The Journal of Chemical Physics

Remarks on the chemical Fokker-Planck and Langevin equations: Nonphysical currents at equilibrium

Cite as: J. Chem. Phys. **148**, 064114 (2018); https://doi.org/10.1063/1.5016158
Submitted: 16 November 2017 . Accepted: 24 January 2018 . Published Online: 14 February 2018

Alessandro Ceccato, and Diego Frezzato

View Online Export Citation CrossMark

ARTICLES YOU MAY BE INTERESTED IN

The chemical Langevin equation
The Journal of Chemical Physics **113**, 297 (2000); https://doi.org/10.1063/1.481811

How accurate are the nonlinear chemical Fokker-Planck and chemical Langevin equations?
The Journal of Chemical Physics **135**, 084103 (2011); https://doi.org/10.1063/1.3625958

The complex chemical Langevin equation
The Journal of Chemical Physics **141**, 024103 (2014); https://doi.org/10.1063/1.4885345

AIP Publishing

© 2018 Author(s).

Remarks on the chemical Fokker-Planck and Langevin equations: Nonphysical currents at equilibrium

Alessandro Ceccato and Diego Frezzato[a)]
Department of Chemical Sciences, University of Padova, Via Marzolo 1, I-35131 Padova, Italy

(Received 16 November 2017; accepted 24 January 2018; published online 14 February 2018)

The chemical Langevin equation and the associated chemical Fokker-Planck equation are well-known continuous approximations of the discrete stochastic evolution of reaction networks. In this work, we show that these approximations suffer from a physical inconsistency, namely, the presence of nonphysical probability currents at the thermal equilibrium even for closed and fully detailed-balanced kinetic schemes. An illustration is given for a model case. *Published by AIP Publishing.* https://doi.org/10.1063/1.5016158

I. INTRODUCTION

Under isothermal conditions and rapid re-distribution of molecules in the available space of fixed volume, chemical reactions involving small numbers of molecules in homogeneous fluid phases are consensually modeled as a Markov process, in which the system's state is specified by the number of molecules of each species. In such a framework, the chemical master equation (CME)[1–3] and Gillespie's stochastic simulation algorithm (SSA)[4] provide the exact description of the evolution in terms of probabilistic expectations and generation of trajectories, respectively. Unfortunately, the CME is analytically hardly tractable apart from simple cases, and even its numerical solution becomes rapidly unfeasible as the number of reactant molecules increases. In parallel, simulations via SSA become lengthy again in the limit of large numbers of molecules and/or in the presence of a large time scale separation between the reaction channels (stiffness); moreover, a very large ensemble of trajectories should be simulated to achieve accurate statistics. To circumvent these issues, one looks for approximate but reliable simplifications of the exact evolution law.

One popular approximate evolution machinery of stochastic reaction networks is the so-called chemical Langevin equation (CLE), introduced by Gillespie in Ref. 5 and compared with the previous approaches of van Kampen[6] and Kurtz.[7–9] In the CLE context, the evolution of the system is described in a coarse-grained fashion on the time variable. This implies turning from integer numbers of molecules (in the following denoted by n_j for the j th species) to their continuous real-valued extension (the η_j in the following). The CLE, which will be reviewed and commented in Sec. II, has the form of a Langevin-like Itô stochastic differential equation for the evolution of the configuration $\boldsymbol{\eta}$ [see Eq. (9)]. The chemical Fokker Planck equation (CFPE) is the corresponding partial-derivative differential equation which rules the evolution of the probability density in the $\boldsymbol{\eta}$-space starting from a given initial condition. In short, the approximate CLE replaces the exact SSA route, whereas the approximate CFPE replaces the exact CME.

The strength of the CLE consists of dealing with continuous dynamical variables and allowing for rapid simulation of single trajectories. In this respect, the CLE is greatly employed in biochemical contexts, for example as in transcriptional regulation,[10] provided that one can switch from the exact SSA to the CLE coarse-grained picture.[11] In addition, the CLE is the suitable intermediate step to bridge stochastic kinetics and macroscopic mass-action rate equations; such a link can be established in the thermodynamic limit in which both the numbers of reactant molecules and the volume increase at fixed volumetric concentrations.[12] We also mention some recent advancements in the adaptation of the computational singular perturbation methodology to achieve dimensional reduction for prototype models of stochastic differential equations;[13] further development of that strategy, with application to the CLE, could lead to set up the machinery for disentangling slow and fast modes of evolution for stochastic chemical networks at the mesoscale between low numbers of molecules and the thermodynamic limit. In parallel, the potential utility of the CFPE consists in the possibility, at least in principle, of detecting directly the slow eigenmodes of evolution (and related rates) of a reaction network. For example, modern strategies like Diffusion Maps[14] might be suitable for constructing the slowest evolution modes even in relatively high-dimensional reaction networks.

Besides these points of strength, the crucial question is: How safely can we rely on the CLE and CFPE as *physically consistent* shortcuts of the SSA and CME? It is not only a matter of having good approximations of the exact solution on quantitative grounds but also, and more importantly, to check if the CLE and CFPE are at least devoid of nonphysical drawbacks (or, if present, to what extent they may be serious). Regarding the numerical consistency, it is known that the statistical properties of the ensemble of trajectories simulated by means of the CLE are fully consistent with the CME/SSA only for networks of unimolecular reactions; otherwise, the consistency is guaranteed only up to the first- and second-order moments of the distribution.[15]

[a)]Author to whom correspondence should be addressed: diego.frezzato@unipd.it

The accuracy of the multivariate CLE and CFPE has been addressed, for instance, in Ref. 16. Previous inspections on model reaction networks reducible to one-dimensional systems featuring bi-stability[17–19] revealed that the CFPE fails in reproducing the probability density on the long time scale. About the physical consistency, Horowitz recently showed[20] that the CLE is consistent with the thermodynamics (in the sense that the rate of entropy production along the trajectories matches the heat flux between the system and the thermal bath) only when the system is close to equilibrium. The global picture is that these continuous approximations of CME/SSA suffer from subtle inconsistencies whose quantitative manifestation cannot be easily assessed without a case-by-case analysis.

In this paper, we focus on a physical inconsistency that emerges from the analysis of the CFPE. Namely, we shall see that nonphysical probability currents may be generally present at equilibrium even for closed and detailed-balanced reaction networks. In practice, this means that a stationary distribution is attained in the long time scale, but a directed circulation (on average) in the configurational space would still be present. This clearly goes against the condition of thermal equilibrium. It must be stressed from the beginning that such an issue regards both the CLE and CFPE. In fact, the CLE and the CFPE are fully consistent with each other in the sense that, at the given boundary conditions in the $\boldsymbol{\eta}$-space, they have the same statistics.

The paper is structured as follows. In Sec. II we specify the physical context and give the essentials about the chemical master equation approach. Section III presents the chemical Langevin equation and the associated Fokker-Planck equation, with special emphasis on their limits of applicability. In Sec. IV we address the nonphysical probability currents that emerge in the chemical Fokker-Planck context, and in Sec. V we illustrate such an issue for a model kinetic scheme. Section VI is devoted to conclusions.

Mathematical notation. Throughout the paper, vectors and matrices will be indicated with bold style. Vectors are implicitly intended as column-vectors. The superscript "T" denotes the transposed array. The symbol "⊗" stands for the dyadic product between two vectors: $\mathbf{a}\otimes\mathbf{b}$ is the matrix with elements $[\mathbf{a}\otimes\mathbf{b}]_{ij} = a_i b_j$.

II. PHYSICAL CONTEXT AND THE CHEMICAL MASTER EQUATION

Let us consider a network of M elementary reactions (labeled by the index m) involving N chemical species (labeled by the index j). Let $\nu_{R_j}^{(m)}$ and $\nu_{P_j}^{(m)}$ be the stoichiometric coefficients of the species j as reactant and product, respectively, in the reaction m. The system's configuration is specified by the array $\mathbf{n}$ whose non-negative integer entries $(n_1, n_2, \ldots, n_N)$ are the numbers of molecules of each species. Finally, the set $\mathbf{n}_0$ specifies the initial configuration.

The dynamics corresponds to stochastic transitions among all possible configurations which are accessible from $\mathbf{n}_0$ due to the moves allowed by the stoichiometry of the reaction channels. In this work, we consider closed networks (neither pure source nor sink processes) composed of reversible reactions. This implies that the number of molecules remains strictly positive for each species, and that the number of achievable configurations is finite. In addition, a number of *a priori* constraints implies that some linear combinations of the molecular numbers are conserved: that is, there exists a constant matrix $\mathbf{S}$ of dimension $d\times N$ with $d < N$ such that

$$\mathbf{S}\,\mathbf{n}(t) \equiv \mathbf{S}\,\mathbf{n}_0 = \mathbf{c} \tag{1}$$

at any time, where $\mathbf{c}$ is a constant vector. By introducing the N-dimensional arrays $\boldsymbol{\nu}_m$ with entries

$$(\boldsymbol{\nu}_m)_j = \nu_{P_j}^{(m)} - \nu_{R_j}^{(m)}, \tag{2}$$

the condition in Eq. (1) corresponds to

$$\mathbf{S}\,\boldsymbol{\nu}_m = \mathbf{0} \quad \text{for each } m. \tag{3}$$

The full array $\mathbf{n}$ is thus redundant since the accessible configurations lie on a $(N-d)$-dimensional hyperplane in the full space. A subset $\tilde{\mathbf{n}}$ of dimension $(N-d)$ suffices to specify the network's state.

In the context defined above, the quantity of interest is the probability $p(\mathbf{n}, t)$ to find the network in the configuration $\mathbf{n}$ at time t; $p(\mathbf{n}, t)$ is normalized as $\sum_{\mathbf{n}} p(\mathbf{n}, t) = 1$ at any time. The initial condition is $p(\mathbf{n}, 0) = \prod_j \delta_{n_j, n_j^0}$, where n_j^0 are the components of $\mathbf{n}_0$ and δ stands for the Kronecker's delta-function.

The evolution of $p(\mathbf{n}, t)$ is specified by the chemical master equation (CME) given below. The CME is built by accounting for both the processes that lead to the realization of the state $\mathbf{n}$ from other states and the processes that take off from it,[1–3]

$$\frac{\partial p(\mathbf{n},t)}{\partial t} = \sum_{m=1}^{M} \left[a_m(\mathbf{n}-\boldsymbol{\nu}_m)p(\mathbf{n}-\boldsymbol{\nu}_m,t) - a_m(\mathbf{n})p(\mathbf{n},t)\right]. \tag{4}$$

The state-dependent factors $a_m(\mathbf{n})$ are the so-called "propensity functions"; the quantity $a_m(\mathbf{n})\delta t$ is the probability that, if the system is presently in the state $\mathbf{n}$, the mth reaction takes place in the subsequent time-interval δt. The general form of a propensity function is $a_m(\mathbf{n}) = c_m f_m(\mathbf{n})$, where the function $f_m(\mathbf{n})$ and the proportionality coefficient c_m with the physical dimension of inverse-of-time are deduced from the molecularity of the elementary reaction on the basis of combinatorial arguments, and from the matching with the deterministic mass-action rate equation when the numbers of reactant molecules are large. In particular, only first- and second-order reactions are of practical relevance. For unimolecular reactions A → Products, the propensity function reads $a_{\text{uni}}(\mathbf{n}) = c_{\text{uni}}\, n_{\text{A}}$, where $c_{\text{uni}} \equiv k_{\text{uni}}$ is the kinetic constant in the deterministic limit. For bimolecular reactions of homo-molecular kind, 2A → Products, one has $a_{\text{bim},1}(\mathbf{n}) = c_{\text{bim},1} n_{\text{A}}(n_A - 1)/2$ with $c_{\text{bim},1} = 2k_{\text{bim},1}V^{-1}$, while for bimolecular reactions of hetero-molecular kind, A + B → Products, one has $a_{\text{bim},2}(\mathbf{n}) = c_{\text{bim},2} n_{\text{A}} n_{\text{B}}$ with $c_{\text{bim},2} = k_{\text{bim},2}V^{-1}$ ($k_{\text{bim},1}$ and $k_{\text{bim},2}$ are the kinetic constants in the deterministic limit and V is the available volume).

The single-trajectory counterpart of the CME is Gillespie's stochastic simulation algorithm (SSA)[4] which generates trajectories whose statistical ensemble is exactly consistent with the CME.

Apart from simple cases, solving the CME is a quite hard task. The most natural way is to convert Eq. (4) into a set of linear ordinary differential equations[21] and solve them by means of strategies that are able to contrast the rapid growth of dimension as the number of accessible configurations increases; among these strategies, we mention the "finite state projection" method[22,23] and its technical variants.[24] Concerning the SSA counterpart, the problem is that the advancement of the reaction network becomes slow when the number of reactant molecules is large and/or in the presence of a large spread in the magnitude of the c_m rate coefficients (stiffness). Because of these criticalities, efficient approximations of the SSA/CME are demanded when treating stiff reaction networks and/or large numbers of molecules (but not large enough to adopt the deterministic rate equations).

III. THE CHEMICAL LANGEVIN AND FOKKER-PLANCK EQUATIONS

In this section we introduce the chemical Langevin equation (CLE) and the associated chemical Fokker-Planck equation (CFPE). It is well known that the CFPE can be derived directly by truncating the Kramers-Moyal expansion of the CME (written in terms of variables $\boldsymbol{\eta}$) at the second-order derivatives; see, for example, Sec. 7.5 of Gardiner's book[1] and Ref. 16 on the same topic. On the other hand, as indicated by Gillespie,[5] *that* form of CFPE is exactly the Fokker-Planck equation whose drift and diffusion terms are parametrized by the CLE. In such a way, the CFPE is supported by the clear physical assumptions that underlie the CLE (see below), rather than deriving from a mere mathematical truncation of the Kramers-Moyal. This is the perspective adopted here.

Before proceeding further, we make a clarification concerning the connection between integer numbers of molecules and their continuous extension to real values. Let $\mathcal{I}$ be the domain of configurations $\mathbf{n}$ which are accessible from the initial condition. Then, let $\mathcal{D}$ be the domain in $\mathbb{R}^N$ which "fills" and "completes" $\mathcal{I}$ in the following sense: by denoting with cell($\mathbf{n}$) the hyper-cube $n_j - 1/2 \leq \eta_j < n_j + 1/2$, we say that $\boldsymbol{\eta} \in \mathcal{D}$ if there exists a unique $\mathbf{n} \in \mathcal{I}$ such that $\boldsymbol{\eta} \in$ cell($\mathbf{n}$). The domain $\mathcal{D}$ is the union of all cell($\mathbf{n}$) for $\mathbf{n} \in \mathcal{I}$.

A. The CLE

Following Gillespie, the CLE is derived directly from the physical assumptions underlying the CME and SSA. We mention an interesting alternative approach[15] in which the CLE emerges as one among several allowed parametric stochastic differential equations, all ensuring the matching of the first- and second-order moments of the molecular populations with those produced by the CME. However, in Gillespie's derivation, such a subjective freedom is absent and the CLE is *the* stochastic differential equation mimicking (it being an approximation) the true evolution of a reaction network.

In order to derive the CLE, two assumptions are required. The first one, termed the "tau-leap condition," consists in assuming that all propensity functions $a_m(\boldsymbol{\eta})$ do not change appreciably in a certain time interval Δt sufficiently short. This allows one to adopt the tau-leaping propagation formula in which several reactions can occur, even several times, in that interval. The number of events of each mth reaction is drawn from the Poisson distribution with mean $a_m(\boldsymbol{\eta})\Delta t$. The second assumption consists in having the possibility to choose Δt sufficiently long so that the first assumption still holds but $a_m(\boldsymbol{\eta})\Delta t \gg 1$ for all reactions. This allows one to approximate the Poisson distributions by Gaussian distributions with mean and variance both equal to $a_m(\boldsymbol{\eta})\Delta t$.

As a whole, the CLE is applicable if it is possible to choose Δt such that

$$\Delta t_{\min}(\boldsymbol{\eta}) \leq \Delta t \leq \Delta t_{\max}(\boldsymbol{\eta}), \tag{5}$$

where we take

$$\Delta t_{\min}(\boldsymbol{\eta}) = \frac{\gamma}{\min_m\{a_m(\boldsymbol{\eta})\}} \tag{6}$$

with $\gamma \gg 1$ subjectively chosen ($\gamma = 3$ is considered to be sufficient by us), and where $\Delta t_{\max}(\boldsymbol{\eta})$ is an estimate of the largest value of the propagation time-step which can be employed in the tau-leaping strategy (for example, one can adopt the efficient τ-selection procedure presented in Ref. 25, as in Sec. V). The applicability of the CLE is thus limited to the regions of the $\boldsymbol{\eta}$-space where

$$\frac{\Delta t_{\min}(\boldsymbol{\eta})}{\Delta t_{\max}(\boldsymbol{\eta})} \leq 1. \tag{7}$$

The condition in Eq. (7) is usually fulfilled for sufficiently large numbers of reactant molecules so that the rate of reactive events is large [$\Delta t_{\min}(\boldsymbol{\eta})$ is small], but even the occurrence of a large number of reactions does not sensibly affect the value of the propensity functions [hence $\Delta t_{\max}(\boldsymbol{\eta})$ can be longer than $\Delta t_{\min}(\boldsymbol{\eta})$].

Gillespie showed that if Δt can be fixed according to Eq. (5) for the current state $\boldsymbol{\eta}$, then the following propagation route is accurate:[5]

$$\boldsymbol{\eta}(t+\Delta t) \simeq \boldsymbol{\eta}(t) + \Delta t \sum_m \boldsymbol{\nu}_m a_m(\boldsymbol{\eta}(t)) + \sum_m \boldsymbol{\nu}_m \sqrt{a_m(\boldsymbol{\eta}(t))\Delta t}\, \mathcal{N}_m(0,1), \tag{8}$$

where $\mathcal{N}_m(0,1)$ are random numbers drawn from independent Standard Normal Distributions (zero mean and unit variance). Equation (8) is the CLE in the form of explicit advancement of the system's state. In the form of Itô stochastic differential equation following from Eq. (8), the CLE reads

$$\frac{d\boldsymbol{\eta}}{dt} \simeq \sum_m \boldsymbol{\nu}_m a_m(\boldsymbol{\eta}) + \sum_m \boldsymbol{\nu}_m \sqrt{a_m(\boldsymbol{\eta})}\, \xi_m, \tag{9}$$

where ξ_m stands for the mth component of the M-dimensional Gaussian white noise.[26] In adopting Eq. (9), caution must be taken since "dt" is a "macroscopic infinitesimal" (Gillespie's terminology[5]) bounded according to Eq. (5).

As stressed by Gillespie and co-workers,[5,25] the propagation via CLE should be halted, in favor of the exact SSA, as soon as the two requirements for the CLE validity are no more fulfilled. Moreover, the evolution scheme of Eq. (8) may give rise to a problem when $\boldsymbol{\eta}(t)$ is a point close to the faces of the positive orthant and the amplitude of the

propagation step is large enough to bring $\boldsymbol{\eta}(t + \Delta t)$ out of the orthant so that for one or more species the number of molecules would become negative. In our opinion, a formal way to incorporate a physical boundary into the CLE scheme is still lacking and *ad hoc* solutions have been proposed to date. An alternative is to accept the occurrence of negative concentrations [and hence of possible imaginary factors multiplying the white noise terms in Eq. (8)] and check that the statistical properties of the ensemble of trajectories are, however, compatible with the CME statistics.[27] On the other hand, when the number of molecules of a reactant species is close to zero, one falls outside the region of applicability of the CLE itself.

Note that Eqs. (8) and (9) fulfill the mass-conservation constraints discussed in Sec. II. In fact, by multiplying both members of these equations by the matrix $\mathbf{S}$ and considering Eq. (3), it follows $d[\mathbf{S}\boldsymbol{\eta}(t)]/dt = \mathbf{0}$ which implies $\mathbf{S}\boldsymbol{\eta}(t) = \mathbf{S}\boldsymbol{\eta}(0) = \mathbf{c}$. This means that a reduced $(N - d)$-dimensional array $\tilde{\boldsymbol{\eta}}(t)$ suffices to describe the system's state once the matrix $\mathbf{S}$ is known and the array $\mathbf{c}$ is given. By introducing a $(N - d) \times N$ matrix $\mathbf{R}$ which selects the independent variables via $\tilde{\boldsymbol{\eta}}(t) = \mathbf{R}\boldsymbol{\eta}(t)$, a reduced form of Eqs. (8) and (9) is readily obtained; for example, Eq. (9) turns into

$$\frac{d\tilde{\boldsymbol{\eta}}}{dt} \simeq \sum_m \mathbf{R}\boldsymbol{\nu}_m \tilde{a}_m(\tilde{\boldsymbol{\eta}}) + \sum_m \mathbf{R}\boldsymbol{\nu}_m \sqrt{\tilde{a}_m(\tilde{\boldsymbol{\eta}})}\, \xi_m, \tag{10}$$

where $\tilde{a}_m(\tilde{\boldsymbol{\eta}}) \equiv a_m(\boldsymbol{\eta})|_{\boldsymbol{\eta}=\boldsymbol{\eta}(\tilde{\boldsymbol{\eta}},\mathbf{c})}$ in which $\boldsymbol{\eta}(\tilde{\boldsymbol{\eta}}, \mathbf{c})$ denotes the full set of variables retrieved from the reduced one.

B. The CFPE

The CLE allows the parametrization of the corresponding chemical Fokker-Planck equation (CFPE) for the evolution of the probability density $\rho(\boldsymbol{\eta}, t)$ normalized as $\int d\boldsymbol{\eta}\ \rho(\boldsymbol{\eta}, t) = 1$. The link between the probability density $\rho(\boldsymbol{\eta}, t)$ and the probability $p(\mathbf{n}, t)$ can be set, on intuitive grounds, to be $\int_{\text{cell}(\mathbf{n})} d\boldsymbol{\eta}\ \rho(\boldsymbol{\eta}, t) = p(\mathbf{n}, t)$.

The CFPE takes the form

$$\frac{\partial \rho(\boldsymbol{\eta}, t)}{\partial t} = -\hat{\Gamma}\rho(\boldsymbol{\eta}, t) \tag{11}$$

with the evolution operator

$$\hat{\Gamma} = \frac{\partial}{\partial \boldsymbol{\eta}}^T \mathbf{v}(\boldsymbol{\eta}) - \frac{1}{2}\sum_{i,j} \frac{\partial^2}{\partial \eta_i \partial \eta_j} B_{ij}(\boldsymbol{\eta}), \tag{12}$$

in which the drift vector $\mathbf{v}(\boldsymbol{\eta}) = \lim_{\Delta t \to 0}\{\langle \Delta\boldsymbol{\eta}(\Delta t)\rangle/\Delta t\}$ and the diffusion matrix $\mathbf{B}(\boldsymbol{\eta}) = \lim_{\Delta t \to 0}\{\langle \Delta\boldsymbol{\eta}(\Delta t) \otimes \Delta\boldsymbol{\eta}(\Delta t)\rangle/\Delta t\}$ are determined by using Eq. (8) for the displacement $\Delta\boldsymbol{\eta}(\Delta t)$, and by considering the statistical properties of the distributions $\mathcal{N}_m(0, 1)$ to evaluate the averages.[28] The resulting expressions, according to Gillespie,[5] are

$$\mathbf{v}(\boldsymbol{\eta}) = \sum_m \boldsymbol{\nu}_m a_m(\boldsymbol{\eta}) \tag{13}$$

and

$$\mathbf{B}(\boldsymbol{\eta}) = \sum_m [\boldsymbol{\nu}_m \otimes \boldsymbol{\nu}_m] a_m(\boldsymbol{\eta}). \tag{14}$$

Such a matching ensures that the solution of Eq. (11) yields a probability density in accord with the one obtainable from the statistical analysis of the ensemble of trajectories generated by means of Eq. (8) (under the application of the same boundary conditions). We remark again that Eqs. (11)–(14) agree with the CFPE that can be obtained directly from the Kramers-Moyal expansion of the CME up to the second-order terms.[1]

It is readily seen that $\mathbf{B}(\boldsymbol{\eta})$ is a $N \times N$ (symmetric) positive semidefinite matrix since $\mathbf{u}^T\boldsymbol{B}(\boldsymbol{\eta})\mathbf{u} \geq 0$ for any vector $\mathbf{u}$ in the N-dimensional space. In fact, $\mathbf{u}^T\mathbf{B}(\boldsymbol{\eta})\mathbf{u} = \sum_m (\mathbf{u}^T\boldsymbol{\nu}_m)^2 a_m(\boldsymbol{\eta})$ is always non-negative and null only for vectors orthogonal to the hyperplane individuated by the mass-conservation constraints ($\mathbf{u}^T\boldsymbol{\nu}_m = 0$ for all m). This implies that the probability spread (diffusion) out of such hyperplane is automatically prohibited by the structure of the CFPE itself. On the other hand, for an easier handling of the CFPE, it may be preferred to get rid *a priori* of these extra dimensions by adopting the reduced CFPE for the essential variables $\tilde{\boldsymbol{\eta}}$. The reduced equation is analogous to Eq. (11) with (12), but with derivatives taken with respect to the components of $\tilde{\boldsymbol{\eta}}$, drift vector $\tilde{\mathbf{v}}(\tilde{\boldsymbol{\eta}}) = \sum_m \mathbf{R}\boldsymbol{\nu}_m \tilde{a}_m(\tilde{\boldsymbol{\eta}})$, and diffusion matrix $\tilde{\mathbf{B}}(\tilde{\boldsymbol{\eta}}) = \sum_m [(\mathbf{R}\boldsymbol{\nu}_m) \otimes (\mathbf{R}\boldsymbol{\nu}_m)]\tilde{a}_m(\tilde{\boldsymbol{\eta}})$. The diffusion matrix now results to be positive definite. The fact that the eigenvalues of $\tilde{\mathbf{B}}(\tilde{\boldsymbol{\eta}})$ are strictly real and positive for any $\tilde{\boldsymbol{\eta}}$ ensures that the stationary state is reached in the long time scale.

Although it does not emerge formally in the derivations of the CFPE, "impenetrable boundaries" should be applied at the faces of the positive orthant or, in the reduced formulation in terms of the variables $\tilde{\boldsymbol{\eta}}$, at the intersections between the hyperplane determined by the mass-conservation constraints and the faces of the positive orthant. The need for such reflecting boundaries, on which the orthogonal component of the probability current must vanish (see the discussion in the following), is related with the need to keep $\boldsymbol{\eta}(t)$ inside the positive orthant when the trajectories are simulated by means of the CLE. We stress that, contrary to other kinds of stochastic dynamics like conformational fluctuations in molecular systems where the boundaries are "natural" and imposed by the energetics, or like diffusive motions in restricted geometries where the boundaries are physical impenetrable barriers externally imposed, here the boundaries are inherent in the dynamics of the system that cannot reach nonphysical configurations by means of finite moves determined by the stoichiometry. To our knowledge, the behavior at the boundaries for the CFPE has not been addressed properly yet.

IV. NONPHYSICAL PROBABILITY CURRENTS AT EQUILIBRIUM

In the typical diffusion equations (i.e., Fokker-Planck equations in the Smoluchowski form[1]) encountered in the physics of overdamped fluctuating systems at thermal equilibrium, the diffusion matrix is tuned in a way that an equilibrium state devoid of probability currents is attained in the long time scale. Here, on the contrary, *both* the drift vector and the diffusion matrix given in Eqs. (13) and (14) are determined by the stochastic evolution law of the reaction network under the approximations at the basis of the CLE. This implies that diffusion and drift might be generally unbalanced in the sense that the vanishing of the probability currents at equilibrium

may not be guaranteed. This is the crucial issue investigated in what follows.

Let us consider the reduced CFPE once the mass-conservation constraints are enforced as described in Sec. III A. The independent variables are $\tilde{\eta}_1, \tilde{\eta}_2, \ldots, \tilde{\eta}_{i'}, \ldots, \tilde{\eta}_{N-d}$, collected in the array $\tilde{\boldsymbol{\eta}}$; in what follows, the indexes with the prime will label such variables. The reduced CFPE can be written in the form

$$\frac{\partial \rho(\tilde{\boldsymbol{\eta}}, t)}{\partial t} = -\frac{\partial}{\partial \tilde{\boldsymbol{\eta}}}^{T} \mathbf{J}(\tilde{\boldsymbol{\eta}}, t), \tag{15}$$

where $\mathbf{J}(\tilde{\boldsymbol{\eta}}, t)$ is the probability current vector whose components are

$$J_{i'}(\tilde{\boldsymbol{\eta}}, t) = \tilde{v}_{i'}(\tilde{\boldsymbol{\eta}})\rho(\tilde{\boldsymbol{\eta}}, t) - \frac{1}{2}\sum_{j'} \frac{\partial[\tilde{B}_{i'j'}(\tilde{\boldsymbol{\eta}})\rho(\tilde{\boldsymbol{\eta}}, t)]}{\partial \tilde{\eta}_{j'}}. \tag{16}$$

The probability current vector is such that, given an oriented surface $\delta\Omega_+$ in the $\tilde{\boldsymbol{\eta}}$-space, the flux $\int_{\delta\Omega_+} d\sigma(\tilde{\boldsymbol{\eta}})\, \hat{\mathbf{s}}(\tilde{\boldsymbol{\eta}})^T \mathbf{J}(\tilde{\boldsymbol{\eta}}, t)$ gives the rate of probability transfer through that surface [in the integral, $d\sigma(\tilde{\boldsymbol{\eta}})$ is the area of a surface element centered in $\tilde{\boldsymbol{\eta}}$ and $\hat{\mathbf{s}}(\tilde{\boldsymbol{\eta}})$ is the unit vector normal to such oriented surface element].

Let us assume that a unique stationary state is reached in the long time scale, with $\lim_{t\to\infty} p(\mathbf{n}, t) = p_{ss}(\mathbf{n})$; correspondingly, $\lim_{t\to\infty} \rho(\tilde{\boldsymbol{\eta}}, t) = \rho_{ss}(\tilde{\boldsymbol{\eta}})$. By elaborating Eq. (16) it follows

$$\mathbf{J}_{ss}(\tilde{\boldsymbol{\eta}}) = [\tilde{\mathbf{v}}(\tilde{\boldsymbol{\eta}}) - \tilde{\mathbf{b}}(\tilde{\boldsymbol{\eta}})]\rho_{ss}(\tilde{\boldsymbol{\eta}}) - \frac{1}{2}\tilde{\mathbf{B}}(\tilde{\boldsymbol{\eta}})\frac{\partial \rho_{ss}(\tilde{\boldsymbol{\eta}})}{\partial \tilde{\boldsymbol{\eta}}}, \tag{17}$$

where $\mathbf{J}_{ss}(\tilde{\boldsymbol{\eta}}) = \lim_{t\to\infty} \mathbf{J}(\tilde{\boldsymbol{\eta}}, t)$ and $\tilde{\mathbf{b}}(\tilde{\boldsymbol{\eta}})$ is the column vector with components

$$\tilde{b}_{i'}(\tilde{\boldsymbol{\eta}}) = \frac{1}{2}\sum_{j'} \frac{\partial \tilde{B}_{i'j'}(\tilde{\boldsymbol{\eta}})}{\partial \tilde{\eta}_{j'}}. \tag{18}$$

The specific topology of $\mathbf{J}_{ss}(\tilde{\boldsymbol{\eta}})$ depends on the features of the reaction network which, in all generality, may lack detailed-balance.

Let us now focus on a closed network of reversible and detailed-balanced reactions.[29] The detailed-balance condition is here referred to the reaction network in the thermodynamic limit, and it consists in having the same rate for each forward/backward pair of elementary processes. In such a condition, the stationary point in the concentration space corresponds to the point of thermodynamic equilibrium. When *the same* network is brought down to the stochastic context, the corresponding CME yields a stationary state which corresponds to the thermal equilibrium. Thus, $p_{ss}(\mathbf{n}) \equiv p_{eq}(\mathbf{n})$ and $\rho_{ss}(\tilde{\boldsymbol{\eta}}) \equiv \rho_{eq}(\tilde{\boldsymbol{\eta}})$. In such a situation, physics imposes that all components of the probability current *must* be zero, i.e., $\mathbf{J}_{ss}(\tilde{\boldsymbol{\eta}}) \equiv \mathbf{J}_{eq}(\tilde{\boldsymbol{\eta}}) = \mathbf{0}$; otherwise, there would be a directed (on average) motion in the $\tilde{\boldsymbol{\eta}}$-space for free. By introducing the scalar field

$$\Phi(\tilde{\boldsymbol{\eta}}) = -\ln \rho_{eq}(\tilde{\boldsymbol{\eta}}) \tag{19}$$

and considering that the positive definite matrix $\tilde{\mathbf{B}}(\tilde{\boldsymbol{\eta}})$ is invertible, the required vanishing of the right-hand side of Eq. (17) implies

$$\frac{\partial \Phi(\tilde{\boldsymbol{\eta}})}{\partial \tilde{\boldsymbol{\eta}}} = \boldsymbol{\Psi}(\tilde{\boldsymbol{\eta}}), \tag{20}$$

where for the sake of compactness we have introduced the new vector

$$\boldsymbol{\Psi}(\tilde{\boldsymbol{\eta}}) = 2\tilde{\mathbf{B}}(\tilde{\boldsymbol{\eta}})^{-1}\left[\tilde{\mathbf{b}}(\tilde{\boldsymbol{\eta}}) - \tilde{\mathbf{v}}(\tilde{\boldsymbol{\eta}})\right]. \tag{21}$$

Equation (20) with (21) is known as the "potential equation"[1] and constitutes the mathematical requirement to have a stationary state with null currents: *if* there exists a scalar field $\Phi(\tilde{\boldsymbol{\eta}})$ such that its gradient generates identically the vector $\boldsymbol{\Psi}(\tilde{\boldsymbol{\eta}})$, *then* $\mathbf{J}_{eq}(\tilde{\boldsymbol{\eta}}) = \mathbf{0}$ can be fulfilled; on the contrary, the dynamics of the system would be such that (as an artifact) the probability current at equilibrium would be non-null.

Since Eq. (20) states that $\boldsymbol{\Psi}(\tilde{\boldsymbol{\eta}})$ must be a conservative vector field, a way to check this property is to verify if the following condition holds identically:

$$\frac{\partial \Psi_{i'}(\tilde{\boldsymbol{\eta}})}{\partial \tilde{\eta}_{j'}} = \frac{\partial \Psi_{j'}(\tilde{\boldsymbol{\eta}})}{\partial \tilde{\eta}_{i'}} \quad \text{for all } i', j'. \tag{22}$$

An alternative route is to verify if the path integrals of $\boldsymbol{\Psi}(\tilde{\boldsymbol{\eta}})$ between any two points A and B arbitrarily chosen are independent of the path. Explicitly, the required condition is

$$\int_0^1 ds\, \hat{\boldsymbol{\gamma}}(\tilde{\boldsymbol{\eta}}^{\gamma}(s))^T \boldsymbol{\Psi}(\tilde{\boldsymbol{\eta}})|_{\tilde{\boldsymbol{\eta}}=\tilde{\boldsymbol{\eta}}^{\gamma}(s)} = \text{const} \tag{23}$$

for any curve γ connecting A with B [here, $0 \le s \le 1$ is a progression variable, $\tilde{\boldsymbol{\eta}}^{\gamma}(s)$ is the corresponding point on the curve in the $\tilde{\boldsymbol{\eta}}$-space, and $\hat{\boldsymbol{\gamma}}(\tilde{\boldsymbol{\eta}}^{\gamma}(s))$ is the tangent versor to the curve in that point].

If the violation of Eq. (22) or Eq. (23) were recognized even by a single check, and even for a single closed and detailed-balanced reaction network, then one would conclude that the CFPE (and the CLE as well) is inconsistent with the condition that at thermal equilibrium the probability current must be identically null. In Sec. V we show, for a simple case, that Eqs. (22) and (23) are indeed violated. However, there may be cases in which the current at equilibrium is unequivocally null. This is certainly the case when the array $\tilde{\boldsymbol{\eta}}$ reduces to a single variable $\tilde{\eta}$, so that the potential equation Eq. (20) is fulfilled since $\Phi(\tilde{\eta})$ can be determined by integration: $\Phi(\tilde{\eta}) = \int^{\tilde{\eta}} d\tilde{\eta}' \Psi(\tilde{\eta}')$. For example, this happens for the dimerization $\mathrm{A} \rightleftharpoons \mathrm{B}$ discussed in Ref. 30, where the conservation constraint $n_{\mathrm{A}} + n_{\mathrm{B}} = c$ allows one to use the sole variable $\tilde{\eta} \equiv \eta_{\mathrm{A}}$ to specify the composition of the system. Going to higher dimensions, however, a potential $\Phi(\tilde{\boldsymbol{\eta}})$ that fulfills Eq. (20) cannot be found in all generality.

V. ILLUSTRATIVE EXAMPLE AND REMARKS

To show that the physical inconsistency addressed in Sec. IV occurs in practice, we adopt the simple kinetic scheme

$$2\mathrm{X} \underset{-1}{\overset{1}{\rightleftarrows}} \mathrm{X} + \mathrm{Y}$$

$$\mathrm{X} + \mathrm{Y} \underset{-2}{\overset{2}{\rightleftarrows}} \mathrm{Z}$$

$$\mathrm{Z} \underset{-3}{\overset{3}{\rightleftarrows}} 2\mathrm{X},$$

where the numbers attached to the arrows label the reaction channels. Such a network is closed, reversible in all the reac-

tion channels, and connected (meaning that each configuration of reactants is directly connected with the others). In addition, we choose the following values for the rate coefficients c_m that enter the propensity functions as expressed in Sec. II: $c_1 = 2$, $c_{-1} = 3$, $c_2 = 2$, $c_{-2} = 500$, $c_3 = 75$, and $c_{-3} = 0.2$. These values are meant to be given in some units of inverse-of-time that are immaterial in this context. By turning to the corresponding kinetic constants, one has that $k_1k_2k_3 = k_{-1}k_{-2}k_{-3}$; hence, the reaction network is also detailed-balanced as discussed in Sec. IV. This implies that the system reaches a stationary state of equilibrium. The components of $\mathbf{n}$ and $\boldsymbol{\eta}$ are, respectively, n_X, n_Y, n_Z and η_X, η_Y, η_Z. The network owns the conservation constraint $n_X + n_Y + 2n_Z = \text{const}$; the value of such a constant was set to 2×10^4 in the present calculations. Such a constraint allows us to take only the species X and Y as independent; in particular, the components of the reduced set $\tilde{\boldsymbol{\eta}}$ are $\tilde{\eta}_1 = \eta_X$ and $\tilde{\eta}_2 = \eta_Y$.

Panel (a) of Fig. 1 shows five trajectories simulated with the standard SSA. The panel also shows the triangular intersection between the hyperplane corresponding to the conservation constraint and the faces of the positive orthant. Note that, on such a scale, the erratic character of the trajectories is hardly detectable. The fluctuations are evident in panel (b), where one of the trajectories is shown on a smaller scale in the reduced space of the species X and Y. The equilibrium distribution $p_{eq}(\mathbf{n})$ is here presented in color-scale. To construct the distribution, 10^6 trajectories were generated by means of the standard SSA. All trajectories were initiated from a state close to the equilibrium point of the deterministic rate equations for unitary volume of the sample; this ensures a good statistics for the configurations mostly visited at the thermal equilibrium. Each simulation was stopped at the time 0.02. By collecting the final states, the equilibrium distribution was obtained by a histogram construction. A convergence check was made by verifying that the distribution is indistinguishable from the one generated by collecting the final states at the shorter time 0.01.

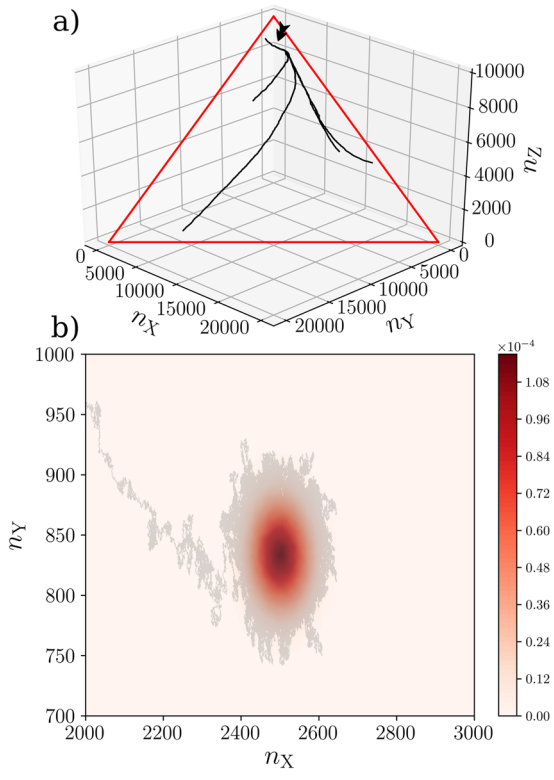

FIG. 1. (a) Five trajectories for the model reaction network simulated by means of the standard SSA method (see the text for details). The trajectories lie on the hyperplane corresponding to the mass conservation constraint $n_X + n_Y + 2n_Z = 2 \times 10^4$; the intersections of such a plane with the faces of the positive orthant are indicated by the red lines. (b) Equilibrium distribution in the reduced space of the species X and Y. The distribution has been obtained from 10^6 trajectories generated with the standard SSA method. The trajectory displayed is the one indicated by the arrow in (a).

As stated in Sec. III A, the CLE is strictly applicable only in the configurational region where Eq. (7) holds. To identify such a region in the reduced $\tilde{\boldsymbol{\eta}}$-space, we computed the ratio $\Delta t_{\min}(\tilde{\boldsymbol{\eta}})/\Delta t_{\max}(\tilde{\boldsymbol{\eta}})$ in the domain with $1 < \eta_X < 10^4$ and $1 < \eta_Y < 10^4$. For each state $\tilde{\boldsymbol{\eta}}$, $\Delta t_{\min}(\tilde{\boldsymbol{\eta}})$ was calculated according to Eq. (6) with $\gamma = 3$. For $\Delta t_{\max}(\tilde{\boldsymbol{\eta}})$ we followed the optimized tau-leaping procedure illustrated in Ref. 25 (see Sec. II C 1 along with Sec. IV A therein) and adopted the same computational parameters employed in that work.[31] With these criteria, the filled area in Fig. 2 represents the points for which $\Delta t_{\min}(\tilde{\boldsymbol{\eta}})/\Delta t_{\max}(\tilde{\boldsymbol{\eta}}) \leq 1$. It can be seen that such a region covers a limited portion of the explored $\tilde{\boldsymbol{\eta}}$-space. Notably, such a portion encloses the cloud of states typically visited by the equilibrium fluctuations. This tells us that the CLE is suitable for simulating the thermal fluctuations of this specific reactive system with the adopted parametrization. We stress that, however, this is not a general situation since the cloud of typically visited states might fall outside the region of applicability of the CLE. Note that other criteria to fix $\Delta t_{\min}(\tilde{\boldsymbol{\eta}})$ and $\Delta t_{\max}(\tilde{\boldsymbol{\eta}})$ would have led to a somehow different outcome; however, our purpose here is mainly to remark that the CLE has a limited region of applicability.

Let us now turn to the main issue, that is, showing that the vector field $\boldsymbol{\Psi}(\tilde{\boldsymbol{\eta}})$ in Eq. (21) is not a conservative field, implying that the vanishing of the probability current at equilibrium cannot be exactly satisfied in the whole accessible $\tilde{\boldsymbol{\eta}}$-space. For a given state $\tilde{\boldsymbol{\eta}}$, the vector $\boldsymbol{\Psi}(\tilde{\boldsymbol{\eta}})$ was obtained via Eq. (21) with

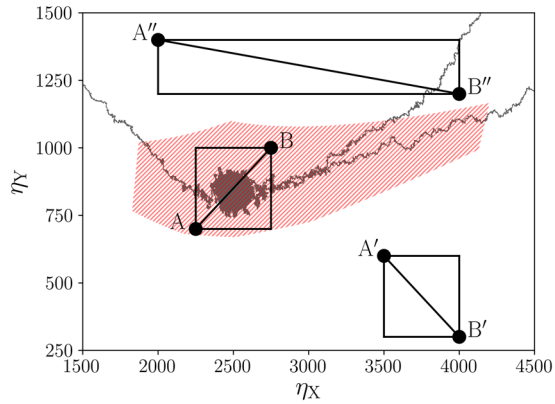

FIG. 2. The filled area corresponds to the domain of applicability of the CLE for the model reaction network under the constraint $\eta_X + \eta_Y + 2\eta_Z = 2 \times 10^4$ (see the text for details). The pairs of points (A, B), (A′, B′), and (A″, B″) are the chosen edge points to compute path integrals along the displayed connecting paths (three paths per each pair of points). The values of the integrals are reported in Table I.

$\tilde{\mathbf{b}}(\tilde{\boldsymbol{\eta}})$ from Eq. (18); the derivatives required in Eq. (18) and the matrix inversion in Eq. (21) were performed analytically. First, it is found that the equivalence Eq. (22) is violated. To see this, we considered the factor

$$\mathcal{R}(\tilde{\boldsymbol{\eta}}) = \frac{\dfrac{\partial \Psi_X(\tilde{\boldsymbol{\eta}})}{\partial \eta_Y} - \dfrac{\partial \Psi_Y(\tilde{\boldsymbol{\eta}})}{\partial \eta_X}}{\left|\dfrac{\partial \Psi_X(\tilde{\boldsymbol{\eta}})}{\partial \eta_Y}\right| + \left|\dfrac{\partial \Psi_Y(\tilde{\boldsymbol{\eta}})}{\partial \eta_X}\right|} \tag{24}$$

as a function of the system's state (the derivatives are here computed by means of finite differences). By construction, such a factor is bounded between −1 and 1, and Eq. (22) would be satisfied only if $\mathcal{R}$ were identically null. In the present case, it is found that $\mathcal{R}$ has a marked variation in the explored domain, as shown by the contour plot in Fig. 3. This means that Eq. (22) is violated and hence the vector field $\boldsymbol{\Psi}(\tilde{\boldsymbol{\eta}})$ is not conservative. Second, we can arrive at the same conclusion by noting that Eq. (23) is also violated. In Fig. 2 are shown the pairs of edge points adopted to compute different path integrals according to Eq. (23).[32] For each pair, three connecting paths are chosen: an upper two-segment path, a straight diagonal path, and a lower two-segment path, as indicated in the figure. The numerical results are presented in Table I. It can be seen that, for all the chosen pairs of edge points, the three integrals are different from each other.

As a whole, the numerical investigations have shown that $\boldsymbol{\Psi}(\tilde{\boldsymbol{\eta}})$ is not a conservative field for this simple case model. Having detected such a fact for a single case implies that this is a concrete issue concerning the CFPE/CLE in all generality.

It might be the case, however, that the non-null spurious probability current has no marked signatures on the solution of the CFPE and on the statistics of the CLE trajectories. In this regard, it is worth noting that the percentage spread between the three path integrals for A → B is much smaller than that for A′ → B′, despite the fact that the distance between the edge points is the same in the two cases. Comparing the two situations, we note that A and B closely surround the cloud

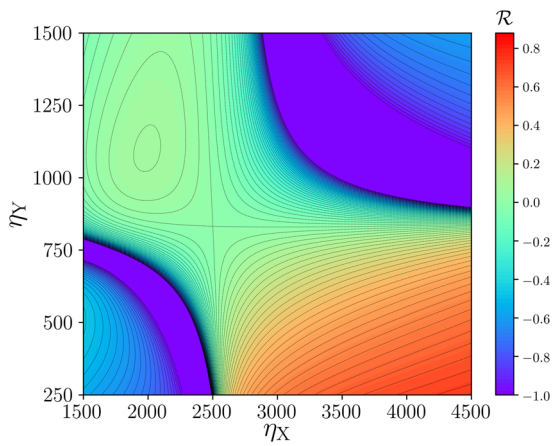

FIG. 3. State dependence of the parameter $\mathcal{R}$, defined in Eq. (24), for the model reaction network under the constraint $\eta_X + \eta_Y + 2\eta_Z = 2 \times 10^4$ (see the text for details). Note that $\mathcal{R}$ is close to zero in the region corresponding to the most visited configurations at the thermal equilibrium [compare with (b) of Fig. 1].

TABLE I. Values of the path integrals for the edge points shown in Fig. 2. For each pair of points, the first value refers to the upper two-segment path, the second value refers to the diagonal path, and the third value refers to the lower two-segment path. The spread is intended as relative percentage dispersion of the extreme values with respect to the average of the three values.

	Value of the path integral	Spread
A → B	4.246 4.229 4.222	0.6%
A′ → B′	349.258 362.724 372.258	6.4%
A″ → B″	274.781 273.510 268.661	2.2%

of states mostly visited by the fluctuations at equilibrium [see panel (b) of Fig. 1]; we also note that in such a region the factor $\mathcal{R}$ in Fig. 3 is nearly zero. As a whole, one might provisionally conclude, at least for this simple reaction network with the adopted parameters, that $\boldsymbol{\Psi}(\tilde{\boldsymbol{\eta}})$ is "almost conservative" inside the region of states typically visited at thermal equilibrium. Conversely, $\boldsymbol{\Psi}(\tilde{\boldsymbol{\eta}})$ is manifestly non-conservative far from the equilibrium cloud; on the other hand, here $\rho_{eq}(\tilde{\boldsymbol{\eta}})$ is nearly flat and low in magnitude. By considering that $\rho_{eq}(\tilde{\boldsymbol{\eta}})$ and its gradient enter the expression of $\mathbf{J}_{eq}(\tilde{\boldsymbol{\eta}})$ [see Eq. (17)], it is reasonable to expect a small probability current far from the typical equilibrium configurations. As a whole, $\mathbf{J}_{eq}(\tilde{\boldsymbol{\eta}})$ could be small everywhere in the $\tilde{\boldsymbol{\eta}}$-space. A quantitative validation of such a statement should be made case by case, but the analysis is hampered by the fact that $p_{eq}(\mathbf{n})$ (generally hardly accessible) and its smooth interpolation $\rho_{eq}(\tilde{\boldsymbol{\eta}})$ would be required.

VI. CONCLUSIONS

In this work we reviewed the chemical Langevin equation (CLE) and the associated chemical Fokker-Planck equation (CFPE) as approximate continuous formulations of the stochastic chemical kinetics. In doing that, we focused on a physical inconsistency, namely, the possible presence of nonphysical probability currents at equilibrium even for closed and fully detailed-balanced networks of elementary reactions. The analysis of the case model reported in Sec. V supports the concreteness of such an issue.

As pointed out at the end of Sec. IV, such an issue may manifest only in multidimensional cases. Previously detailed analyses of the CFPE were focused on chemical networks reducible to one-dimensional problems[17–19,30] for which the nonphysical currents are absent. Other studies on multidimensional systems mostly regarded the accuracy of the CLE/CFPE, with respect to the CME, in terms of mean concentrations and variance of the fluctuations about the mean.[16] Although all inspections on the CLE/CFPE explore different facets of the same problem, to the best of our knowledge the issue of nonphysical probability flow at equilibrium was not inspected so far.

We emphasize once again that the presence of nonphysical probability currents regards both the CLE and CFPE, since they are fully consistent with each other. However, the use of the CLE in the production of a stochastic trajectory can be locally suspended (switching to the exact SSA propagation), while the CFPE, be it a partial-derivative equation for the evolution of the probability density field on the *global* scale, has to be solved without the possibility to impose a delimitation *a priori* of the configurational space. Thus, such a "flexibility" of the CLE might allow one, through a suitable algorithmic implementation, to produce an ensemble of trajectories for which the impact of the nonphysical currents is attenuated. On the contrary, the solution of the CFPE should be taken with caution. At any rate, the solution of the CFPE is quite demanding, both in terms of computational cost (especially when the number of independent species is just above a few units) and in terms of difficulty of enforcing reflecting conditions at the boundaries. For this reason, the CFPE appears to us more as a formal construction to be further inspected rather than a practical tool for describing the dynamics of stochastic reaction networks.

[1] C. W. Gardiner, *Handbook of Stochastic Methods*, 3rd ed. (Springer-Verlag, Berlin, 2004).

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

[3] D. T. Gillespie, A. Hellander, and L. R. Petzold, "Perspective: Stochastic algorithms for chemical kinetics," J. Chem. Phys. **138**, 170901 (2013).

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

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

[6] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, 1992).

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

[8] T. G. Kurtz, "Limit theorems and diffusion approximations for density dependent Markov chains," in *Stochastic Systems: Modeling, Identification and Optimization, I*, Volume 5 of Mathematical Programming Studies, edited by R. J. B. Wets (Springer, Berlin, Heidelberg, 1976).

[9] T. G. Kurtz, "Strong approximation theorems for density dependent Markov chains," Stochastic Processes Appl. **6**, 223–240 (1978).

[10] D. Adalsteinsson, D. McMillen, and T. C. Elston, "Biochemical network stochastic simulator (BioNetS): Software for stochastic modeling of biochemical networks," BMC Bioinf. **5**, 24 (2004).

[11] D. Schnoerr, G. Sanguinetti, and R. Grima, "Approximation and inference methods for stochastic biochemical kinetics—A tutorial review," J. Phys. A: Math. Theor. **50**, 093001 (2017).

[12] D. T. Gillespie, "Deterministic limit of stochastic chemical kinetics," J. Phys. Chem. B **113**, 1640–1644 (2009).

[13] L. Wang, X. Han, Y. Cao, and H. N. Najm, "Computational singular perturbation analysis of stochastic chemical systems with stiffness," J. Comput. Phys. **335**, 404–425 (2017).

[14] R. R. Coifman, I. G. Kevrekidis, S. Lafon, M. Maggioni, and B. Nadler, "Diffusion maps, reduction coordinates, and low dimensional representation of stochastic systems," Multiscale Model. Simul. **7**(2), 842–864 (2008).

[15] B. Mélykúti, K. Burrage, and K. C. Zygalakis, "Fast stochastic simulation of biochemical reaction systems by alternative formulations of the chemical Langevin equation," J. Phys. Chem. **132**, 164109 (2010).

[16] R. Grima, P. Thomas, and A. V. Straube, "How accurate are the nonlinear chemical Fokker-Planck and chemical Langevin equations?," J. Chem. Phys. **135**(8), 084103 (2011).

[17] P. Hanggi, H. Grabert, P. Talkner, and H. Thomas, "Bistable systems: Master equation versus Fokker-Planck modeling," Phys. Rev. A **29**(1), 371–378 (1984).

[18] M. Velleda and H. Qian, "Stochastic dynamics and non-equilibrium thermodynamics of a bistable chemical system: The Schlögl model revisited," J. R. Soc., Interface **6**, 925–940 (2009).

[19] D. Zhou and H. Qian, "Fixation, transient landscape, and diffusion dilemma in stochastic evolutionary game dynamics," Phys. Rev. E **84**, 031907 (2011).

[20] J. M. Horowitz, "Diffusion approximations to the chemical master equation only have a consistent stochastic thermodynamics at chemical equilibrium," J. Chem. Phys. **143**(4), 044111 (2015).

[21] The set of equations takes the form $\dot{\mathbf{p}} = -\mathbf{K}\,\mathbf{p}$, where $\mathbf{p}$ is the column-vector with entries $p_{\mathbf{n}}(t) \equiv p(\mathbf{n}, t)$ and $\mathbf{K}$ is the matrix with elements $K_{\mathbf{n},\mathbf{n}'} = a_0(\mathbf{n}')\delta_{\mathbf{n},\mathbf{n}'} - \sum_{m=1}^{M} \delta_{\mathbf{n}',\mathbf{n}-\boldsymbol{\nu}_m}\, a_m(\mathbf{n}')$, where δ is the Kronecker's delta-function and $a_0(\mathbf{n}) = \sum_{m=1}^{M} a_m(\mathbf{n})$. The dimension of the arrays to be handled is fixed by the number N_{conf} of system's configurations which are reachable by the starting point $\mathbf{n}_0$. The formal solution of the CME is thus $p(\mathbf{n}, t) = \left[e^{-\mathbf{K}t}\right]_{\mathbf{n},\mathbf{n}_0}$.

[22] B. Musky and M. Khammash, "The finite state projection algorithm for the solution of the chemical master equation," J. Chem. Phys. **124**(4), 044104 (2006).

[23] Z. Fox, G. Neuert, and B. Munsky, "Finite state projection based bounds to compare chemical master equation models using single-cell data," J. Chem. Phys. **145**(7), 074101 (2016).

[24] S. MacNamara, K. Burrage, and R. B. Sidje, "Multiscale modeling of chemical kinetics via the master equation," Multiscale Model. Simul. **6**(4), 1146–1168 (2008).

[25] Y. Cao, D. T. Gillespie, and L. R. Petzold, "Efficient step size selection for the tau-leaping simulation method," J. Chem. Phys. **124**(4), 044109 (2006).

[26] $\langle \xi_m(t) \rangle = 0$ and $\langle \xi_m(t)\xi_{m'}(t') \rangle = \delta_{m,m'}\delta_D(t - t')$, where "$\delta$" is the Kronecker's delta and "δ_D" is the Dirac's delta-function; the averages $\langle \cdots \rangle$ are meant to be taken over the statistical ensemble of realizations.

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

[28] Note that if the differential form in Eq. (9) were adopted to parametrize the Fokker-Planck equation without any information about how Eq. (9) itself was derived, there would be the ambiguity of choosing between Itô and Stratonovich methods for stochastic integration. However, the fact that Eq. (9) follows from the integrated form Eq. (8) indicates that Itô's route is the natural choice [in this way, one can go back from Eq. (9) to the discrete explicit advancement in Eq. (8)]. At any rate, Stratonovich integration would lead to an alternative Fokker-Planck equation in which the drift term is given by $\boldsymbol{v}(\boldsymbol{\eta})$ in Eq. (13) plus a correction smaller than $\boldsymbol{v}(\boldsymbol{\eta})$ by a factor of the order of the average number of reactant molecules. Thus, since the CLE is valid for large numbers of molecules, such a different Fokker-Planck would reduce to the CFPE in that limit.

[29] R. Rao and M. Esposito, "Nonequilibrium thermodynamics of chemical reaction networks: Wisdom from stochastic thermodynamics," Phys. Rev. X **6**, 041064 (2016).

[30] D. T. Gillespie, "The chemical Langevin and Fokker-Planck equations for the reversible isomerization reaction," J. Phys. Chem. A **106**, 5063–5071 (2002).

[31] Namely, $n_c = 10$ for the identification of the "critical reactions" and $\epsilon = 0.03$ in Eq. (33) of Ref. 25. A "critical reaction" is any reaction for which there are at most n_c firings left before one of its reactants disappears. The parameter $0 < \epsilon < 1$ approximately bounds the relative change of each propensity function. We did not consider steps (3) and (6) in the procedure of Ref. 25, since these steps are important only for the generation of stochastic trajectories.

[32] The integrations along the paths were performed numerically by means of a FORTRAN routine employing Romberg's method; a convergence check was made with respect to the variation of the parameters of accuracy and tolerance. In addition, the integration route was tested by checking the invariance of the path integrals for benchmark conservative fields.