1}^{a}D^A_\mu\rho_{a,b,c} + \sum_{\nu=1}^{b}D^B_\nu\rho_{a,b,c} + \sum_{\xi=1}^{c}D^C_\xi\rho_{a,b,c}\\ &+ \frac{(a+1)(b+1)}{c}\sum_{\xi=1}^{c}\Big(\int_{\mathbb{X}^2}\lambda\Big(x^{(c)}_\xi;z,z'\Big)\rho_{a+1,b+1,c-1}\Big((x^{(a)},z),(x^{(b)},z'),x^{(c)}_{\setminus\{\xi\}}\Big)dzdz'\Big) - \rho_{a,b,c}\sum_{\nu_1=1}^{a}\sum_{\nu_2=1}^{b}\int_{\mathbb{X}}\lambda(y;x^{(a)}_{\nu_1},x^{(b)}_{\nu_2})dy,\end{aligned} \tag{62}$$

which is the same as Eq. (26).

C. Enzyme kinetics

Before closing, we develop the CDME for a real-world example, namely, the Michaelis–Menten scheme for enzyme kinetics, which consists of three reactions and involves four species,

$$R_1 : E + S \to C, \tag{63a}$$

$$R_2 : C \to E + S, \tag{63b}$$

$$R_3 : C \to E + P. \tag{63c}$$

The scheme describes an enzyme E that can bind a substrate molecule S to form the complex C. This complex can either dissociate again or yield a product P while releasing the original enzyme. The rate functions corresponding to these reactions are $\lambda_1(y_C;\ x_E,x_S)$, $\lambda_2(y_E,y_S;\ x_C)$, and $\lambda_3(y_E,y_P;\ x_C)$, respectively. The CDME is an evolution equation for the family ρ of densities of the form $\rho_{e,s,p,c}(x^{(e)},x^{(s)},x^{(p)},x^{(c)})$ for all possible values of e, s, p, and c, and it takes the form

$$\frac{\partial\rho}{\partial t} = \mathcal{D}\rho + \Bigg(\sum_{r=1}^{3}\mathcal{R}^{(r)}\Bigg)\rho \tag{64}$$

with the diffusion and reaction operators

$$\mathcal{D} = \mathrm{e}^+\ \mathrm{D}^E\ \mathrm{e}^- + \mathrm{s}^+\ \mathrm{D}^S\ \mathrm{s}^- + \mathrm{c}^+\ \mathrm{D}^C\ \mathrm{c}^- + \mathrm{p}^+\ \mathrm{D}^P\ \mathrm{p}^-, \tag{65a}$$

$$\mathcal{R}^{(1)} = \mathrm{c}^+\ \mathrm{G}_1\ \mathrm{e}^-\mathrm{s}^- - \mathrm{e}^+\mathrm{s}^+\ \mathrm{L}_1\ \mathrm{e}^-\mathrm{s}^-, \tag{65b}$$

$$\mathcal{R}^{(2)} = \mathrm{e}^+\mathrm{s}^+\ \mathrm{G}_2\ \mathrm{c}^- - \mathrm{c}^+\ \mathrm{L}_2\ \mathrm{c}^-, \tag{65c}$$

$$\mathcal{R}^{(3)} = \mathrm{e}^+\mathrm{p}^+\ \mathrm{G}_3\ \mathrm{c}^- - \mathrm{c}^+\ \mathrm{L}_3\ \mathrm{c}^-. \tag{65d}$$

The expansions of the operators, as well as the corresponding loss and gain operators for each reaction, are completely analogous to the previous examples. By virtue of the Appendix, we obtain the CDME in its integral notation,

$$\begin{aligned}\frac{\partial\rho_{e,s,p,c}}{\partial t} &= \sum_{\mu=1}^{e}D^E_\mu\rho_{e,s,p,c} + \sum_{\mu=1}^{s}D^S_\mu\rho_{e,s,p,c} + \sum_{\mu=1}^{p}D^P_\mu\rho_{e,s,p,c} + \sum_{\mu=1}^{c}D^C_\mu\rho_{e,s,p,c}\\ &+ \frac{(e+1)(s+1)}{c}\sum_{\xi=1}^{c}\Big(\int_{\mathbb{X}^2}\rho_{e+1,s+1,p,c-1}\Big((x^{(e)},z),(x^{(s)},z'),x^{(p)},x^{(c)}_{\setminus\{\xi\}}\Big)\lambda_1\Big(x^{(c)}_\xi;z,z'\Big)dzdz'\Big)\\ &+ \frac{(c+1)}{es}\sum_{\mu=1}^{e}\sum_{\eta=1}^{s}\Big(\int_{\mathbb{X}}\rho_{e-1,s-1,p,c+1}\Big(x^{(e)}_{\setminus\{\mu\}},x^{(s)}_{\setminus\{\eta\}},x^{(p)},(x^{(c)},z)\Big)\lambda_2\Big(x^{(e)}_\mu,x^{(s)}_\eta;z\Big)dz\Big)\\ &+ \frac{(c+1)}{ep}\sum_{\mu=1}^{e}\sum_{\eta=1}^{p}\Big(\int_{\mathbb{X}}\rho_{e-1,s,p-1,c+1}\Big(x^{(e)}_{\setminus\{\mu\}},x^{(s)},x^{(p)}_{\setminus\{\eta\}},(x^{(c)},z)\Big)\lambda_3\Big(x^{(e)}_\mu,x^{(p)}_\eta;z\Big)dz\Big)\\ &- \rho_{e,s,p,c}\Bigg(\sum_{\nu_1=1}^{e}\sum_{\nu_2=1}^{s}\int_{\mathbb{X}}\lambda_1(y;x^{(e)}_{\nu_1},x^{(s)}_{\nu_2})dy + \sum_{\nu=1}^{c}\int_{\mathbb{X}^2}\lambda_2(y_1,y_2;x^{(c)}_\nu)dy_1dy_2 + \sum_{\nu=1}^{c}\int_{\mathbb{X}^2}\lambda_3(y_1,y_3;x^{(c)}_\nu)dy_1dy_3\Bigg).\end{aligned} \tag{66}$$

D. Non-rigorous extension to Dirac δ-distributions

According to (45) of the annihilation and creation operators, a particle is inserted with a spatial probability density $w(x)$ and removed with a position-dependent rate function $f(x)$. From a physics perspective, classical particles have a defined position, and so it should be possible to add and delete particles at a single point $y \in \mathbb{X}$ (in this case, w would correspond to a point measure). To this end, we formally extend these operators to accept Dirac δ-distributions as their arguments, ignoring here any mathematical difficulties associated with it. For $\delta_y(x) := \delta(x - y)$, we define

$$(a^+\{\delta_y\}\rho_n)(x^{(n+1)}) = \frac{1}{n+1}\sum_{j=1}^{n+1} \delta(x_j^{(n+1)} - y)\rho_n(x_{\setminus\{j\}}^{(n+1)}), \tag{67a}$$

$$(a^-\{\delta_y\}\rho_n)(x^{(n-1)}) = n\int_{\mathbb{X}} \delta(z-y)\rho_n\Big(x^{(n-1)}, z\Big)\ dz = n\ \rho_n(x^{(n-1)}, y). \tag{67b}$$

For brevity, we will write $a^+(y) = a^+\{\delta_y\}$ and $a^-(y) = a^-\{\delta_y\}$ in the following. By direct substitution and straightforward calculations analogous to the ones in Ref. 16, one proves that these operators satisfy the commutation relations [see also Eq. (46)],

$$\big[a^-(y_1), a^+(y_2)\big] = \delta(y_1 - y_2), \qquad \big[a^-(y_1), a^-(y_2)\big] = \big[a^+(y_1), a^+(y_2)\big] = 0, \tag{68}$$

which agree with the corresponding expressions in quantum field theory.[28]

In general, for an operator A acting on a single particle at position y, such as diffusion, or an operator B acting on two particles at positions y_1 and y_2, we obtain the following representations of the corresponding Fock space operators [see also Eqs. (60) and (64) in Ref. 16]:

$$\mathcal{A} = \int_{\mathbb{X}\times\mathbb{X}} dxdy\ a^+(x)\ \tilde{A}(x;y)\ a^-(y), \tag{69}$$

$$\mathcal{B} = \frac{1}{2!}\int_{\mathbb{X}^2\times\mathbb{X}^2} dx_1dx_2dy_1dy_2\ a^+(x_1)a^+(x_2)\ \tilde{B}(x_1,x_2;y_1,y_2)\ a^-(y_1)a^-(y_2). \tag{70}$$

As a rule of thumb, given a basis expansion, such as Eq. (47), the functions u_α are replaced by δ_{x_α}, and the sums over α and β are replaced by integrals over the continuous variables x_α and y_β, respectively. The integral kernels $\tilde{A}$ and $\tilde{B}$ generalize the coefficient matrices and read $\tilde{A}(x;y) := \langle\delta_x, A\delta_y\rangle$ and $\tilde{B}(x_1,x_2;y_1,y_2) := \langle\delta_{(x_1,x_2)}, B\delta_{(y_1,y_2)}\rangle$, respectively; here, $\delta_{(x_1,\ldots,x_k)}(z_1,\ldots,z_k) := \delta(x_1 - z_1)\cdots\delta(x_k - z_k)$ denotes the k-dimensional Dirac δ-distribution.

In case of a "diagonal" operator, such as the loss operator L, the one-particle kernel reduces to $\tilde{A}(x;y) = \tilde{A}(y)\ \delta(x-y)$ for $\tilde{A}(y) := \tilde{A}(y,y)$ and Eq. (69) simplifies to [cf. Eq. (24) in Ref. 17]

$$\mathcal{A} = \int_{\mathbb{X}} dy\ a^+(y)\ \tilde{A}(y)\ a^-(y). \tag{71}$$

If A is a differential operator (e.g., the diffusion operator D), we note that $\tilde{A}(x;y)$ has to be interpreted in a distributional sense,

$$\int dy\ \varphi(y)\tilde{A}(x;y) = \int dy\ \varphi(y)\ \langle\delta_x, A\delta_y\rangle = \Big\langle\delta_x, A\Big(\int dy\ \varphi(y)\ \delta_y\Big)\Big\rangle = \langle\delta_x, A\varphi\rangle = (A\varphi)(x) \tag{72}$$

for suitable test functions φ.

For the *global* gain and loss operators of the reaction $kA \to lA$, we apply the same rules, starting from (51) and (52), respectively,

$$\mathcal{L} = \frac{1}{k!}\int_{\mathbb{X}^k\times\mathbb{X}^k} dx^{(k)}dy^{(k)}a^+(x^{(k)})\ \tilde{L}(x^{(k)};y^{(k)})\ a^-(y^{(k)}), \tag{73}$$

$$\mathcal{G} = \frac{1}{k!}\int_{\mathbb{X}^l\times\mathbb{X}^k} dx^{(l)}dy^{(k)}a^+(x^{(l)})\ \tilde{G}(x^{(l)};y^{(k)})\ a^-(y^{(k)}), \tag{74}$$

where $a^+(x^{(k)}) := a^+(x_1^{(k)})\cdots a^+(x_k^{(k)})$ yields the insertion of k particles at positions $x^{(k)} = (x_1^{(k)},\ldots,x_k^{(k)})$, and analogously $a^-(x^{(k)})$ for the removal of k particles; we note that the factors in these products commute. The coefficient functions are readily calculated from the definitions of the *local* loss and gain operators, L and G,

$$\tilde{L}(x^{(k)};y^{(k)}) := \langle\delta_{x^{(k)}}, L\delta_{y^{(k)}}\rangle \overset{(49)}{=} \delta(x^{(k)} - y^{(k)})\int_{\mathbb{X}^l} dz^{(l)}\ \lambda(z^{(l)};y^{(k)}), \tag{75}$$

$$\tilde{G}(x^{(l)};y^{(k)}) := \langle\delta_{x^{(l)}}, G\delta_{y^{(k)}}\rangle \overset{(50)}{=} \lambda(x^{(l)};y^{(k)}). \tag{76}$$

These results together with Eqs. (73) and (74) agree with Doi's work.[17]

For the action of products of the creation and annihilation operators, we find from Eq. (67) by induction

$$
\begin{aligned}
(a^{+}(y^{(k)})\rho_{n-k})(x^{(n)}) &= \Big(a^{+}(y_1^{(k)})a^{+}(y_{\setminus\{1\}}^{(k)})\rho_n\Big)(x^{(n)}) \\
&= \frac{1}{n}\sum_{j_1=1}^{n}\delta(x_{j_1}^{(n)}-y_1^{(k)})\Big(a^{+}(y_{\setminus\{1\}}^{(k)})\rho_{n-k}\Big)(x_{\setminus\{j_1\}}^{(n)}) \\
&= \frac{1}{n(n-1)}\sum_{j_1=1}^{n}\sum_{\substack{j_2=1\\ j_2\neq j_1}}^{n}\delta(x_{j_1}^{(n)}-y_1^{(k)})\delta(x_{j_2}^{(n)}-y_2^{(k)})\Big(a^{+}(y_{\setminus\{1,2\}}^{(k)})\rho_{n-k}\Big)(x_{\setminus\{j_1,j_2\}}^{(n)}) \\
&= \frac{(n-k)!}{n!}\sum_{j_1=1}^{n}\cdots\sum_{\substack{j_k=1\\ j_k\neq j_1,\ldots,j_{k-1}}}^{n}\delta(x_{j_1}^{(n)}-y_1^{(k)})\cdots\delta(x_{j_k}^{(n)}-y_k^{(k)})\rho_{n-k}(x_{\setminus\{j_1,\ldots,j_k\}}^{(n)}) \\
&= \frac{k!(n-k)!}{n!}\sum_{1\leqslant j_1<\cdots<j_k\leqslant n}\delta(x_{j_1}^{(n)}-y_1^{(k)})\cdots\delta(x_{j_k}^{(n)}-y_k^{(k)})\rho_{n-k}(x_{\setminus\{j_1,\ldots,j_k\}}^{(n)})
\end{aligned}
\tag{77}
$$

and, more immediately,

$$
(a^{-}(y^{(k)})\rho_n)(x^{(n-k)}) = \frac{n!}{(n-k)!}\rho_n(x^{(n-k)},y^{(k)}).
\tag{78}
$$

In combination with Eqs. (73) and (75), these results deliver the explicit form of the loss term of the CDME,

$$
\begin{aligned}
(\mathcal{L}\rho_n)(x^{(n)}) &= \frac{1}{k!}\int_{\mathbb{X}^l\times\mathbb{X}^k}\lambda(z^{(l)};y^{(k)})\Big(a^{+}(y^{(k)})a^{-}(y^{(k)})\rho_n\Big)(x^{(n)})dz^{(l)}dy^{(k)} \\
&\overset{(77)}{=} \frac{(n-k)!}{n!}\sum_{1\leqslant j_1<\cdots<j_k\leqslant n}\int_{\mathbb{X}^k}\Big(\int_{\mathbb{X}^l}\lambda(z^{(l)};y^{(k)})dz^{(l)}\Big) \\
&\quad\times\delta(x_{j_1}^{(n)}-q_1^{(k)})\cdots\delta(x_{j_k}^{(n)}-q_k^{(k)})\Big(a^{-}(y^{(k)})\rho_n\Big)(x_{\setminus\{j_1,\ldots,j_k\}}^{(n)})dy^{(k)} \\
&= \frac{(n-k)!}{n!}\sum_{1\leqslant j_1<\cdots<j_k\leqslant n}\Big(\int_{\mathbb{X}^l}\lambda(z^{(l)};x_{j_1,\ldots,j_k}^{(n)})dz^{(l)}\Big)\Big(a^{-}(x_{j_1,\ldots,j_k}^{(n)})\rho_n\Big)(x_{\setminus\{j_1,\ldots,j_k\}}^{(n)}) \\
&\overset{(78)}{=} \sum_{1\leqslant j_1<\cdots<j_k\leqslant n}\Big(\int_{\mathbb{X}^l}\lambda(z^{(l)};x_{j_1,\ldots,j_k}^{(n)})dz^{(l)}\Big)\rho_n(x^{(n)}).
\end{aligned}
\tag{79}
$$

Thereby, we have recovered Eq. (10), showing consistency between this approach and the one introduced in Sec. II. We can repeat this exercise for the gain operator using Eq. (74),

$$
\begin{aligned}
(\mathcal{G}\rho_{n+k-l})(x^{(n)}) &= \frac{1}{k!}\int_{\mathbb{X}^l\times\mathbb{X}^k}\lambda(z^{(l)};y^{(k)})\Big(a^{+}(z^{(l)})a^{-}(y^{(k)})\rho_{n+k-l}\Big)(x^{(n)})\ dz^{(l)}dy^{(k)} \\
&\overset{(77)}{=} \frac{l!(n-l)!}{n!k!}\sum_{1\leqslant j_1<\cdots<j_l\leqslant n}\int_{\mathbb{X}^l\times\mathbb{X}^k}\lambda(z^{(l)};y^{(k)}) \\
&\quad\times\delta(x_{j_1}^{(n)}-z_1^{(l)})\cdots\delta(x_{j_l}^{(n)}-z^{(l)_l})\Big(a^{-}(y^{(k)})\rho_{n+k-l}\Big)(x_{\setminus\{j_1\cdots j_l\}}^{(n)})\ dz^{(l)}dy^{(k)} \\
&= \frac{l!(n-l)!}{n!k!}\sum_{1\leqslant j_1<\cdots<j_l\leqslant n}\int_{\mathbb{X}^k}\lambda(x_{j_1,\ldots,j_l}^{(n)};y^{(k)})\Big(a^{-}(y^{(k)})\rho_{n+k-l}\Big)(x_{\setminus\{j_1\cdots j_l\}}^{(n)})\ dy^{(k)} \\
&\overset{(78)}{=} \frac{l!(n+k-l)!}{n!k!}\sum_{1\leqslant j_1<\cdots<j_l\leqslant n}\int_{\mathbb{X}^k}\lambda(x_{j_1,\ldots,j_l}^{(n)};y^{(k)})\rho_{n+k-l}(x_{\setminus\{j_1\cdots j_l\}}^{(n)},y^{(k)})\ dy^{(k)} \\
&= \binom{n}{l}^{-1}\binom{n+k-l}{k}\sum_{1\leqslant j_1<\cdots<j_l\leqslant n}\int_{\mathbb{X}^k}\lambda(x_{j_1,\ldots,j_l}^{(n)};y^{(k)})\rho_{n+k-l}(x_{\setminus\{j_1\cdots j_l\}}^{(n)},y^{(k)})\ dy^{(k)},
\end{aligned}
\tag{80}
$$

once again, recovering Eq. (13) from Sec. II. The relations in the dictionary from the Appendix are proved in a similar fashion, but using the expansions of Sec. IV as shown in Ref. 16.

We can further obtain a relation between the rate functions and the many-particle propensities by comparing the resulting loss from Eq. (79) with the many particle propensity in Eq. (27a). This relation holds regardless of the density,

$$\int_{\mathbb{X}^{n-k+l}} \Lambda_n(y^{(n-k+l)}; x^{(n)}) dy^{(n-k+l)} = \sum_{1 \leqslant j_1 < \cdots < j_k \leqslant n} \int_{\mathbb{X}^l} \lambda(y^{(l)}; x^{(n)}_{j_1,\ldots,j_k}) dy^{(l)}. \tag{81}$$

This establishes a connection with Sec. III. We can prove this identity independently by deriving the expressions of the many-particle propensities for the reaction $kA \to lA$.

V. DISCUSSION

We presented three approaches to formulate the CDME, the governing equation of stochastic particle-based reaction–diffusion dynamics. In general, the CDME consists of a diffusion operator, which describes the spatial transport of particles, and several reaction operators each corresponding to a chemical reaction in the system. Every reaction operator can further be separated into a loss operator and a gain operator for the probabilistic outflow and inflow, respectively.

In the first approach, these *global* loss and gain operators have been expressed as combinations of *local* loss and gain operators referring to reactions of subsets of reactants and products within the system. The central combinatorial factors, which come into play due to the particle exchange symmetry for molecules of the same species, have been justified by carefully applying combinatorial arguments for the random selection of subsets of particles out of a larger set. Although this approach is intuitive and relatively straightforward, it requires computing the combinatorial factors of the reaction operators by hand, and it is error-prone when writing the equations for complicated systems.

The second approach (Sec. III) works directly at the many-particle level by focusing on many-particle propensities, leaving the counting/combinatorial details as a secondary task, albeit still a cumbersome one. The global many-particle propensities are derived as explicit expressions (in terms of sums and products) of the local rate functions using permutations and Dirac δ-distributions, which provide a method to select the required particles. One of its main advantages is that, as it works directly with many-particle propensities, it is capable of incorporating crowding effects in a more straightforward manner than the other approaches.

In the third approach (Sec. IV), the operators arise in the form of expansions containing single-particle creation and annihilation operators, which encode the combinatorics of particle selections. This allows us to focus on formulating only the operators per reaction, yielding a fast method to write down the CDME in a compact way for any reaction system, which can be a big advantage from a practical point of view. The resulting equation can be employed to perform analytical calculations, for instance, one can directly apply Galerkin discretizations,[16,29] opening the door for ready-to-use software libraries for numerical implementation, and to apply methodologies from quantum field theory.[14,18] In addition, the actions of the operators a^+ and a^- have immediate interpretations within the stochastic Malliavin calculus,[19] which may open a new perspective on the stochastic description of reaction–diffusion systems. However, the compact version of the CDME can appear obscure for practitioners used to more classical formulations in terms of integrals. To mitigate this issue, we added a dictionary (see the Appendix) to translate the short-hand notation for expansions in terms of creation and annihilation operators to concrete algebraic expressions, which explicitly include the combinatorial factors, sums, and integrals. This could be further automatized using a symbolic algebra software. We finally explore a special case using δ-distributions (Sec. IV D), which simplify the original expansions into simple integrals. Although the ease to derive discretizations—as well as some mathematical rigor—is lost, some practitioners might find this approach more suitable.

From a mathematical perspective, the CDME is formulated in terms of density functions. Another question of interest for future research is how to formulate a corresponding equation in terms of probability measures as in Ref. 30. This is of relevance since such a formulation might be more familiar to some mathematicians working on tangential fields, where one requires analogous models to reaction and diffusion, such as social dynamics.[31–33]

One of the main future prospective applications of the CDME is to unify most of the well-known reaction–diffusion models at different scales, establishing the relationships between them and yielding a theoretical and computational framework for multiscale modeling of biochemical reaction systems. For instance, we believe that the well-known models of diffusion-influenced reactions,[14,15,34–38] as well as recent developments,[39–41] can be recovered as special cases of the CDME. One can also, in principle, recover the chemical master equation as the well-mixed limit of the CDME. In this limiting case, it has been shown that the classical law of mass action emerges from the large copy number limit.[42] If the spatial dependence is kept, it has been shown that macroscopic reaction–diffusion models emerge as the large copy number limit of certain particle-based models,[9,43,44] which can be considered as spatial discretizations of the CDME. This further yields a precise relation between the macroscopic parameters and those at the particle level, allowing for consistent multiscale simulations.[9,45] In ongoing work, we are aiming at a general derivation of reaction–diffusion PDEs from the corresponding CDME without using an intermediate discrete model. To this end, it is essential to consider the correct combinatorial factors in the loss and gain terms. Another example is given by a recent simulation scheme to couple Markov models of molecular kinetics with particle-based reaction–diffusion simulations,[46,47] where the root model used to derive the schemes is once again a special case of the CDME. Similarly, in Ref. 48, the authors used a hierarchy of Fokker–Planck equations to model the variable number of ions in an ion channel; a model that we also believe is a special case of the CDME.

All in all, the CDME has the potential to unify a diverse range of reaction–diffusion models at different scales, yielding mathematical relationships that serve as the key ingredient to derive novel hybrid multiscale simulations for biochemical dynamics that capture the cascades of interactions across scales.

ACKNOWLEDGMENTS

We acknowledge the support from Deutsche Forschungsgemeinschaft (DFG) through the Collaborative Research Center (Grant No. SFB 1114) "Scaling Cascades in Complex Systems," Project No. 235221301, Sub-Projects C01 and C03, and under Germany's Excellence Strategy—MATH+: The Berlin Mathematics Research Center (Grant No. EXC-2046/1)—Project No. 390685689 (Subproject AA1-1). M.J.d.R. acknowledges the support from DFG (Grant No. RA 3601/1-1) and from the Dutch Institute for Emergent Phenomena (DIEP) cluster at the University of Amsterdam.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Mauricio J. del Razo: Conceptualization (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Stefanie Winkelmann**: Conceptualization (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Rupert Klein**: Conceptualization (equal); Funding acquisition (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Felix Höfling**: Funding acquisition (equal); Investigation (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

APPENDIX: EXPANSION DICTIONARY

Although using the notation presented in Sec. IV results in writing the CDME at once, it is not evident to find the connection to the more classical form of the equation. In this appendix, we present a dictionary for the most used cases, where we match the expansions in terms of creation and annihilation operators in the compact notation with their corresponding expressions in an explicit integral form. These expressions, although non-trivial, are straightforward to prove along the lines given in Ref. 16. At first, we present the expansions for the diffusion and then for loss operators, where the form is simpler as compared to the gain terms because it only depends on the reactants. Finally, we proceed with the gain operators. For the purpose of generality, we use the notation $\rho_{...}$, with the dots in the subindex indicating the unknown species involved in the reaction, e.g., we write $\rho_{a,b,...}(x^{(a)}, x^{(b)}, \dots)$, where the dots represent numbers and positions of other species, respectively.

1. Diffusion operators

In the absence of physical interactions, the diffusion operators only act on one particle at a time, so they are the most simple ones,

$$\mathrm{a}^{+}\ \mathrm{D}\ \mathrm{a}^{-}\ \rho_{n,\dots} = \sum_{\nu=1}^{n} D_\nu \rho_{n,\dots}. \tag{A1}$$

2. Loss operators

For the loss operators, only the reactants are relevant, while the products just determine the variables of integration. Thus, we denote the positions of the l products by $y^{(l)}$, regardless of their species.

(i) Reactions of the form $\emptyset \to$ *(l products).* The reaction rate function is given by $\lambda(y^{(l)};)$. Here, we leave the semicolon inside the rate function in order to emphasize that there are no reactants. The expression is simply given by

$$1\ \mathrm{L}\ 1\ \rho_{\dots} = \rho_{\dots} \int_{\mathbb{X}^l} \lambda(y^{(l)};)\, dy^{(l)}. \tag{A2}$$

(ii) Reactions of the form $A \to$ (l products). The reaction rate function is given by $\lambda(y^{(l)};\, x)$. Let the number of A-particles be n, and denote by $x_1^{(n)}, \ldots, x_n^{(n)}$ the positions of the n possible reactants. Then,

$$\left(\mathrm{a}^+ \;\mathrm{L}\; \mathrm{a}^- \;\rho_{n,\ldots}\right)(x^{(n)}, \ldots) = \rho_{n,\ldots}(x^{(n)}, \ldots) \sum_{\nu=1}^{n} \int_{\mathbb{X}^l} \lambda(y^{(l)}; x_\nu^{(n)})\, dy^{(l)}. \tag{A3}$$

(iii) Reactions of the form $A + A \to$ (l products). The reaction rate function is given by $\lambda(y^{(l)};\, x_1, x_2)$, where x_1 and x_2 are the positions of the reactants. Then,

$$\left((\mathrm{a}^+)^2 \;\mathrm{L}\; (\mathrm{a}^-)^2 \;\rho_{n,\ldots}\right)(x^{(n)}, \ldots) = \rho_{n,\ldots}(x^{(n)}, \ldots) \sum_{1\leqslant \nu_1 < \nu_2 \leqslant n} \int_{\mathbb{X}^l} \lambda(y^{(l)}; x_{\nu_1}^{(n)}, x_{\nu_2}^{(n)})\, dy^{(l)}, \tag{A4}$$

where again $x_1^{(n)}, \ldots, x_n^{(n)}$ denote the positions of the n possible reactants.

(iv) Reactions of the form $A + B \to$ (l products). The reaction rate function is given by $\lambda(y^{(l)};\, x, z)$, where x is the position of the A reactant and z is the position of the B reactant. Let a and b be the numbers of A and B particles and $x_1^{(a)}, \ldots, x_a^{(a)}$ and $x_1^{(b)}, \ldots, x_b^{(b)}$ be their positions, respectively. Then,

$$\left(\mathrm{a}^+\mathrm{b}^+ \;\mathrm{L}\; \mathrm{a}^-\mathrm{b}^- \;\rho_{a,b,\ldots}\right)(x^{(a)}, x^{(b)}, \ldots) = \rho_{a,b,\ldots}(x^{(a)}, x^{(b)}, \ldots) \sum_{\nu_1=1}^{a} \sum_{\nu_2=1}^{b} \int_{\mathbb{X}^l} \lambda(y^{(l)}; x_{\nu_1}^{(a)}, x_{\nu_2}^{(b)})\, dy^{(l)}. \tag{A5}$$

(v) Reactions of the form $k_1 A + k_2 B \to$ (l products). As a generalization of all the previous examples, we can write the loss for an arbitrary reaction involving two species in their reactants. The reaction rate function is given by $\lambda(y^{(l)}; x^{(k_1)}, z^{(k_2)})$, where $x^{(k_1)}$ are the positions of the A-reactants and $z^{(k_2)}$ are the positions of the B-reactants; a and b are the numbers of A and B particles, respectively. Then,

$$\left((\mathrm{a}^+)^{k_1}(\mathrm{b}^+)^{k_2} \;\mathrm{L}\; (\mathrm{a}^-)^{k_1}(\mathrm{b}^-)^{k_2} \;\rho_{a,b,\ldots}\right)(x^{(a)}, x^{(b)}, \ldots) = \rho_{a,b,\ldots}(x^{(a)}, x^{(b)}, \ldots) \sum_{\substack{1\leqslant \nu_1 < \cdots < \nu_{k_1} \leqslant a \\ 1\leqslant \mu_1 < \cdots < \mu_{k_2} \leqslant b}} \int_{\mathbb{X}^l} \lambda(y^{(l)}; x_{\nu_1,\ldots,\nu_{k_1}}^{(a)}, x_{\mu_1,\ldots,\mu_{k_2}}^{(b)})\, dy^{(l)}, \tag{A6}$$

where $x_{\nu_1,\ldots,\nu_{k_1}}^{(n)} := (x_{\nu_1}^{(n)}, \ldots, x_{\nu_{k_1}}^{(n)})$.

3. Gain operators

For the gain operators, both the reactants and the products are relevant, so we need to take both into account. Once again, as the number of species will, in general, not be known, we indicate particle numbers and position arguments referring to non-participating species by an ellipsis, ...

(i) Reactions of the form $k_1 A + k_2 B \to l_1 A + l_2 B$. The reaction rate function is given by $\lambda(y_A^{(l_1)}, y_B^{(l_2)}; x_A^{(k_1)}, x_B^{(k_2)})$; a and b are the numbers of A and B particles, respectively. The expression for the gain is then given by

$$\begin{aligned}
&\left((\mathrm{a}^+)^{l_1}(\mathrm{b}^+)^{l_2} \;\mathrm{G}\; (\mathrm{a}^-)^{k_1}(\mathrm{b}^-)^{k_2} \;\rho_{a+k_1-l_1, b+k_2-l_2,\ldots}\right)(x^{(a)}, x^{(b)}, \ldots) \\
&\quad = C_{ab} \sum_{\substack{1\leqslant \mu_1 < \cdots < \mu_{l_1} \leqslant a \\ 1\leqslant \eta_1 < \cdots < \eta_{l_2} \leqslant b}} \int_{\mathbb{X}^{k_1}\times\mathbb{X}^{k_2}} \rho_{a+k_1-l_1, b+k_2-l_2,\ldots}\left((x_{\setminus\{\mu_1,\ldots,\mu_{l_1}\}}^{(a)}, z^{(k_1)}), (x_{\setminus\{\eta_1,\ldots,\eta_{l_2}\}}^{(b)}, \hat{z}^{(k_2)}), \ldots\right) \\
&\quad\quad \times \lambda\left(x_{\mu_1,\ldots,\mu_{l_1}}^{(a)}, x_{\eta_1,\ldots,\eta_{l_2}}^{(b)}; z^{(k_1)}, \hat{z}^{(k_2)}\right) dz^{(k_1)}\, d\hat{z}^{(k_2)}
\end{aligned} \tag{A7}$$

with the combinatorial factor

$$C_{ab} = \binom{a}{l_1}^{-1} \binom{b}{l_2}^{-1} \binom{a + k_1 - l_1}{k_1} \binom{b + k_2 - l_2}{k_2}. \tag{A8}$$

(ii) Reactions of the form $k_1 A + k_2 B + C \to l_1 A + l_2 B$. The reaction rate function is given by $\lambda(y_A^{(l_1)}, y_B^{(l_2)}; x_A^{(k_1)}, x_B^{(k_2)}, x_C)$; a, b, and c are the numbers of A, B, and C particles, respectively. The expression for the gain is then given by

$$
\begin{aligned}
&\left((\mathrm{a}^+)^{l_1}(\mathrm{b}^+)^{l_2}\ \mathrm{G}\ (\mathrm{a}^-)^{k_1}(\mathrm{b}^-)^{k_2}\mathrm{c}^-\ \rho_{a+k_1-l_1,b+k_2-l_2,c+1,\dots}\right)(x^{(a)},x^{(b)},x^{(c)},\dots)\\
&\quad= C_{abc}\sum_{\substack{1\leqslant\mu_1<\dots<\mu_{l_1}\leqslant a\\ 1\leqslant\eta_1<\dots<\eta_{l_2}\leqslant b}}\int_{\mathbb{X}^{k_1}\times\mathbb{X}^{k_2}\times\mathbb{X}}\rho_{a+k_1-l_1,b+k_2-l_2,c+1,\dots}\left((x^{(a)}_{\backslash\{\mu_1,\dots,\mu_{l_1}\}},z^{(k_1)}),(x^{(b)}_{\backslash\{\eta_1,\dots,\eta_{l_2}\}},\hat z^{(k_2)}),(x^{(c)},z'),\dots\right)\\
&\qquad\times\lambda\left(x^{(a)}_{\mu_1,\dots,\mu_{l_1}},x^{(b)}_{\eta_1,\dots,\eta_{l_2}};z^{(k_1)},\hat z^{(k_2)},z'\right)dz^{(k_1)}d\hat z^{(k_2)}dz'
\end{aligned}
\tag{A9}
$$

with

$$
C_{abc}=\binom{a}{l_1}^{-1}\binom{b}{l_2}^{-1}\binom{a+k_1-l_1}{k_1}\binom{b+k_2-l_2}{k_2}\binom{c+1}{1}.
\tag{A10}
$$

(iii) Reactions of the form $k_1A+k_2B\to l_1A+l_2B+C$. The reaction rate function is given by $\lambda(y_A^{(l_1)},y_B^{(l_2)},y_C;x_A^{(k_1)},x_B^{(k_2)})$; a, b, and c are the numbers of A, B, and C particles, respectively. The expression for the gain is then given by

$$
\begin{aligned}
&\left((\mathrm{a}^+)^{l_1}(\mathrm{b}^+)^{l_2}\mathrm{c}^+\ \mathrm{G}\ (\mathrm{a}^-)^{k_1}(\mathrm{b}^-)^{k_2}\rho_{a+k_1-l_1,b+k_2-l_2,c-1,\dots}\right)(x^{(a)},x^{(b)},x^{(c)},\dots)\\
&\quad=\tilde C_{abc}\sum_{\substack{1\leqslant\mu_1<\dots<\mu_{l_1}\leqslant a\\ 1\leqslant\eta_1<\dots<\eta_{l_2}\leqslant b}}\sum_{\xi=1}^{c}\int_{\mathbb{X}^{k_1}\times\mathbb{X}^{k_2}}\rho_{a+k_1-l_1,b+k_2-l_2,c-1,\dots}\left((x^{(a)}_{\backslash\{\mu_1,\dots,\mu_{l_1}\}},z^{(k_1)}),(x^{(b)}_{\backslash\{\eta_1,\dots,\eta_{l_2}\}},\hat z^{(k_2)}),x^{(c)}_{\backslash\{\xi\}},\dots\right)\\
&\qquad\times\lambda\left(x^{(a)}_{\mu_1,\dots,\mu_{l_1}},x^{(b)}_{\eta_1,\dots,\eta_{l_2}},x^{(c)}_{\xi};z^{(k_1)},\hat z^{(k_2)}\right)dz^{(k_1)}d\hat z^{(k_2)}
\end{aligned}
\tag{A11}
$$

with

$$
\tilde C_{abc}=\frac{1}{c}\binom{a}{l_1}^{-1}\binom{b}{l_2}^{-1}\binom{a+k_1-l_1}{k_1}\binom{b+k_2-l_2}{k_2}.
\tag{A12}
$$

(iv) Reactions of the form $C\to A+B$. This is a special case of example (ii), putting $k_1=k_2=0$ and $l_1=l_2=1$. The reaction rate function is given by $\lambda(y_A,y_B;\ x_C)$; a, b, and c are the numbers of A, B, and C particles, respectively. The expression for the gain is then given by

$$
\begin{aligned}
&\left(\mathrm{a}^+\mathrm{b}^+\ \mathrm{G}\ \mathrm{c}^-\rho_{a-1,b-1,c+1,\dots}\right)(x^{(a)},x^{(b)},x^{(c)},\dots)\\
&\quad=\frac{c+1}{ab}\sum_{\mu=1}^{a}\sum_{\eta=1}^{b}\int_{\mathbb{X}}\rho_{a-1,b-1,c+1,\dots}\left(x^{(a)}_{\backslash\{\mu\}},x^{(b)}_{\backslash\{\eta\}},(x^{(c)},z),\dots\right)\lambda\left(x^{(a)}_{\mu},x^{(b)}_{\eta};z\right)dz.
\end{aligned}
\tag{A13}
$$

(v) Reactions of the form $A+B\to C$. This is a special case of example (iii) with $k_1=k_2=1$ and $l_1=l_2=0$. The reaction rate function is given by $\lambda(y_C;\ x_A,x_B)$; a, b, and c are the numbers of A, B, and C particles, respectively. The expression for the gain is then given by

$$
\begin{aligned}
&\left(\mathrm{c}^+\ \mathrm{G}\ \mathrm{a}^-\mathrm{b}^-\rho_{a+1,b+1,c-1,\dots}\right)(x^{(a)},x^{(b)},x^{(c)},\dots)\\
&\quad=\frac{(a+1)(b+1)}{c}\sum_{\xi=1}^{c}\int_{\mathbb{X}\times\mathbb{X}}\rho_{a+1,b+1,c-1,\dots}\left((x^{(a)},z),(x^{(b)},\hat z),x^{(c)}_{\backslash\{\xi\}}\dots\right)\lambda\left(x^{(c)}_{\xi};z,\hat z\right)dzd\hat z.
\end{aligned}
\tag{A14}
$$

REFERENCES

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

[2] B. Drawert, S. Engblom, and A. Hellander, "URDME: A modular framework for stochastic simulation of reaction-transport processes in complex geometries," BMC Syst. Biol. **6**, 76 (2012).

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

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

[5] S. Smith and R. Grima, "Spatial stochastic intracellular kinetics: A review of modelling approaches," Bull. Math. Biol. **81**, 2960–3009 (2019).

[6] P. Grindrod, *Patterns and Waves: The Theory and Applications of Reaction-Diffusion Equations* (Oxford University Press, 1991).

[7] J. D. Murray, *Mathematical Biology II: Spatial Models and Biomedical Applications* (Springer, New York, 2001), Vol. 3.

[8] F. Brauns, J. Halatek, and E. Frey, "Phase-space geometry of mass-conserving reaction-diffusion dynamics," Phys. Rev. X **10**, 041036 (2020).

[9] M. Kostré, C. Schütte, F. Noé, and M. J. del Razo, "Coupling particle-based reaction-diffusion simulations with reservoirs mediated by reaction-diffusion PDEs," Multiscale Model. Simul. **19**, 1659–1683 (2021).

[10] C. Kim, A. Nonaka, J. B. Bell, A. L. Garcia, and A. Donev, "Stochastic simulation of reaction-diffusion systems: A fluctuating-hydrodynamics approach," J. Chem. Phys. **146**, 124110 (2017).

[11] S. A. Isaacson, J. Ma, and K. Spiliopoulos, "Mean field limits of particle-based stochastic reaction-diffusion models," SIAM J. Math. Anal. **54**, 453–511 (2022).

[12] F. Höfling and T. Franosch, "Anomalous transport in the crowded world of biological cells," Rep. Prog. Phys. **76**, 046602 (2013).

[13] M. Weiss, "Crowding, diffusion, and biochemical reactions," in *New Models of the Cell Nucleus: Crowding, Entropic Forces, Phase Separation, and Fractals*, International Review of Cell and Molecular Biology Vol. 307, edited by R. Hancock and K. W. Jeon (Academic Press, 2014), Chap. 11, pp. 383–417.
[14] M. Doi, "Stochastic theory of diffusion-controlled reaction," J. Phys. A: Math. Gen. **9**, 1479 (1976).
[15] M. von Smoluchowski, "Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Lösungen," Z. Phys. Chem. **92**(1), 129–168 (1918).
[16] M. J. del Razo, D. Frömberg, A. V. Straube, C. Schütte, F. Höfling, and S. Winkelmann, "A probabilistic framework for particle-based reaction-diffusion dynamics using classical Fock space representations," Lett. Math. Phys. **112**, 49 (2022).
[17] M. Doi, "Second quantization representation for classical many-particle system," J. Phys. A: Math. Gen. **9**, 1465 (1976).
[18] P. Grassberger and M. Scheunert, "Fock-space methods for identical classical objects," Fortschr. Phys. **28**, 547–578 (1980).
[19] A. Lanconelli, "Using Malliavin calculus to solve a chemical diffusion master equation," arXiv:2203.14676 [math.PR].
[20] D. T. Gillespie, "A rigorous derivation of the chemical master equation," Physica A **188**, 404–425 (1992).
[21] D. A. McQuarrie, "Stochastic approach to chemical kinetics," J. Appl. Probab. **4**, 413–478 (1967).
[22] H. Qian and L. M. Bishop, "The chemical master equation approach to nonequilibrium steady-state of open biochemical systems: Linear single-molecule enzyme kinetics and nonlinear biochemical reaction networks," Int. J. Mol. Sci. **11**, 3472–3500 (2010).
[23] S. Winkelmann and C. Schütte, *Stochastic Dynamics in Computational Biology* (Springer, 2020), Vol. 645.
[24] R. Klein and L. Delle Site, "Derivation of Liouville-like equations for the n-state probability density of an open system with thermalized particle reservoirs and its link to molecular simulation," J. Phys. A: Math. Theor. **55**, 155002 (2022).
[25] L. Delle Site and R. Klein, "Liouville-type equations for the n-particle distribution functions of an open system," J. Math. Phys. **61**, 083102 (2020).
[26] L. Schimansky-Geier, M. Mieth, H. Rosé, and H. Malchow, "Structure formation by active Brownian particles," Phys. Lett. A **207**, 140–146 (1995).
[27] F. Schweitzer and J. D. Farmer, *Brownian Agents and Active Particles: Collective Dynamics in the Natural and Social Sciences* (Springer, 2003), Vol. 1.
[28] M. E. Peskin, *An Introduction to Quantum Field Theory* (CRC Press, 2018).
[29] D. Frömberg and F. Höfling, "Generalized master equation for first-passage problems in partitioned spaces," J. Phys. A: Math. Theor. **54**, 215601 (2021).
[30] V. P. Belavkin and V. N. Kolokol'tsov, "On a general kinetic equation for many–particle systems with interaction, fragmentation and coagulation," Proc. R. Soc. London, Ser. A **459**, 727–748 (2003).
[31] N. D. Conrad, D. Furstenau, A. Grabundžija, L. Helfmann, M. Park, W. Schier, B. Schütt, C. Schütte, M. Weber, N. Wulkow *et al.*, "Mathematical modeling of the spreading of innovations in the ancient world," eTopoi **7**, 1 (2018).
[32] N. Djurdjevac Conrad, L. Helfmann, J. Zonker, S. Winkelmann, and C. Schütte, "Human mobility and innovation spreading in ancient times: A stochastic agent-based simulation approach," EPJ Data Sci. **7**, 24 (2018).
[33] T. Veloz, P. Razeto-Barry, P. Dittrich, and A. Fajardo, "Reaction networks and evolutionary game theory," J. Math. Biol. **68**, 181–206 (2014).
[34] N. Agmon and A. Szabo, "Theory of reversible diffusion-influenced reactions," J. Chem. Phys. **92**, 5270–5284 (1990).
[35] F. C. Collins and G. E. Kimball, "Diffusion-controlled reaction rates," J. Colloid Sci. **4**, 425–437 (1949).
[36] P. Hänggi, P. Talkner, and M. Borkovec, "Reaction-rate theory: Fifty years after Kramers," Rev. Mod. Phys. **62**, 251 (1990).
[37] A. Szabo, K. Schulten, and Z. Schulten, "First passage time approach to diffusion controlled reactions," J. Chem. Phys. **72**, 4350–4357 (1980).
[38] A. Szabo, "Theory of diffusion-influenced fluorescence quenching," J. Phys. Chem. **93**, 6929–6939 (1989).
[39] I. V. Gopich and A. Szabo, "Theory of diffusion-influenced reaction networks," J. Phys. Chem. B **122**, 11338–11354 (2018).
[40] I. V. Gopich and A. Szabo, "Diffusion-induced competitive two-site binding," J. Chem. Phys. **150**, 094104 (2019).
[41] M. Dibak, C. Fröhner, F. Noé, and F. Höfling, "Diffusion-influenced reaction rates in the presence of pair interactions," J. Chem. Phys. **151**, 164105 (2019).
[42] D. F. Anderson and T. G. Kurtz, *Stochastic Analysis of Biochemical Systems* (Springer, 2015), Vol. 674.
[43] M. J. del Razo and H. Qian, "A discrete stochastic formulation for reversible bimolecular reactions via diffusion encounter," Commun. Math. Sci. **14**, 1741–1772 (2016).
[44] M. J. del Razo, H. Qian, and F. Noé, "Grand canonical diffusion-influenced reactions: A stochastic theory with applications to multiscale reaction–diffusion simulations," J. Chem. Phys. **149**, 044102 (2018).
[45] M. B. Flegg, S. J. Chapman, and R. Erban, "The two-regime method for optimizing stochastic reaction–diffusion simulations," J. R. Soc. Interface **9**, 859–868 (2012).
[46] M. Dibak, M. J. del Razo, D. De Sancho, C. Schütte, and F. Noé, "MSM/RD: Coupling Markov state models of molecular kinetics with reaction-diffusion simulations," J. Chem. Phys. **148**, 214107 (2018).
[47] M. J. del Razo, M. Dibak, C. Schütte, and F. Noé, "Multiscale molecular kinetics by coupling Markov state models and reaction-diffusion dynamics," J. Chem. Phys. **155**, 124109 (2021).
[48] W. Chen, R. Erban, and S. J. Chapman, "From Brownian dynamics to Markov chain: An ion channel example," SIAM J. Appl. Math. **74**, 208–235 (2014).