PHYSICAL REVIEW E **101**, 032417 (2020)

Chemical Langevin equation: A path-integral view of Gillespie's derivation

John J. Vastola[*]
Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee, USA
and Quantitative Systems Biology Center, Vanderbilt University, Nashville, Tennessee 37235, USA

William R. Holmes
Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee, USA;
Quantitative Systems Biology Center, Vanderbilt University, Nashville, Tennessee 37235, USA;
and Department of Mathematics, Vanderbilt University, Nashville, Tennessee 37235, USA

(Received 22 October 2019; revised manuscript received 22 February 2020; accepted 25 February 2020; published 31 March 2020)

In 2000, Gillespie rehabilitated the chemical Langevin equation (CLE) by describing two conditions that must be satisfied for it to yield a valid approximation of the chemical master equation (CME). In this work, we construct an original path-integral description of the CME and show how applying Gillespie's two conditions to it directly leads to a path-integral equivalent to the CLE. We compare this approach to the path-integral equivalent of a large system size derivation and show that they are qualitatively different. In particular, both approaches involve converting many sums into many integrals, and the difference between the two methods is essentially the difference between using the Euler-Maclaurin formula and using Riemann sums. Our results shed light on how path integrals can be used to conceptualize coarse-graining biochemical systems and are readily generalizable.

DOI: 10.1103/PhysRevE.101.032417

I. INTRODUCTION

Gillespie's classic paper [1] on how to derive the chemical Langevin equation (CLE) from the chemical master equation (CME) proceeds differently than by naively truncating the Kramers-Moyal expansion of the CME [2–4] or by invoking the largeness of the system volume Ω *a la* van Kampen [5,6]; instead, he argues based on the existence of a *timescale* with certain properties. In particular, his derivation avoids rewriting discrete number variables n as concentration variables $x := n/\Omega$.

By writing down two precise conditions that control the validity of the CLE (to be reviewed in Sec. II), he rehabilitated it as a well-founded approach to approximating stochastic dynamics described by the CME (in the face of ostensible no-go results like the Pawula theorem [7]), and directly inspired the τ-leaping algorithm [8] and its many modifications [9–14] for speeding up numerical simulations of biochemical reactions.

Path integrals offer a way to think about stochastic processes that is somewhat independent from the usual differential equations perspective [15]. This means that—at least in principle—there should be a way to translate Gillespie's derivation into path-integral language. Because path integrals (along with associated technology like the renormalization group [16–19]) are known to be useful for understanding coarse-grained descriptions of systems (e.g., effective field theories [19,20]), such a translation should contribute meaningfully to our understanding of how to intelligently coarse-grain biochemical systems.

In this paper, we show how Gillespie's two conditions translate to a path-integral-based derivation of the chemical Langevin equation. Our approach here builds upon the path-integral descriptions of Langevin/Fokker-Planck equations described in Ref. [15]. We will proceed with little mathematical rigor (as is typical in physics), but with enough clarity that our arguments could in principle be made mathematically precise.

The paper is organized as follows. In Sec. II, we review Gillespie's derivation of the CLE. In Sec. III, we construct a path-integral description of CME dynamics. In Sec. IV, we apply Gillespie's conditions to our path-integral formulation to obtain the CLE, and we also discuss an alternative method based on a large system volume argument. Finally, we discuss consequences of our work for understanding coarse-grained biochemical systems in Sec. V.

II. REVIEW OF GILLESPIE'S CHEMICAL LANGEVIN EQUATION DERIVATION

In this section, we review Gillespie's derivation [1] of the chemical Langevin equation from the chemical master equation. We use the same notation Gillespie used in his paper, although we will not require that the same physical assumptions (i.e., well-stirred, dilute chemicals in a fixed volume and at constant temperature) hold, because the derivation does not depend on them.

Consider a system with N species and M reactions. Denote the propensity function of the jth reaction by a_j, and the

---

[*] john.j.vastola@vanderbilt.edu

corresponding stoichiometry vector by $\boldsymbol{\nu}_j$. The chemical master equation reads

$$\frac{\partial P(\mathbf{n},t)}{\partial t} = \sum_{j=1}^{M} a_j(\mathbf{n}-\boldsymbol{\nu}_j)P(\mathbf{n}-\boldsymbol{\nu}_j,t) - a_j(\mathbf{n})P(\mathbf{n},t), \quad (1)$$

where $P(\mathbf{n},t)$ is the probability that the state of the system is $\mathbf{n} = (n_1, ..., n_N) \in \mathbb{N}^N$ at time t.

Gillespie's derivation requires the existence of a timescale τ for which the following two conditions hold:

(i) The propensity functions do not change their values appreciably,

i.e., $a_j(\mathbf{n}(t)) \approx a_j(\mathbf{n}(t'))$ for all j and all $t' \in [t, t+\tau]$.

(ii) The average number of firings of each reaction over a time τ is much larger than 1.

Due to their connection with the τ-leaping algorithm [8–14] for approximately simulating CME dynamics, Gillespie later called these the first leap condition and the second leap condition [21]. They are in practice easily satisfied in the case of large molecule numbers, and they are exactly satisfied in the thermodynamic limit [22], where the system volume Ω is taken to infinity while keeping all concentrations fixed.

Consider $n_i(t)$, the number of molecules corresponding to species i at time t. It changes in a small time Δt according to

$$n_i(t+\Delta t) = n_i(t) + \sum_{j=1}^{M} \nu_{ji} K_j(a_j, \Delta t), \quad (2)$$

where ν_{ji} is the ith component of the stoichiometry vector $\boldsymbol{\nu}_j$ (i.e., the change in number of species i due to reaction j firing once), and $K_j(a_j, \Delta t)$ is a random variable that describes the number of times reaction j fires in Δt.

For an arbitrary CME and arbitrary length of time Δt, K_j might be taken from a complicated distribution. But if condition (i) holds in a length of time τ, each reaction fires independently of each other reaction, because no reactions significantly change any propensity functions. Because (by definition) the probability of reaction j firing in an infinitesimal time dt is $a_j(\mathbf{n}(t))dt$, and because that probability will not significantly change during the time length τ, the number of times reaction j fires in τ is well-approximated as a Poisson random variable with mean $a_j(\mathbf{n}(t))\tau$, which we will denote by $\mathcal{P}_j\{a_j(\mathbf{n}(t))\tau\}$.

This means that when condition (i) holds we can write the time evolution of $n_i(t)$ over a length of time τ as

$$n_i(t+\tau) = n_i(t) + \sum_{j=1}^{M} \nu_{ji} \mathcal{P}_j\{a_j(\mathbf{n}(t))\tau\}. \quad (3)$$

This equation is the basis for the τ-leaping approach first described by Gillespie in 2001 [8], and later modified and extended by himself and others [9–14].

If condition (ii) holds, then the average number of times reaction j fires in τ (i.e., $a_j(\mathbf{n}(t))\tau$) is much larger than 1, so the Poisson random variables are well-approximated by normal random variables:

$$\mathcal{P}_j\{a_j(\mathbf{n}(t))\tau\} \approx \mathcal{N}_j\{a_j(\mathbf{n}(t))\tau, a_j(\mathbf{n}(t))\tau\}, \quad (4)$$

where $\mathcal{N}_j\{a_j(\mathbf{n}(t))\tau, a_j(\mathbf{n}(t))\tau\}$ is a normal random variable with mean and variance both equal to $a_j(\mathbf{n}(t))\tau$. If we also note that each normal random variable can be decomposed as

$$\begin{aligned}&\mathcal{N}_j\{a_j(\mathbf{n}(t))\tau, a_j(\mathbf{n}(t))\tau\}\\ &\quad = a_j(\mathbf{n}(t))\tau + \sqrt{a_j(\mathbf{n}(t))\tau}\, \mathcal{N}_j(0,1),\end{aligned} \quad (5)$$

then we can write the time evolution of $n_i(t)$ in a time τ as

$$\begin{aligned} n_i(t+\tau) = n_i(t) &+ \sum_{j=1}^{M} \nu_{ji} a_j(\mathbf{n}(t))\tau \\ &+ \sum_{j=1}^{M} \nu_{ji} \sqrt{a_j(\mathbf{n}(t))\tau}\, \mathcal{N}_j(0,1). \end{aligned} \quad (6)$$

Because this equation has the form of an Euler-Maruyama time step, we can identify the dynamics of the system on the timescale τ with the set of N stochastic differential equations (SDEs)

$$\dot{x}_i = \sum_{j=1}^{M} \nu_{ji} a_j(\mathbf{x}) + \sum_{j=1}^{M} \nu_{ji} \sqrt{a_j(\mathbf{x})}\, \Gamma_j, \quad (7)$$

where the Γ_j are M independent Gaussian white noise terms, and where we have relabeled each n_i as x_i to emphasize that we are now working with continuous variables.

Our chemical Langevin equation corresponds to a chemical Fokker-Planck equation

$$\begin{aligned}\frac{\partial P(\mathbf{x},t)}{\partial t} &= \sum_{i=1}^{N} -\frac{\partial}{\partial x_i}\left[\left(\sum_{j=1}^{M} \nu_{ji} a_j(\mathbf{x})\right) P(\mathbf{x},t)\right] \\ &+ \frac{1}{2}\sum_{i=1}^{N}\sum_{i'=1}^{N} \frac{\partial^2}{\partial x_i \partial x_{i'}}\left[\left(\sum_{j=1}^{M} \nu_{ji}\nu_{ji'} a_j(\mathbf{x})\right) P(\mathbf{x},t)\right],\end{aligned} \quad (8)$$

which serves an approximation to the CME [Eq. (1)]. As Gillespie notes, this is exactly what one would get from truncating the Kramers-Moyal expansion of the CME at second order, so his derivation in some sense justifies the naive one.

The CLE [Eq. (7)], and the associated chemical Fokker-Planck equation [Eq. (8)] describing how the system's probability density will evolve in time, are not without problems. They generically predict negative concentrations [23] (although the hope is that the system has a negligibly small probability of occupying these states, and this is often borne out in practice), can be inaccurate for systems far from equilibrium [24], may not always exhibit multistability when the CME is multistable [25], and can give rise to nonphysical probability currents at equilibrium [26].

Despite these shortcomings, utilizing the CLE can help speed up simulations of CME dynamics when some species have large molecule numbers [27–30] or when there is a clear separation of timescales [31–33]. Moreover, alternative schemes like the deterministic reaction rate equations and the linear noise approximation [6] can profitably be viewed as approximations to the CLE [34], and moment-closure approximations have comparable accuracy [35].

The CLE, and Langevin equations more generally, have become standard approaches to modeling noisy gene regulation [36–40]. They have also been used to analyze noise-driven oscillations [41], model intracellular calcium dynamics [42–44], study ion-channel gating [45], and understand spiking neurons [46]. While only approximate, the CLE is unquestionably useful.

III. PATH-INTEGRAL FORMULATION OF CME DYNAMICS

Although path integrals [47] are most well-known in the context of quantum mechanics and quantum field theory [19,48–52], they have also proven useful for understanding classical stochastic phenomena like Brownian motion [53–55], conformational transitions [56–58], quantitative finance [50,59,60], population dynamics [61–65], neuron firing [66–72], gene regulation [36,73–77], and chemical kinetics [78–82].

In this section, we will develop a straightforward path-integral formulation of chemical master equation dynamics. Our path integral is constructed to closely resemble the formalism we used to describe SDE/Fokker-Planck dynamics in Ref. [15]. To our knowledge, it is original, although certain aspects also resemble the approach used by Lazarescu *et al.* [83]. The approach presented in this section is somewhat distinct from the often used Doi-Peliti approach [84–87], which involves integrating over so-called coherent states and yields integrals instead of sums.

A. States and operators

Our main objective is to solve the CME, Eq. (1). Instead of solving it directly, we will solve a related problem phrased in terms of states and operators in a certain Hilbert space; this allows us to construct a path integral just as one does in quantum mechanics.

Consider an infinite-dimensional Hilbert space spanned by the $|\mathbf{n}\rangle$ vectors (where $\mathbf{n} = (n_1, ..., n_N) \in \mathbb{N}^N$), in which an arbitrary state $|\phi\rangle$ is written

$$|\phi\rangle = \sum_{n_1=0}^{\infty} \cdots \sum_{n_N=0}^{\infty} c(\mathbf{n})|\mathbf{n}\rangle, \tag{9}$$

for some generally complex-valued coefficients $c(\mathbf{n})$. To ease notation, we will write

$$\sum_{\mathbf{n}} := \sum_{n_1=0}^{\infty} \cdots \sum_{n_N=0}^{\infty}, \tag{10}$$

so that an arbitrary state reads

$$|\phi\rangle = \sum_{\mathbf{n}} c(\mathbf{n})|\mathbf{n}\rangle. \tag{11}$$

Define an inner product in this space by

$$\langle \mathbf{m}|\mathbf{n}\rangle = \delta_{\mathbf{mn}}, \tag{12}$$

for all basis vectors $|\mathbf{m}\rangle$ and $|\mathbf{n}\rangle$, so that the inner product of two arbitrary states reads

$$\langle \phi_2|\phi_1\rangle = \sum_{\mathbf{n}} c_2^*(\mathbf{n})c_1(\mathbf{n}). \tag{13}$$

Using the inner product defined by Eqs. (12) and (13), we can show that there is a resolution of the identity

$$1 = \sum_{\mathbf{n}} |\mathbf{n}\rangle\langle \mathbf{n}|, \tag{14}$$

since $\langle \mathbf{n}|\phi\rangle = c(\mathbf{n})$. Define the state operators $\hat{n}_i$ by

$$\hat{n}_i|\mathbf{n}\rangle := n_i|\mathbf{n}\rangle, \tag{15}$$

for all $i = 1, ..., N$. We will associate any function $f(\mathbf{n}) = f(n_1, ..., n_N)$ with the operator $f(\hat{\mathbf{n}})$, whose action on a basis vector $|\mathbf{n}\rangle$ is

$$f(\hat{\mathbf{n}})|\mathbf{n}\rangle := f(\hat{n}_1, ..., \hat{n}_N)|\mathbf{n}\rangle, \tag{16}$$

where there is no operator ordering ambiguity because the $\hat{n}_i$ all commute with one another. Also define the propensity function operators $\hat{a}_j$ via

$$\hat{a}_j|\mathbf{n}\rangle := a_j(\boldsymbol{n})|\mathbf{n} + \boldsymbol{\nu}_j\rangle, \tag{17}$$

for all $j = 1, ..., M$, where $\boldsymbol{\nu}_j$ denotes the stoichiometry vector of the jth reaction.

B. Generating function and equation of motion

In the spirit of Peliti [86], define the generating function

$$|\psi(t)\rangle := \sum_{\mathbf{n}} P(\mathbf{n}, t)|\mathbf{n}\rangle, \tag{18}$$

where, as in the previous section, $P(\mathbf{n}, t)$ is the probability that the state of the system is $\mathbf{n} = (n_1, ..., n_N)$ at time t. Note that

$$\begin{aligned}\frac{\partial|\psi\rangle}{\partial t} &= \sum_{\mathbf{n}} \frac{\partial P(\mathbf{n}, t)}{\partial t}|\mathbf{n}\rangle \\ &= \sum_{\mathbf{n}} \left[\sum_{j=1}^{M} a_j(\mathbf{n} - \boldsymbol{\nu}_j)P(\mathbf{n} - \boldsymbol{\nu}_j, t) - a_j(\mathbf{n})P(\mathbf{n}, t)\right]|\mathbf{n}\rangle \\ &= \sum_{\mathbf{n}} \sum_{j=1}^{M} a_j(\mathbf{n})P(\mathbf{n}, t)|\mathbf{n} + \boldsymbol{\nu}_j\rangle \\ &\quad - \sum_{\mathbf{n}} \sum_{j=1}^{M} a_j(\mathbf{n})P(\mathbf{n}, t)|\mathbf{n}\rangle,\end{aligned} \tag{19}$$

where we have reindexed the left sum in the last step. Now we have

$$\begin{aligned}\frac{\partial|\psi\rangle}{\partial t} &= \sum_{\mathbf{n}} \left[\sum_{j=1}^{M} a_j(\mathbf{n})|\mathbf{n} + \boldsymbol{\nu}_j\rangle - a_j(\mathbf{n})|\mathbf{n}\rangle\right] P(\mathbf{n}, t) \\ &= \sum_{\mathbf{n}} \left[\sum_{j=1}^{M} \hat{a}_j|\mathbf{n}\rangle - a_j(\mathbf{n})\right] P(\mathbf{n}, t)|\mathbf{n}\rangle \\ &= \left[\sum_{j=1}^{M} \hat{a}_j - a_j(\hat{\mathbf{n}})\right] |\psi\rangle.\end{aligned} \tag{20}$$

If we define the operator

$$\hat{H} := \sum_{j=1}^{M} \hat{a}_j - a_j(\hat{\mathbf{n}}), \tag{21}$$

which we will call (in analogy with quantum mechanics) the Hamiltonian, then we can write the equation describing the time evolution of the generating function as

$$\frac{\partial|\psi\rangle}{\partial t} = \hat{H}|\psi\rangle. \tag{22}$$

It is this equation that we will solve instead of the CME; since $\langle \mathbf{n}|\psi(t)\rangle = P(\mathbf{n}, t)$, a solution to the CME can be extracted out of a solution to this equation.

C. Deriving the CME path integral

The formal solution to Eq. (22) is

$$|\psi(t_f)\rangle = e^{\hat{H}(t_f - t_0)}|\psi(t_0)\rangle. \tag{23}$$

At this point (following the usual procedure for deriving path integrals [15]), we write the length of time $(t_f - t_0)$ as $T\Delta t$ for some large number of time steps T, split the propagator into many pieces, and insert many resolutions of the identity:

$$\begin{aligned} |\psi(t_f)\rangle &= e^{\hat{H}\Delta t}\cdots e^{\hat{H}\Delta t}|\psi(t_0)\rangle \\ &= \sum_{\mathbf{n}_0}\cdots\sum_{\mathbf{n}_T}|\mathbf{n}_T\rangle\langle\mathbf{n}_T|e^{\hat{H}\Delta t}|\mathbf{n}_{T-1}\rangle\cdots\langle\mathbf{n}_1|e^{\hat{H}\Delta t}|\mathbf{n}_0\rangle \\ &\quad\times\langle\mathbf{n}_0||\psi(t_0)\rangle. \end{aligned} \tag{24}$$

We are specifically interested in the transition probability $P(\mathbf{n}_f, t_f; \mathbf{n}_0, t_0)$. To obtain an expression for it, note that if $|\psi(t_0)\rangle = |\mathbf{n}_0\rangle$, then $P(\mathbf{n}_f, t_f; \mathbf{n}_0, t_0) = \langle\mathbf{n}_f|\psi(t_f)\rangle$. Hence, we have

$$P(\mathbf{n}_f, t_f; \mathbf{n}_0, t_0) = \sum_{\mathbf{n}_1}\cdots\sum_{\mathbf{n}_{T-1}}\langle\mathbf{n}_T|e^{\hat{H}\Delta t}|\mathbf{n}_{T-1}\rangle\cdots\langle\mathbf{n}_1|e^{\hat{H}\Delta t}|\mathbf{n}_0\rangle, \tag{25}$$

where $\mathbf{n}_T = \mathbf{n}_f$. Now we just need to evaluate these matrix elements and put them together. Choose Δt sufficiently small so that

$$\begin{aligned} \langle\mathbf{n}_k|e^{\hat{H}\Delta t}|\mathbf{n}_{k-1}\rangle &\approx \langle\mathbf{n}_k|1 + \hat{H}\Delta t|\mathbf{n}_{k-1}\rangle \\ &= \delta_{\mathbf{n}_k,\mathbf{n}_{k-1}} + \langle\mathbf{n}_k|\hat{H}|\mathbf{n}_{k-1}\rangle\Delta t. \end{aligned} \tag{26}$$

We will take $\Delta t \to 0$ at the end of the calculation, so this equality will hold exactly. Using the specific form of $\hat{H}$ [Eq. (21)], we have

$$\begin{aligned} \langle\mathbf{n}_k|\hat{H}|\mathbf{n}_{k-1}\rangle &= \langle\mathbf{n}_k|\sum_{j=1}^{M}\hat{a}_j - a_j(\hat{\mathbf{n}})|\mathbf{n}_{k-1}\rangle \\ &= \sum_{j=1}^{M}\langle\mathbf{n}_k|\hat{a}_j - a_j(\hat{\mathbf{n}})|\mathbf{n}_{k-1}\rangle \\ &= \sum_{j=1}^{M}a_j(\mathbf{n}_{k-1})[\langle\mathbf{n}_k|\mathbf{n}_{k-1} + \boldsymbol{\nu}_j\rangle - \langle\mathbf{n}_k|\mathbf{n}_{k-1}\rangle] \\ &= \sum_{j=1}^{M}a_j(\mathbf{n}_{k-1})[\delta_{\mathbf{n}_k,\mathbf{n}_{k-1}+\boldsymbol{\nu}_j} - \delta_{\mathbf{n}_k,\mathbf{n}_{k-1}}]. \end{aligned} \tag{27}$$

Recall that the usual integral representation of the Dirac δ function reads

$$\delta_{\mathbf{m},\mathbf{n}} = \int\frac{d\mathbf{p}}{(2\pi)^N}\,e^{-i\mathbf{p}\cdot(\mathbf{m}-\mathbf{n})}, \tag{28}$$

where $d\mathbf{p} = dp_1\cdots dp_N$ and each p_i is integrated over the whole real line. Using this representation, $\langle\mathbf{n}_k|\hat{H}|\mathbf{n}_{k-1}\rangle$ becomes

$$\int\frac{d\mathbf{p}_k}{(2\pi)^N}\,e^{-i\mathbf{p}_k\cdot(\mathbf{n}_k-\mathbf{n}_{k-1})}\left\{\sum_{j=1}^{M}[e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j} - 1]a_j(\mathbf{n}_{k-1})\right\}, \tag{29}$$

where we have labeled the integration variable $\mathbf{p}_k$ to anticipate there being one integral for each matrix element in the final answer. Using Eq. (26), $\langle\mathbf{n}_k|e^{\hat{H}\Delta t}|\mathbf{n}_{k-1}\rangle$ is approximately equal to

$$\int\frac{d\mathbf{p}_k}{(2\pi)^N}\,e^{-i\mathbf{p}_k\cdot(\mathbf{n}_k-\mathbf{n}_{k-1})}\left\{1 + \Delta t\sum_{j=1}^{M}[e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j} - 1]a_j(\mathbf{n}_{k-1})\right\}. \tag{30}$$

Noting that Δt is small enough for the bracketed expression to be approximately equal to the corresponding exponential, our final expression for $\langle\mathbf{n}_k|e^{\hat{H}\Delta t}|\mathbf{n}_{k-1}\rangle$ becomes

$$\int\frac{d\mathbf{p}_k}{(2\pi)^N}\,e^{-i\mathbf{p}_k\cdot(\mathbf{n}_k-\mathbf{n}_{k-1}) + \Delta t\sum_{j=1}^{M}[\exp(i\mathbf{p}_k\cdot\boldsymbol{\nu}_j) - 1]a_j(\mathbf{n}_{k-1})}. \tag{31}$$

Using Eqs. (25) and (31), we find that $P(\mathbf{n}_f, t_f; \mathbf{n}_0, t_0)$ can be written as the path integral

$$\begin{aligned} P = \lim_{T\to\infty}\sum_{\mathbf{n}_1}\cdots\sum_{\mathbf{n}_{T-1}}\int\frac{d\mathbf{p}_1}{(2\pi)^N}\cdots\int\frac{d\mathbf{p}_T}{(2\pi)^N}\;\exp\Bigg\{\sum_{k=1}^{T} -i\mathbf{p}_k \\ \cdot(\mathbf{n}_k - \mathbf{n}_{k-1}) + \Delta t\sum_{j=1}^{M}[e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j} - 1]a_j(\mathbf{n}_{k-1})\Bigg\}, \end{aligned} \tag{32}$$

which resembles the Martin-Siggia-Rose-De Dominicis (MSRJD) path-integral description [88–92] of the Fokker-Planck equation. Again, while the Doi-Peliti path integral involves integrating over coherent states, this path integral involves integrating over every possible discrete path through $\mathbb{N}^N$ that goes from $\mathbf{n}_0$ to $\mathbf{n}_f$.

Although our primary interest in this paper is to use Eq. (32) to derive the CLE, this path-integral representation of the CME has utility in its own right. See Appendix A for how it can be used to exactly solve for time-dependent transition probabilities associated with simple chemical reaction systems.

IV. PATH-INTEGRAL DERIVATION OF THE CHEMICAL LANGEVIN EQUATION

In this section, we reinterpret Gillespie's derivation of the CLE in the context of stochastic path integrals and show explicitly how his two conditions translate in the path-integral context. Our central tool will be the Euler-Maclaurin formula [93,94], which allows one to approximate sums as integrals (plus correction terms). It says that

$$\sum_{n=a}^{b} f(n) \sim \int_a^b f(x)\,dx + \frac{f(b)+f(a)}{2} + \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!}[f^{(2k-1)}(b) - f^{(2k-1)}(a)], \quad (33)$$

where B_{2k} is the $(2k)$th Bernoulli number, and the "$\sim$" symbol is to indicate that we are to interpret the right-hand side as an asymptotic expansion (generically, the infinite sum may not be convergent, but retaining a finite number of terms still usually provides a good approximation to the left-hand side).

The Euler-Maclaurin formula is not an unfamiliar tool in chemical physics, given that it is often used to approximate partition functions [95–97] to good accuracy in certain regimes (e.g., the high temperature limit). It has also been used for other interesting purposes, like computing Fermi-Dirac integrals [98], and proving the asymptotic equivalence of two descriptions of Coulombic systems in certain potentials [99].

Roughly speaking, we will proceed as follows. Condition (i) will allow us to approximate each sum in Eq. (32) as an integral, and to argue that the correction terms are small; meanwhile, condition (ii) will allow us to Taylor expand the $\exp(i\mathbf{p}_k \cdot \boldsymbol{\nu}_j)$ terms in Eq. (32) to second order in the momenta $\mathbf{p}_k$. The result of these two approximations will be a MSRJD path integral, which we know from studies of stochastic path integrals [15] to be equivalent to a system of Langevin equations. In particular, it will be equivalent to Eq. (7), the CLE.

A. Only some paths satisfy Gillespie's conditions

Gillespie's first condition (see Sec. II) says that, in a period of time τ, the propensity functions do not change appreciably. Upon some reflection, we realize that this cannot be true for *all* possible trajectories the system might have, assuming the propensity functions have some state-dependence (which, in general, they do). In principle, it is possible that the number of molecules of some species jumps between 1 and 10^{100}, wildly and irregularly, so that there does not exist any timescale on which the propensity functions do not change appreciably. Indeed, all sorts of crazy trajectories are possible *in principle*—but they are overwhelmingly unlikely in practice.

While there certainly exist crazy and pathological paths for which it is hard or impossible to find a timescale τ that satisfies Gillespie's first condition, the requirement is not so stringent for most of the trajectories the system might take. In other words, we will suppose that the first condition is satisfied for the *dominant paths* rather than for all paths.

A similar argument applies to the second condition. This means that, in applying our two conditions, we will no longer be summing over all possible paths [cf. Eq. (32)]. Instead, we will be summing over all possible paths that satisfy the two conditions, a collection which we will assume includes the dominant or most likely paths.

If we are not summing over *all* possible paths, then what does our region of integration look like? To understand this, it is helpful to consider the simple case of a CME with one species and one reaction. Label the number of that species by n, and the propensity function of the single reaction by a.

Imagine starting the system in the state with n_0 molecules and thinking about where it will go (i.e., all possible states n_1) in the next time length τ. For the dominant paths, we assume that the difference $|a(n_1) - a(n_0)|$ is small, so that the propensity function did not change appreciably. But what do we mean by "appreciably"?

In a paper showing his two conditions hold in the thermodynamic limit [22], Gillespie assumed that his first condition meant

$$\frac{|a(n_1) - a(n_0)|}{a(n_0)} \ll 1, \quad (34)$$

i.e., that the change in the propensity function on the timescale τ is negligible compared to the size of its original value. This matches the intuition we have about what constitutes a negligible change in population size: for example, if the population size changed by 100 molecules, but the total number of molecules is on the order of 10^5, then we imagine that change not to be noticeable.

Here, we can be a little bit more precise than Eq. (34). We generally assume that our propensity functions are nicely behaved—in particular, that they are continuous, that they are infinitely differentiable, and that we may freely Taylor expand them. That is, we assume the a_j are analytic functions throughout our domain. Because most propensity functions of interest are polynomials (or at worst, rational functions like Hill functions), and because expressions like the Kramers-Moyal expansion already assume the a_j are smooth, these assumptions do not turn out to be particularly strong.

Suppose $a(n_0) > 0$, which is always true in the regime we care about, since we will usually need n sufficiently large; generic monomolecular and bimolecular propensity functions have zeros at $n = 0$ and $n = 1$. Because a is continuous, for any $\epsilon > 0$ we can find a $\delta > 0$ such that

$$|a(n_1) - a(n_0)| < \epsilon a(n_0), \quad (35)$$

provided $|n_1 - n_0| < \delta$. For the correction terms that arise from applying the Euler-Maclaurin formula [Eq. (33)] to be negligible, we also want to bound the derivatives of a in a similar fashion.

Analogous conditions apply in the general case, where the a_j may be functions of more than one variable. The moral of the story is that, because of the assumed smooth behavior of the propensity functions, we can find a region where they (and their derivatives) do not vary appreciably. In the simple one-dimensional case, this is an 'interval' $[n_0 - \delta_0^-, n_0 + \delta_0^+] \subseteq \mathbb{N}$ (where we let $\delta_0^- \neq \delta_0^+$ in general since we need $n_0 - \delta_0^-$ and $n_0 + \delta_0^+$ to both be natural numbers); in general, this is the intersection of an open set with a lattice: $U_0(\delta) \cap \mathbb{N}^N \subseteq \mathbb{N}^N$. For convenience, we will use U_0 to denote both the open set and its lattice intersection.

Hence, for a one-dimensional system, we restrict ourselves to paths

$$\sum_{n_1=0}^{\infty} \cdots \sum_{n_{N-1}=0}^{\infty} \to \sum_{n_1=n_0-\delta_0^-}^{n_0+\delta_0^+} \cdots \sum_{n_{N-1}=n_{N-2}-\delta_{N-2}^-}^{n_{N-2}+\delta_{N-2}^+}, \quad (36)$$

where the δ_i^+ and δ_i^-, as in the discussion above, are chosen so that the propensity functions and their derivatives vary within acceptable bounds. For an arbitrary CME, we restrict ourselves to paths

$$\sum_{\mathbf{n}_1}\cdots\sum_{\mathbf{n}_{T-1}} \to \sum_{\mathbf{n}_1\in U_0}\cdots\sum_{\mathbf{n}_{T-1}\in U_{T-2}}, \tag{37}$$

where the sets $U_i \subseteq \mathbb{N}^N$ are chosen similarly.

B. Coarse-graining time

There is another "philosophical" point we need to address. Earlier, we imagined breaking up the propagator into T time steps of length Δt, and choosing T to be large enough (or equivalently, Δt to be small enough) that each piece of the propagator was well-approximated by its first-order Taylor expansion [cf. Eq. (26)]. However, Gillespie's two conditions only apply on the "coarser" timescale τ. How do we go from time steps of size Δt to time steps of size τ in Eq. (32)?

There are two straightforward ways we can imagine. The simpler way is to say that, since we are in the business of making approximations *anyway*, we may as well make the approximation that Eq. (32) is valid on the timescale τ, and that the terms we neglected when Taylor expanding the propagator do not matter much in the regime where Gillespie's conditions apply.

But there is a more intellectually honest way to proceed. Suppose we originally broke the propagator into $S \cdot T$ time steps, for some natural number S large enough for our derivation to go through without issue. This means that the time step in our path integral has size $\Delta t := t/(S \cdot T)$. We want to rewrite our path integral in terms of a "macroscopic" timescale $\tau := t/T$, which corresponds to breaking up the overall time t into T time steps of length τ.

Schematically, this means we want to make the following identifications:

$$\begin{aligned}
\mathbf{n}_0 \xrightarrow{\Delta t} \mathbf{n}_1 \xrightarrow{\Delta t} \cdots \xrightarrow{\Delta t} \mathbf{n}_S \ &:\ \mathbf{n}_0 \xrightarrow{\tau} \mathbf{n}_1\\
\mathbf{n}_S \xrightarrow{\Delta t} \mathbf{n}_{S+1} \xrightarrow{\Delta t} \cdots \xrightarrow{\Delta t} \mathbf{n}_{2S} \ &:\ \mathbf{n}_1 \xrightarrow{\tau} \mathbf{n}_2\\
&\ \vdots\\
\mathbf{n}_{S\cdot(T-1)} \xrightarrow{\Delta t} \mathbf{n}_{S\cdot(T-1)+1} \xrightarrow{\Delta t} \cdots \xrightarrow{\Delta t} \mathbf{n}_{S\cdot T} \ &:\ \mathbf{n}_{T-1} \xrightarrow{\tau} \mathbf{n}_T.
\end{aligned} \tag{38}$$

The argument of the exponential in Eq. (32) reads

$$\sum_{k=1}^{S\cdot T} -i\mathbf{p}_k \cdot (\mathbf{n}_k - \mathbf{n}_{k-1}) + \Delta t \sum_{j=1}^{M} [e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j} - 1] a_j(\mathbf{n}_{k-1}). \tag{39}$$

Consider the following small piece of this expression:

$$\sum_{k=1}^{S\cdot T} [e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j} - 1] a_j(\mathbf{n}_{k-1}). \tag{40}$$

Assuming (on the dominant paths) that the propensity function a_j only changes appreciably on the timescale $\tau = S\Delta t$, we can make the approximation that

$$\begin{gathered}
a_j(\mathbf{n}_0) \approx a_j(\mathbf{n}_1) \approx \cdots \approx a_j(\mathbf{n}_{S-1})\\
a_j(\mathbf{n}_S) \approx a_j(\mathbf{n}_{S+1}) \approx \cdots \approx a_j(\mathbf{n}_{2S-1})\\
\vdots\\
a_j(\mathbf{n}_{S\cdot(T-1)}) \approx a_j(\mathbf{n}_{S\cdot(T-1)+1}) \approx \cdots \approx a_j(\mathbf{n}_{S\cdot T-1})
\end{gathered} \tag{41}$$

and rewrite Eq. (40) in terms of $a_j(\mathbf{n}_0)$, $a_j(\mathbf{n}_S)$, $a_j(\mathbf{n}_{2S})$, ..., $a_j(\mathbf{n}_{S\cdot T})$ only. This means that the only places the "intermediate" time steps (e.g., $\mathbf{n}_1$, ..., $\mathbf{n}_{S-1}$, or $\mathbf{n}_{S+1}$, ..., $\mathbf{n}_{2S-1}$) will appear are in the piece that reads

$$\sum_{k=1}^{S\cdot T} -i\mathbf{p}_k \cdot (\mathbf{n}_k - \mathbf{n}_{k-1}). \tag{42}$$

Happily, this means that all of the intermediate time steps can be summed over. For example,

$$\begin{aligned}
&\sum_{\mathbf{n}_1}\cdots\sum_{\mathbf{n}_{S-1}} \exp\left\{\sum_{k=1}^{S} -i\mathbf{p}_k \cdot (\mathbf{n}_k - \mathbf{n}_{k-1})\right\}\\
&\quad\approx \delta(\mathbf{p}_1 - \mathbf{p}_2)\delta(\mathbf{p}_2 - \mathbf{p}_3)\cdots\delta(\mathbf{p}_{S-1} - \mathbf{p}_S),
\end{aligned} \tag{43}$$

where the right-hand side is approximate because, due to our restriction of the sum domain in the previous section, the sum representation of the Dirac δ function

$$\frac{1}{(2\pi)^N}\sum_{\mathbf{n}} \exp\{-i\mathbf{n}\cdot(\mathbf{p} - \mathbf{p}')\} = \delta(\mathbf{p} - \mathbf{p}') \tag{44}$$

only approximately applies. After summing over all intermediate time steps and integrating out extraneous $\mathbf{p}_k$ using the δ functions that appear, Eq. (40) reads

$$\begin{aligned}
&\sum_{k=1}^{T} -i\mathbf{p}_k \cdot (\mathbf{n}_k - \mathbf{n}_{k-1}) + S\Delta t \sum_{j=1}^{M} [e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j} - 1] a_j(\mathbf{n}_{k-1})\\
&\quad= \sum_{k=1}^{T} -i\mathbf{p}_k \cdot (\mathbf{n}_k - \mathbf{n}_{k-1}) + \tau \sum_{j=1}^{M} [e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j} - 1] a_j(\mathbf{n}_{k-1}).
\end{aligned} \tag{45}$$

Hence, using Gillespie's first condition, we have successfully gone from a path integral with timescale Δt to a path integral with a "coarser" timescale τ.

C. Applying condition 1

In this section, we will apply condition (i) to convert the sums in Eq. (32) to integrals. After restricting our domain to the dominant paths (see Sec. IV A) and coarse-graining time (see Sec. IV B), our current path-integral description of the CME reads

$$P \approx \sum_{\mathbf{n}_1\in U_0}\cdots\sum_{\mathbf{n}_{T-1}\in U_{T-2}} \int \frac{d\mathbf{p}_1}{(2\pi)^N}\cdots\int \frac{d\mathbf{p}_T}{(2\pi)^N} \exp\{-S\tau\}, \tag{46}$$

where we recall that the sets U_0, ..., U_{T-2} cover all trajectories on which Gillespie's two conditions apply, and where we have defined the function (which we can call the "action," in analogy with quantum mechanics)

$$S := \sum_{k=1}^{T} i\mathbf{p}_k \cdot \left(\frac{\mathbf{n}_k - \mathbf{n}_{k-1}}{\tau}\right) - \sum_{j=1}^{M} [e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j} - 1]a_j(\mathbf{n}_{k-1}) \tag{47}$$

to ease notation. We will proceed using the Euler-Maclaurin formula [Eq. (33)]. As a starting point, consider Eq. (46) in one-dimension:

$$P \approx \sum_{n_1=n_0-\delta_0^-}^{n_0+\delta_0^+} \cdots \sum_{n_{N-1}=n_{N-2}-\delta_{N-2}^-}^{n_{N-2}+\delta_{N-2}^+} \int \frac{dp_1}{2\pi} \cdots \int \frac{dp_T}{2\pi} \exp\{-S\tau\}, \tag{48}$$

where the δ_i^- and δ_i^+ are as described in Sec. IV A. Using the Euler-Maclaurin formula, we have

$$\sum_{n_1=n_0-\delta_0^-}^{n_0+\delta_0^+} \exp\{-S\tau\} \approx \int_{n_0-\delta_0^-}^{n_0+\delta_0^+} \exp\{-S\tau\}\, dn_1 + \frac{e^{-S(n_0+\delta_0^+)\tau} + e^{-S(n_0-\delta_0^-)\tau}}{2} + \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!} \frac{d^{2k-1}}{dn_1^{2k-1}} [e^{-S\tau}]_{n_0-\delta_0^-}^{n_0+\delta_0^+}. \tag{49}$$

Now we need to argue that the correction terms can safely be neglected. Define $\delta := \delta_0^+ + \delta_0^-$. Because the propensity functions don't change vary much in the interval $[n_0 - \delta_0^-, n_0 + \delta_0^+]$ (by Gillespie's first condition), the integral term is roughly

$$\exp\{-S(n_0)\tau\}\delta. \tag{50}$$

Meanwhile, the next term is roughly

$$\exp\{-S(n_0)\tau\}, \tag{51}$$

which should be negligible compared to the first as long as $\delta \gg 1$. This should certainly be true; if $\delta \sim 1$, then our conditions are either too strict or we are in a regime with too small molecule numbers.

Because the propensity functions a_j do not change much (and because the a_j are nicely behaved, usually monotonic functions in the regime we care about), they are approximately "flat." This means that their derivatives $a_j^{(2k-1)}$ are small. For example,

$$\frac{d}{dn_1}[e^{-S\tau}] = \left[i(p_2 - p_1) + \tau \sum_{j=1}^{M} [e^{ip_k\nu_j} - 1]a_j' \right] e^{-S\tau}, \tag{52}$$

so

$$\begin{aligned} \frac{d}{dn_1}[e^{-S\tau}]_{n_0-\delta_0^-}^{n_0+\delta_0^+} &= \tau \sum_{j=1}^{M} [e^{ip_k\nu_j} - 1][a_j'(n_0 + \delta_0^+)e^{-S(n_0+\delta_0^+)\tau} \\ &\quad - a_j'(n_0 - \delta_0^-)e^{-S(n_0-\delta_0^-)\tau}] \\ &\approx \tau e^{-S(n_0)\tau} \sum_{j=1}^{M} [e^{ip_k\nu_j} - 1][a_j'(n_0 + \delta_0^+) \\ &\quad - a_j'(n_0 - \delta_0^-)] \\ &\approx 0. \end{aligned} \tag{53}$$

In summary, we have

$$\sum_{n_1=n_0-\delta_0^-}^{n_0+\delta_0^+} \exp\{-S\tau\} \approx \int_{n_0-\delta_0^-}^{n_0+\delta_0^+} \exp\{-S\tau\}\, dn_1, \tag{54}$$

which means we have successfully converted a sum into an integral. Apply this argument many more times to obtain

$$\begin{aligned} &\sum_{n_1=n_0-\delta_0^-}^{n_0+\delta_0^+} \cdots \sum_{n_{T-1}=n_{T-2}-\delta_{T-2}^-}^{n_{T-2}+\delta_{T-2}^+} \exp\{-S\tau\} \\ &\approx \int_{n_0-\delta_0^-}^{n_0+\delta_0^+} dn_1 \cdots \int_{n_{T-2}-\delta_{T-2}^-}^{n_{T-2}+\delta_{T-2}^+} dn_{T-1}\, \exp\{-S\tau\}. \end{aligned} \tag{55}$$

A similar argument applies to the N species path integral [Eq. (46)]; the only difference is that the Euler-Maclaurin formula must be applied N times for each time step, because we would like to convert N sums to an N-variable integral.

Alternatively, one can argue using the appropriate many sum generalization of the Euler-Maclaurin formula [Eq. (33)]. There is some literature on generalizations of it to sums over polytopes [100–102] (schematically, shapes in N-dimensional space whose vertices we can imagine as living in $\mathbb{Z}^N$). The main challenge for this approach would be to show that satisfying Gillespie's first condition corresponds to satisfying the requirements associated with the approximation being accurate (which are somewhat more technical than those for the single sum Euler-Maclaurin formula).

The end result of all this is

$$P \approx \int_{U_0} d\mathbf{x}_1 \cdots \int_{U_{T-2}} d\mathbf{x}_{T-1} \int \frac{d\mathbf{p}_1}{(2\pi)^N} \cdots \int \frac{d\mathbf{p}_T}{(2\pi)^N} \exp\{-S\tau\}, \tag{56}$$

where we have relabeled each $\mathbf{n}_k$ as $\mathbf{x}_k$ to (as in Sec. II) emphasize that we are now working with continuous variables. We remark that, if not for the bounds, we would have a Kramers-Moyal path integral (see Sec. V of our earlier paper [15]).

D. Applying condition 2

Consider the terms in the action S [Eq. (47)] that look like

$$[e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j} - 1]a_j(\mathbf{x}_{k-1})\tau. \tag{57}$$

Condition (ii) tells us that, for the dominant paths, $a_j(\mathbf{x}_{k-1})\tau \gg 1$. In particular, we will assume that it is *so* large that Taylor expanding the term it is multiplied by will have a negligible effect on the overall value,[1] i.e.,

$$\begin{aligned} &[e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j} - 1]a_j(\mathbf{x}_{k-1})\tau \\ &\approx \left[i\mathbf{p}_k \cdot \boldsymbol{\nu}_j - \frac{1}{2}\sum_{\ell=1}^{N}\sum_{\ell'=1}^{N} p_k^{\ell} p_k^{\ell'} \nu_{j\ell}\nu_{j\ell'} \right] a_j(\mathbf{x}_{k-1})\tau, \end{aligned} \tag{58}$$

---

[1] See Appendix B for a somewhat more rigorous argument justifying this approximation.

where p_k^ℓ is the ℓ th component of the vector $\mathbf{p}_k$. Thus, we finally obtain

$$S \approx \sum_{k=1}^{T} i\mathbf{p}_k \cdot \left[\frac{\mathbf{x}_k - \mathbf{x}_{k-1}}{\tau} - \sum_{j=1}^{M} \boldsymbol{\nu}_j a_j(\mathbf{x}_{k-1}) \right] + \frac{1}{2} \sum_{\ell=1}^{N} \sum_{\ell'=1}^{N} \sum_{j=1}^{M} p_k^\ell p_k^{\ell'} \nu_{j\ell} \nu_{j\ell'} a_j(\mathbf{x}_{k-1}), \tag{59}$$

which looks just like the action for the MSRJD path integral (see Sec. V of Ref. [15]) corresponding to the chemical Fokker-Planck equation [Eq. (8)]. Our final result for the whole path integral reads

$$P \approx \int_{U_0} d\mathbf{x}_1 \cdots \int_{U_{T-2}} d\mathbf{x}_{T-1} \int \frac{d\mathbf{p}_1}{(2\pi)^N} \cdots \int \frac{d\mathbf{p}_T}{(2\pi)^N} \times \exp\left\{ -\sum_{k=1}^{T} \left[i\mathbf{p}_k \cdot \left(\frac{\mathbf{x}_k - \mathbf{x}_{k-1}}{\tau} - \sum_{j=1}^{M} \boldsymbol{\nu}_j a_j(\mathbf{x}_{k-1}) \right) + \frac{1}{2} \sum_{\ell=1}^{N} \sum_{\ell'=1}^{N} \sum_{j=1}^{M} p_k^\ell p_k^{\ell'} \nu_{j\ell} \nu_{j\ell'} a_j(\mathbf{x}_{k-1}) \right] \tau \right\}, \tag{60}$$

which looks like the usual Fokker-Planck path integral but with restricted integration bounds.

The appearance of Eq. (60) can be compacted somewhat if we define the diffusion tensor $D_{\ell\ell'}$:

$$D_{\ell\ell'}(\mathbf{x}) := \frac{1}{2} \sum_{j=1}^{M} \nu_{j\ell} \nu_{j\ell'} a_j(\mathbf{x}). \tag{61}$$

At the CLE/Fokker-Planck level, the diffusion tensor captures all information about a system's noise. It must be positive semidefinite for the Fokker-Planck equation and its corresponding path integral to make sense [103,104].

Finish the derivation by enlarging our integration domain as much as possible (while keeping the diffusion tensor positive semidefinite), assuming that permitting these additional paths does not substantially contribute to transition probabilities, since they were small enough to neglect in the first place. In general, we do not expect that the appropriate domain for our new continuous variables will be $[0, \infty)^N$, despite the fact that our original domain was $\mathbb{N}^N$. For example, the chemical Langevin equation [1] for the birth-death process (with birth rate k, death rate γ, and steady state mean $\mu := k/\gamma$) reads

$$\dot{x} = k - \gamma x + \sqrt{k + \gamma x}\, \eta(t) \tag{62}$$

and is naturally defined on $[-\mu, \infty)$, because there is always some nonzero probability that the noise term will push the system into negative concentrations while its magnitude is greater than or equal to zero, i.e., when $k + \gamma x = \gamma(\mu + x) \geqslant 0$.

E. Comparison with the system volume approach

We have shown in the previous few sections how Gillespie's derivation works in a path-integral context. Because Gillespie himself [1] compared his approach to ones which rely upon the largeness of the system volume Ω, it is interesting to do that here also. Let us translate the typical system volume approach into path-integral language and see how it compares with the approach we described earlier.

Consider again a CME with N species and M reactions [Eq. (1)], but this time with the additional physical context that the chemicals interact inside a very large volume Ω. Suppose we rewrite the CME in terms of concentration variables $x_i := n_i/\Omega$ for all $i = 1, \ldots, N$. The change in variables will lead to the probability density function $P(\mathbf{n}, t)$ increasing by a factor of Ω^N:

$$\begin{aligned} P(\mathbf{n}, t) d\mathbf{n} &= \Omega^N P(\mathbf{n}, t) d\mathbf{x} = P(\mathbf{x}, t) d\mathbf{x} \\ \Rightarrow\ P(\mathbf{x}, t) &= \Omega^N P(\mathbf{n}, t). \end{aligned} \tag{63}$$

Gillespie used rigorous microphysical arguments [1,105–107] to show that the volume-dependence of the propensity functions for monomolecular, bimolecular, and trimolecular reactions goes like

$$a_j(\mathbf{n}) = \Omega\, \tilde{a}_j(\mathbf{x}), \tag{64}$$

where the adjusted propensity functions $\tilde{a}_j$ are volume-independent. Using Eqs. (63) and (64), our original CME path integral [Eq. (32)] can be rewritten as

$$P = \lim_{T\to\infty} \Omega^N \sum_{\mathbf{n}_1} \cdots \sum_{\mathbf{n}_{T-1}} \int \frac{d\mathbf{p}_1}{(2\pi)^N} \cdots \int \frac{d\mathbf{p}_T}{(2\pi)^N} \times \exp\left\{ \sum_{k=1}^{T} -i\Omega\, \mathbf{p}_k \cdot (\mathbf{x}_k - \mathbf{x}_{k-1}) + \Omega \Delta t \sum_{j=1}^{M} [e^{i\mathbf{p}_k \cdot \boldsymbol{\nu}_j} - 1] \tilde{a}_j(\mathbf{x}_{k-1}) \right\}. \tag{65}$$

Now add in $T-1$ factors of Ω^N/Ω^N:

$$P = \lim_{T\to\infty} \left[\frac{1}{\Omega^N} \sum_{\mathbf{n}_1} \right] \cdots \left[\frac{1}{\Omega^N} \sum_{\mathbf{n}_{T-1}} \right] \times \int \left(\frac{\Omega}{2\pi} \right)^N d\mathbf{p}_1 \cdots \int \left(\frac{\Omega}{2\pi} \right)^N d\mathbf{p}_T \times \exp\left\{ \sum_{k=1}^{T} -i\Omega\, \mathbf{p}_k \cdot (\mathbf{x}_k - \mathbf{x}_{k-1}) + \Omega \Delta t \sum_{j=1}^{M} [e^{i\mathbf{p}_k \cdot \boldsymbol{\nu}_j} - 1] \tilde{a}_j(\mathbf{x}_{k-1}) \right\}. \tag{66}$$

Riemann sums will play the role that the Euler-Maclaurin formula did (i.e., converting sums to integrals) in our earlier derivation. Recall that the (right endpoint) Riemann sum for a function f on $[0, b]$ reads [108]

$$\int_0^b f(x)dx \approx \sum_{i=0}^{N} f(i\Delta x)\Delta x, \tag{67}$$

where $\Delta x = b/N$. If we take $b \to \infty$ and $N \to \infty$ in such a way that Δx remains constant, then we can write

$$\int_0^\infty f(x) dx \approx \sum_{i=0}^\infty f(i\Delta x)\Delta x. \tag{68}$$

The corresponding N-dimensional result is

$$\int_0^\infty dx_1 \cdots \int_0^\infty dx_N\, f(x) \approx \sum_{i_1=0}^\infty \cdots \sum_{i_N=0}^\infty f(i_1\Delta x, \cdots, i_N\Delta x)(\Delta x)^N. \tag{69}$$

Since the inverse system volume $1/\Omega$ seems to play the role of Δx in Eq. (66), we can use this Riemann sum result to approximate each sum as

$$\frac{1}{\Omega^N}\sum_{\mathbf{n}} \approx \int_0^\infty dx_1 \cdots \int_0^\infty dx_N, \tag{70}$$

so that our path integral is now

$$\begin{aligned} P = \lim_{T\to\infty} &\int d\mathbf{x}_1 \cdots \int d\mathbf{x}_{T-1} \\ &\times \int \left(\frac{\Omega}{2\pi}\right)^N d\mathbf{p}_1 \cdots \int \left(\frac{\Omega}{2\pi}\right)^N d\mathbf{p}_T \\ &\times \exp\left\{\sum_{k=1}^T -i\Omega\, \mathbf{p}_k\cdot(\mathbf{x}_k - \mathbf{x}_{k-1}) \right. \\ &\left. + \Omega\Delta t \sum_{j=1}^M [e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j} - 1]\tilde{a}_j(\mathbf{x}_{k-1})\right\}. \end{aligned} \tag{71}$$

Now we can argue just as we did in Sec. IV D. Because we are taking Ω to be extraordinarily large in the thermodynamic limit,

$$\begin{aligned} &\Omega[e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j} - 1]\tilde{a}_j(\mathbf{x}_{k-1})\tau \\ &\approx \Omega\left[i\mathbf{p}_k\cdot\boldsymbol{\nu}_j - \frac{1}{2}\sum_{\ell=1}^N\sum_{\ell'=1}^N p_k^\ell p_k^{\ell'} \nu_{j\ell}\nu_{j\ell'}\right]\tilde{a}_j(\mathbf{x}_{k-1})\tau, \end{aligned} \tag{72}$$

i.e., Ω is so large that above term does not change much in value when Taylor expanded to second order in $\mathbf{p}_k$. Finally, we have

$$\begin{aligned} P = \lim_{T\to\infty} &\int d\mathbf{x}_1 \cdots \int d\mathbf{x}_{T-1} \\ &\times \int \left(\frac{\Omega}{2\pi}\right)^N d\mathbf{p}_1 \cdots \int \left(\frac{\Omega}{2\pi}\right)^N d\mathbf{p}_T \\ &\times \exp\left\{-\Omega\sum_{k=1}^T \left[i\mathbf{p}_k\cdot\left(\frac{\mathbf{x}_k - \mathbf{x}_{k-1}}{\Delta t} - \sum_{j=1}^M \boldsymbol{\nu}_j\tilde{a}_j(\mathbf{x}_{k-1})\right)\right.\right. \\ &\left.\left. + \frac{1}{2}\sum_{\ell=1}^N\sum_{\ell'=1}^N\sum_{j=1}^M p_k^\ell p_k^{\ell'}\nu_{j\ell}\nu_{j\ell'}\tilde{a}_j(\mathbf{x}_{k-1})\right]\Delta t\right\}, \end{aligned} \tag{73}$$

which is the same as the result from Sec. IV D [cf. Eq. (59)] but with additional factors of Ω. It also exactly matches the system volume MSRJD path integral for the Fokker-Planck equation (cf. Eq. (94) in Sec. V of Ref. [15]). In other words, we have indeed derived a path-integral equivalent to a set of Langevin equations/a Fokker-Planck equation. Moreover, it is equivalent to the *same* set of Langevin equations that Eq. (59) is (as is easily seen after changing back from concentration variables to the original number variables)—although the integration bounds on the path integral are different here.

Given that this approach was significantly simpler (in both a technical and conceptual sense), why bother with Gillespie's derivation? There are a few good reasons.

- The approximations provided by Eqs. (68) and (69) are more mathematically dubious than the Euler-Maclaurin formula [Eq. (33)], which is well-studied and has precisely expressed error bounds.
- The thermodynamic limit may not apply to most biochemical systems of interest, given that molecule numbers are often large but not overwhelmingly so, and that the system volume (for example, of a cell) is not large enough to prevent crowding [109–112] and boundary effects [113–115] from being important.
- The system volume approach only *applies* when our CME describes a well-stirred, dilute mix of chemicals held at fixed temperature in a very large box—but the CLE is known to be a useful approximate description of all sorts of other stochastic systems (e.g., spiking neurons, fluctuating population dynamics models, stock options). In these other situations, there is no clear notion of a control parameter analogous to Ω.
- The system volume approach misses the subtlety of the integration bounds associated with the chemical Langevin/chemical Fokker-Planck equations; as we pointed out at the end of the previous section, it is a nontrivial issue that the domain of the approximating CLE will generally not be $[0, \infty)^N$.

V. DISCUSSION

We constructed an original path-integral description of the CME, and applied Gillespie's conditions (suitably interpreted) to it to derive a path integral known to be equivalent to the CLE. In some sense, the difference between the system size approach and Gillespie's approach to deriving the CLE is the difference between approximating sums as integrals via Riemann sums, and via the Euler-Maclaurin formula. As discussed at the end of the previous section, while both approximation techniques can be valid in the appropriate circumstances, the Euler-Maclaurin formula is more generally applicable and has better characterized correction terms.

It is interesting to note that, although we began with an exact path integral that involved taking the limit $\Delta t \to 0$ [Eq. (32)], we coarse-grained time to end up with a path integral with fixed time step τ that does not get taken to zero [Eq. (60)]. This leads to another sense in which the CLE is only an approximate description, since a true CLE/Fokker-Planck path integral (see Ref. [15]) also involves taking the limit $\Delta t \to 0$. However, the idea of a "macroscopic" timescale was addressed by Gillespie himself in his original paper [1]. There, he offered an analogy to current in an electric circuit: We can freely write and manipulate the derivative $I := dq/dt$, and think about the limit $dt \to 0$, *provided* we understand that we are not taking it to be *so* small that shot noise effects start to matter.

Because our argument applied to each reaction/propensity function separately, it can in principle be used construct path integrals for hybrid systems. In other words, just as Harris *et al.* [28–30] do, we can suppose that Gillespie's two conditions apply only to a subset of all reactions or species, and construct a path integral in which some species/reactions are treated CLE-style, while others are treated CME-style. Indeed, there should be a path-integral way to view all of the hybrid constructions—based on molecule numbers or separations of timescales—referenced in Sec. II. These path integrals could then be used to extract large deviation results.

It is unclear if Gillespie's conditions could be applied to the Doi-Peliti path integral [84–87] to recover a CLE-equivalent path integral. Part of the difficulty is that the Doi-Peliti construction involves integrating over coherent states, which contribute integrals over the whole real line in the expression for the propagator [86]; it is not necessarily straightforward to associate these with sums or integrals over state space.

Although we only used the path-integral representation of the CME constructed in Sec. III to derive the CLE, it can also be used as a tool in its own right for the same purposes other path-integral representations are often used for: namely, finding exact solutions [116–118], constructing asymptotic or perturbative approximations to transition probabilities and moments [119], computing least action paths associated with particular state transitions, and enabling a variational method for numerically computing transition probabilities and least action paths [70]. In cases where Gillespie's first condition applies, and the domain of the path integral can be restricted, the numerical evaluation of Eq. (32) becomes even easier. Though our argument does not offer a constructive prescription for the restricted integration domains $U_0, ..., U_{T-2}$, one can in principle bootstrap the path integral by running Gillespie algorithm simulations beforehand to estimate reasonable domains.

VI. CONCLUSION

The chemical Langevin equation is usually derived using Gillespie's two conditions, or large system volume arguments; as we described, both methods have clear path-integral analogues. Our results suggest that path integrals offer a useful and mathematically precise way of thinking about the relationship between different levels of approximation (e.g., CME and CLE), and about coarse-graining biochemical models more generally.

ACKNOWLEDGMENT

This work was supported by NSF Grant No. DMS 1562078.

APPENDIX A: SAMPLE PATH-INTEGRAL CALCULATIONS

One can get a feel for a given path integral by using it to exactly calculate transition probabilities for simple systems. In this Appendix, we use the path integral described by Eq. (32) to compute exact time-dependent transition probabilities for a (i) pure birth process, a (ii) pure death process, and a (iii) chemical birth-death process. For more information on the chemical birth-death process, and for analogous path-integral calculations that are valid in the limit where molecule numbers can be treated as continuous, see Refs. [116,117].

These systems can also be solved using the Doi-Peliti path integral, the most commonly used path-integral description of the CME—see Vastola [118] for a guide. It is interesting to note that the transition probability derivations presented in this Appendix are simpler in some ways; for example, one does not need to introduce coherent states or special scalar products, and transition probabilities can be computed without computing the generating function first.

The following path-integral calculations involve the evaluation of many contour integrals. The resolution of identity introduced in Eq. (28), which leads to the "momentum" integrals going from negative infinity to positive infinity, yields contour integrals with an infinite number of poles. To make these calculations somewhat simpler, we instead use the equally valid resolution of the identity given by

$$\delta(n_1 - n_2) = \int_{-\pi}^{\pi} \frac{dp}{2\pi}\, e^{-ip(n_1-n_2)} \tag{A1}$$

in one dimension, and

$$\delta(\mathbf{n}_1 - \mathbf{n}_2) = \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \frac{dp_1 \cdots dp_N}{(2\pi)^N}\, e^{-i\mathbf{p}\cdot(\mathbf{n}_1-\mathbf{n}_2)} \tag{A2}$$

in more than one dimension. This means that our CME path integral is still given by Eq. (32), but with the $\mathbf{p}_k$ integrals all going from $-\pi$ to π.

1. The pure birth process

The pure birth process models a species that is randomly created at some rate. It is characterized by the chemical reaction

$$\varnothing \xrightarrow{k} X, \tag{A3}$$

where k is the birth rate. This reaction corresponds to the CME

$$\frac{\partial P(n,t)}{\partial t} = k[P(n-1,t) - P(n,t)], \tag{A4}$$

where $P(n,t)$ is the probability that the system has n X molecules at time t (with $n \in \{0, 1, 2, ...\}$). For this system, our path integral for the transition probability $P(n_f, t; n_0, 0)$ [cf. Eq. (32)] reads

$$P = \lim_{T\to\infty} \sum_{n_1} \cdots \sum_{n_{T-1}} \int \frac{dp_1}{2\pi} \cdots \int \frac{dp_T}{2\pi} \times \exp\left\{ \sum_{\ell=1}^{T} -ip_\ell(n_\ell - n_{\ell-1}) + k\Delta t(e^{ip_\ell} - 1) \right\}, \tag{A5}$$

where $\Delta t := t/T$. First, since

$$-k\Delta t \sum_{\ell=1}^{T} = -kt, \tag{A6}$$

the overall answer has a prefactor $\exp(-kt)$. Next, organizing terms by the n_ℓ, we have

$$\begin{aligned} &ip_1 n_0 \\ &\quad - ip_T n_f + in_1(p_2 - p_1) + in_2(p_3 - p_2) \\ &\quad + \cdots + in_{T-1}(p_T - p_{T-1}). \end{aligned} \tag{A7}$$

Each sum over n_ℓ (for $\ell = 1, ..., T-1$) just corresponds to summing a geometric series:

$$\sum_{n_\ell=0}^{\infty} [e^{i(p_{\ell+1}-p_\ell)}]^{n_\ell} = \frac{1}{1 - e^{i(p_{\ell+1}-p_\ell)}}. \tag{A8}$$

Now consider the integral over p_1, which reads

$$\int_{-\pi}^{\pi} \frac{dp_1}{2\pi} \frac{e^{ip_1 n_0 + k\Delta t e^{ip_1}}}{1 - e^{i(p_2 - p_1)}}. \tag{A9}$$

Switching variables to $z = e^{ip_1}$, we have the contour integral

$$\oint \frac{dz}{2\pi i} \frac{z^{n_0} e^{k\Delta t z}}{z - e^{ip_2}}, \tag{A10}$$

whose integrand has a simple pole at e^{ip_2}. Although Cauchy's theorem does not technically apply in this case, since the pole lies on the circular contour, we can imagine deforming the contour or integrand slightly [e.g., by taking $e^{ip_2} \to (1-\epsilon)e^{ip_2}$] so that Cauchy's theorem does apply, allowing us to evaluate the integral in the usual way.[2] Doing so, we obtain

$$e^{ip_2 n_0 + k\Delta t e^{ip_2}}, \tag{A11}$$

which indicates that the overall effect of doing the contour integral was to implement the constraint that $p_1 = p_2$. Similarly, the overall effect of doing the integral over p_ℓ (for $\ell = 2, ..., T-1$), which schematically reads

$$\int_{-\pi}^{\pi} \frac{dp_\ell}{2\pi} \frac{f(p_\ell)}{1 - e^{i(p_{\ell+1}-p_\ell)}}, \tag{A12}$$

for some function $f(p_\ell)$, is to implement the constraint that $p_\ell = p_{\ell+1}$. Hence, after doing the integrals over $p_1, ..., p_{T-1}$, we have

$$\int_{-\pi}^{\pi} \frac{dp_T}{2\pi}\, e^{-ip_T(n_f - n_0) + kte^{ip_T}}. \tag{A13}$$

Define $\Delta n := n_f - n_0$ and change variables to $z = e^{ip_T}$. Then we have

$$\oint \frac{dz}{2\pi i} \frac{e^{ktz}}{z^{\Delta n + 1}}. \tag{A14}$$

For $\Delta n < 0$, this integral has no poles, so $P(n_f, t; n_0, 0) = 0$ in that case. For $\Delta n \geqslant 0$, a standard application of Cauchy's integral formula yields that the result is

$$\frac{1}{\Delta n!} \frac{d^{\Delta n}}{dz^{\Delta n}} [e^{ktz}]_{z=0} = \frac{(kt)^{\Delta n}}{\Delta n!}. \tag{A15}$$

Hence, including the prefactor e^{-kt}, our result for the transition probability is

$$P(n_f, t; n_0, 0) = \frac{(kt)^{\Delta n} e^{-kt}}{\Delta n!}, \tag{A16}$$

i.e., a Poisson distribution.

2. The pure death process

The pure death process models a species that randomly degrades at some rate. It is characterized by the chemical reaction

$$X \xrightarrow{\gamma} \varnothing, \tag{A17}$$

where γ is the death rate. This reaction corresponds to the CME

$$\frac{\partial P(n,t)}{\partial t} = \gamma[(n+1)P(n+1,t) - nP(n,t)], \tag{A18}$$

where $P(n,t)$ is the probability that the system has n X molecules at time t (with $n \in \{0, 1, 2, ...\}$). This time, our path integral for the transition probability $P(n_f, t; n_0, 0)$ reads

$$\begin{aligned} P = \lim_{T\to\infty} \sum_{n_1} \cdots \sum_{n_{T-1}} \int \frac{dp_1}{2\pi} \cdots \int \frac{dp_T}{2\pi} \\ \times \exp\left\{ \sum_{\ell=1}^{T} -ip_\ell(n_\ell - n_{\ell-1}) + \gamma\Delta t\, n_{\ell-1}(e^{-ip_\ell} - 1) \right\}. \end{aligned} \tag{A19}$$

The terms that involve n_ℓ (for some $\ell = 1, ..., T-1$) look like

$$in_\ell[p_{\ell+1} - p_\ell - i\gamma\Delta t(e^{-ip_{\ell+1}} - 1)], \tag{A20}$$

so the sum over n_ℓ yields

$$\begin{aligned} &\sum_{n_\ell=0}^{\infty} [e^{i[p_{\ell+1} - p_\ell - i\gamma\Delta t(e^{-ip_{\ell+1}} - 1)]}]^{n_\ell} \\ &\quad = \frac{1}{1 - e^{i[p_{\ell+1} - p_\ell - i\gamma\Delta t(e^{-ip_{\ell+1}} - 1)]}}. \end{aligned} \tag{A21}$$

The integral over p_1 reads

$$\int_{-\pi}^{\pi} \frac{dp_1}{2\pi} \frac{e^{ip_1 n_0 + \gamma\Delta t n_0(e^{-ip_1} - 1)}}{1 - e^{i[p_2 - p_1 - i\gamma\Delta t(e^{-ip_2} - 1)]}}. \tag{A22}$$

We can argue just as in the previous calculation to show that the net effect of doing this integral is to implement the constraint

$$p_1 = p_2 - i\gamma\Delta t(e^{-ip_2} - 1). \tag{A23}$$

Similarly, the effect of doing the integrals over $p_2, ..., p_{T-1}$ is to implement the constraints

$$p_\ell = p_{\ell+1} - i\gamma\Delta t(e^{-ip_{\ell+1}} - 1), \tag{A24}$$

for $\ell = 2, ..., T-1$. While this recurrence relation is probably not solvable in closed form, we can do a trick to evaluate

---

[2] These arguments can be made more rigorous if one wishes. Because path integrals are somewhat mathematically dubious in the first place, regularization techniques like these are required for the integrals involved to be well-defined. Ultimately, the proof is "in the pudding": If we get the right answer, which can easily be verified by substituting it directly into Eq. (A4), then these abuses can be excused.

p_ℓ (for $\ell = 1, ..., T-1$) in terms of p_T (the only momentum variable not yet integrated over). Notice that Eq. (A24) looks like a "backwards" Euler time step corresponding to the ordinary differential equation (ODE)

$$\dot{p} = -i\gamma(e^{-ip} - 1). \tag{A25}$$

This approximation becomes exact in the $\Delta t \to 0$ limit, which is the limit we are interested in. Eq. (A25) has the solution

$$e^{ip(t)} = 1 - Ce^{-\gamma t}, \tag{A26}$$

where C must be determined from the initial condition $p(0) = p_T$ (since the ODE runs "backwards"). Doing so, we have

$$e^{ip_0} = 1 - (1 - e^{ip_T})e^{-\gamma t} \tag{A27}$$

in the small Δt limit, where p_0 is defined via

$$e^{ip_0} := e^{ip_1 + \gamma\Delta t(e^{-ip_1} - 1)}. \tag{A28}$$

The intuition is that p_0 is what we reach after starting at p_T and taking T time steps of size Δt. Our last integral reads

$$\begin{aligned} &\int_{-\pi}^{\pi} \frac{dp_T}{2\pi}\, e^{ip_0 n_0} e^{-ip_T n_T} \\ &= \int_{-\pi}^{\pi} \frac{dp_T}{2\pi}\, [(1 - e^{-\gamma t}) + e^{-\gamma t} e^{ip_T}]^{n_0} e^{-ip_T n_T}. \end{aligned} \tag{A29}$$

Define $q := 1 - e^{-\gamma t}$ and change variables to $z = e^{ip_T}$ so that we have

$$\begin{aligned} &\oint \frac{dz}{2\pi i} \frac{[q + (1-q)z]^{n_0}}{z^{n_T+1}} \\ &= \sum_{j=0}^{n_0} \binom{n_0}{j} (1-q)^j q^{n_0 - j} \oint \frac{dz}{2\pi i} \frac{1}{z^{n_T - j + 1}}. \end{aligned} \tag{A30}$$

The contour integral is nonzero only for the term with $j = n_T$. Then $P(n_f, t; n_0, 0) = 0$ for $n_T > n_0$, and

$$P(n_f, t; n_0, 0) = \binom{n_0}{n_f} [e^{-\gamma t}]^{n_f} [1 - e^{-\gamma t}]^{n_0 - n_f} \tag{A31}$$

otherwise, i.e., we have a binomial distribution.

3. The chemical birth-death process

The chemical birth-death process models a species that is randomly created and randomly degrades. Its list of chemical reactions is

$$\varnothing \xrightarrow{k} X, \quad X \xrightarrow{\gamma} \varnothing, \tag{A32}$$

where k is the birth rate and γ is the death rate. This reaction corresponds to the CME

$$\begin{aligned} \frac{\partial P(n,t)}{\partial t} &= k[P(n-1,t) - P(n,t)] \\ &\quad + \gamma[(n+1)P(n+1,t) - nP(n,t)], \end{aligned} \tag{A33}$$

where $P(n,t)$ is the probability that the system has n X molecules at time t (with $n \in \{0, 1, 2, ...\}$). This path integral reads

$$\begin{aligned} &\lim_{T\to\infty} \sum_{n_1} \cdots \sum_{n_{T-1}} \int \frac{dp_1}{2\pi} \cdots \int \frac{dp_T}{2\pi} \\ &\times \exp\left\{ \sum_{\ell=1}^{T} -ip_\ell(n_\ell - n_{\ell-1}) + k\Delta t(e^{ip_\ell} - 1) \right. \\ &\left. + \gamma\Delta t\, n_{\ell-1}(e^{-ip_\ell} - 1) \right\}. \end{aligned} \tag{A34}$$

The terms that involve n_ℓ (for some $\ell = 1, ..., T-1$) look just the same as in the previous subsection, so the sums over the n_ℓ evaluate to the same answer. The same ODE constraint is also enforced by the integrals over the p_ℓ, leaving the only difference between this problem and the previous one in the evaluation of the final integral. We have

$$\begin{aligned} &\int_{-\pi}^{\pi} \frac{dp_T}{2\pi}\, e^{ip_0 n_0} e^{-ip_T n_T} e^{k\sum_{\ell=1}^{T}(e^{ip_\ell} - 1)\Delta t} \\ &= \int_{-\pi}^{\pi} \frac{dp_T}{2\pi}\, [q + (1-q)e^{ip_T}]^{n_0} e^{-ip_T n_T} e^{k\sum_{\ell=1}^{T}(e^{ip_\ell} - 1)\Delta t}, \end{aligned} \tag{A35}$$

where p_0 is defined as in Eq. (A28), and $q := 1 - e^{-\gamma t}$ as before. Observe that there is a term in the integrand, due to the birth reaction, that was not present in Eq. (A29). To proceed, we must evaluate that term in the small Δt limit. Note,

$$k\sum_{\ell=1}^{T}(e^{ip_\ell} - 1)\Delta t \approx k\int_0^t dt\, (e^{ip(t)} - 1) = -[1 - e^{ip_T}]\mu q, \tag{A36}$$

where we have exploited the link between Riemann sums and integrals, used the expression for $p(t)$ from Eq. (A26), and defined $\mu := k/\gamma$. Now our integral reads

$$e^{-\mu q} \int_{-\pi}^{\pi} \frac{dp_T}{2\pi} [q + (1-q)e^{ip_T}]^{n_0} e^{-ip_T n_T} e^{\mu q e^{ip_T}}. \tag{A37}$$

Changing variables to $z = e^{ip_T}$ yields

$$\begin{aligned} &e^{-\mu q} \oint \frac{dz}{2\pi i} \frac{[q + (1-q)z]^{n_0} e^{\mu q z}}{z^{n_T+1}} \\ &= e^{-\mu q} \sum_{j=0}^{n_0} \binom{n_0}{j} q^{n_0 - j}(1-q)^j \oint \frac{dz}{2\pi i} \frac{e^{\mu q z}}{z^{n_T - j + 1}}, \end{aligned} \tag{A38}$$

which can be evaluated via Cauchy's integral formula to obtain the result

$$P(n_f, t; n_0, 0) = \sum_{j=0}^{\min(n_0, n_f)} \binom{n_0}{j} q^{n_0 - j}(1-q)^j \frac{(\mu q)^{n_f - j} e^{-\mu q}}{(n_f - j)!}. \tag{A39}$$

This agrees with the result obtained using the Doi-Peliti path integral (cf. Eq. (77) of Ref. [118]), although the calculation was arguably less complicated here.

APPENDIX B: AN ALTERNATIVE ARGUMENT FOR TAYLOR EXPANDING THE ACTION

In Sec. IV D, we argued that Gillespie's second condition allows us to say that

$$[e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j}-1]a_j(\mathbf{x}_{k-1})\tau \approx \left[i\mathbf{p}_k\cdot\boldsymbol{\nu}_j-\frac{1}{2}\sum_{\ell=1}^{N}\sum_{\ell'=1}^{N}p_k^{\ell}p_k^{\ell'}\nu_{j\ell}\nu_{j\ell'}\right]a_j(\mathbf{x}_{k-1})\tau. \tag{B1}$$

In this Appendix, we present a slightly more careful alternative argument for the validity of this important approximation. Because the term we are approximating appears in the argument of the path integral's exponential [cf. Eq. (32)], we are *really* trying to say that

$$\exp\left\{\sum_{j=1}^{M}[e^{i\mathbf{p}_k\cdot\boldsymbol{\nu}_j}-1]a_j(\mathbf{n}_{k-1})\tau\right\} \approx \exp\left\{i\mathbf{p}_k\cdot\left(\sum_{j=1}^{M}\boldsymbol{\nu}_j a_j(\mathbf{n}_{k-1})\tau\right) -\frac{1}{2}\sum_{j=1}^{M}(\mathbf{p}_k\cdot\boldsymbol{\nu}_j)^2 a_j(\mathbf{n}_{k-1})\tau\right\}. \tag{B2}$$

Let us now show that this approximation holds. To ease notation, rewrite this as

$$\exp\left\{\sum_{j=1}^{M}[e^{i\mathbf{p}\cdot\boldsymbol{\nu}_j}-1]a_j\right\} \approx \exp\left\{i\mathbf{p}\cdot\left(\sum_{j=1}^{M}\boldsymbol{\nu}_j a_j\right)-\frac{1}{2}\sum_{j=1}^{M}(\mathbf{p}\cdot\boldsymbol{\nu}_j)^2 a_j\right\}, \tag{B3}$$

where we have removed the superfluous index on $\mathbf{p}$, and used the shorthand $a_j := a_j(\mathbf{n}_{k-1})\tau$. Start by Taylor expanding the exponentials:

$$\exp\left\{\sum_{j=1}^{M}[e^{i\mathbf{p}\cdot\boldsymbol{\nu}_j}-1]a_j\right\} = e^{-(a_1+\cdots a_M)}e^{e^{i\mathbf{p}\cdot\boldsymbol{\nu}_1}a_1}\cdots e^{e^{i\mathbf{p}\cdot\boldsymbol{\nu}_M}a_M} = \sum_{x_1,\ldots,x_M}\frac{(a_1)^{x_1}}{x_1!}e^{-a_1}\cdots\frac{(a_M)^{x_M}}{x_M!}e^{-a_M}e^{i\mathbf{p}\cdot(\boldsymbol{\nu}_1x_1+\cdots+\boldsymbol{\nu}_Mx_M)}, \tag{B4}$$

where the sum over each x_j runs from zero to infinity. Notice that the factor corresponding to the jth reaction looks like a Poisson distribution; since $a_j \gg 1$ (condition (ii)), it can be approximated as Gaussian in the usual way:

$$\frac{(a_j)^{x_j}}{x_j!}e^{-a_j} \approx \frac{1}{\sqrt{2\pi a_j}}e^{-\frac{(x_j-a_j)^2}{2a_j}}. \tag{B5}$$

Now the sum over each x_j can be approximated as an integral using the Euler-Maclaurin formula, and we can expand the domain of integration without significantly changing the result since the Gaussian function is sharply peaked:

$$\sum_{x_j=0}^{\infty}\frac{1}{\sqrt{2\pi a_j}}e^{-\frac{(x_j-a_j)^2}{2a_j}+i\mathbf{p}\cdot\boldsymbol{\nu}_jx_j} \approx \int_0^{\infty}dx_j\,\frac{1}{\sqrt{2\pi a_j}}e^{-\frac{(x_j-a_j)^2}{2a_j}+i\mathbf{p}\cdot\boldsymbol{\nu}_jx_j} \approx \int_{-\infty}^{\infty}dx_j\,\frac{1}{\sqrt{2\pi a_j}}e^{-\frac{(x_j-a_j)^2}{2a_j}+i\mathbf{p}\cdot\boldsymbol{\nu}_jx_j} = e^{i\mathbf{p}\cdot\boldsymbol{\nu}_ja_j-\frac{(\mathbf{p}\cdot\boldsymbol{\nu}_j)^2}{2}a_j}. \tag{B6}$$

Putting these factors together for each j yields Eq. (B3), the desired approximation. While the argument we just presented is somewhat more technical than the one in the main text, it also mirrors Gillespie's original derivation of the CLE more closely: We essentially approximated many Poisson distributions as normal distributions, just as Gillespie did [cf. Eq. (4)].

---

[1] D. T. Gillespie, J. Chem. Phys. **113**, 297 (2000).
[2] H. Risken and H. D. Vollmer, Z. Phys. B Condens. Matter **66**, 257 (1987).
[3] J. Łuczka, P. Hänggi, and A. Gadomski, Phys. Rev. E **51**, 2933 (1995).
[4] H. Risken, Fokker-planck equation, in *The Fokker-Planck Equation: Methods of Solution and Applications* (Springer, Berlin, 1996), pp. 63–95.
[5] N. G. v. Kampen, Can. J. Phys. **39**, 551 (1961).
[6] N. G. V. Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed., North-Holland Personal Library (Elsevier, Amsterdam, 2007).
[7] R. F. Pawula, Phys. Rev. **162**, 186 (1967).
[8] D. T. Gillespie, J. Chem. Phys. **115**, 1716 (2001).
[9] D. T. Gillespie and L. R. Petzold, J. Chem. Phys. **119**, 8229 (2003).
[10] M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie, J. Chem. Phys. **119**, 12784 (2003).
[11] Y. Cao, L. R. Petzold, M. Rathinam, and D. T. Gillespie, J. Chem. Phys. **121**, 12169 (2004).
[12] Y. Cao, D. T. Gillespie, and L. R. Petzold, J. Chem. Phys. **123**, 054104 (2005).
[13] Y. Cao, D. T. Gillespie, and L. R. Petzold, J. Chem. Phys. **124**, 044109 (2006).
[14] Y. Cao, D. T. Gillespie, and L. R. Petzold, J. Chem. Phys. **126**, 224101 (2007).
[15] J. J. Vastola and W. R. Holmes, arXiv:1909.12990 (2019).
[16] K. G. Wilson, Phys. Rev. B **4**, 3174 (1971).
[17] K. G. Wilson, Phys. Rev. B **4**, 3184 (1971).
[18] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, Reading, MA, 1995).

[19] M. D. Schwartz, *Quantum Field Theory and the Standard Model* (Cambridge University Press, Cambridge, UK, 2014).
[20] H. Georgi, Ann. Rev. Nucl. Part. Sci. **43**, 209 (1993).
[21] D. T. Gillespie, A. Hellander, and L. R. Petzold, J. Chem. Phys. **138**, 170901 (2013).
[22] D. T. Gillespie, J. Phys. Chem. B **113**, 1640 (2009).
[23] D. Schnoerr, G. Sanguinetti, and R. Grima, J. Chem. Phys. **141**, 024103 (2014).
[24] R. Grima, P. Thomas, and A. V. Straube, J. Chem. Phys. **135**, 084103 (2011).
[25] A. Duncan, S. Liao, T. Vejchodský, R. Erban, and R. Grima, Phys. Rev. E **91**, 042111 (2015).
[26] A. Ceccato and D. Frezzato, J. Chem. Phys. **148**, 064114 (2018).
[27] S. Winkelmann and C. Schutte, J. Chem. Phys. **147**, 114115 (2017).
[28] L. A. Harris and P. Clancy, J. Chem. Phys. **125**, 144107 (2006).
[29] L. A. Harris, A. M. Piccirilli, E. R. Majusiak, and P. Clancy, Phys. Rev. E **79**, 051906 (2009).
[30] K. A. Iyengar, L. A. Harris, and P. Clancy, J. Chem. Phys. **132**, 094101 (2010).
[31] E. L. Haseltine and J. B. Rawlings, J. Chem. Phys. **117**, 6959 (2002).
[32] H. Salis and Y. Kaznessis, J. Chem. Phys. **122**, 054103 (2005).
[33] F. Wu, T. Tian, J. B. Rawlings, and G. Yin, J. Chem. Phys. **144**, 174112 (2016).
[34] E. W. Wallace, D. T. Gillespie, K. R. Sanft, and L. R. Petzold, IET Syst. Biol. **6**, 102 (2012).
[35] R. Grima, J. Chem. Phys. **136**, 154105 (2012).
[36] J. Wang, K. Zhang, L. Xu, and E. Wang, Proc. Natl. Acad. Sci. USA **108**, 8257 (2011).
[37] J. X. Zhou, M. D. S. Aliyu, E. Aurell, and S. Huang, J. R. Soc., Interface **9**, 3539 (2012).
[38] P. Zhou and T. Li, J. Chem. Phys. **144**, 094109 (2016).
[39] M. L. Simpson, C. D. Cox, and G. S. Sayler, J. Theor. Biol. **229**, 383 (2004).
[40] M. Hinczewski and D. Thirumalai, J. Phys. Chem. B **120**, 6166 (2016).
[41] T. Brett and T. Galla, J. Chem. Phys. **140**, 124112 (2014).
[42] T. Choi, M. R. Maurya, D. M. Tartakovsky, and S. Subramaniam, J. Chem. Phys. **133**, 165101 (2010).
[43] S. Rudiger, Phys. Rep. **534**, 39 (2014).
[44] H. Li, Z. Hou, and H. Xin, Phys. Rev. E **71**, 061916 (2005).
[45] Y. Huang, S. Rudiger, and J. Shuai, Phys. Biol. **12**, 061001 (2015).
[46] E. Wallace, M. Benayoun, W. van Drongelen, and J. D. Cowan, PLoS One **6**, e14804 (2011).
[47] R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948).
[48] R. Feynman, A. Hibbs, and D. Styer, *Quantum Mechanics and Path Integrals*, Dover Books on Physics (Dover Publications, Mineola, NY, 2010).
[49] L. S. Schulman, *Techniques and Applications of Path Integration* (Dover Publications, Mineola, NY, 1981).
[50] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets*, 5th ed. (World Scientific, Singapore, 2009).
[51] R. Shankar, *Principles of Quantum Mechanics*, 2nd ed. (Springer US, 2011).
[52] R. Shankar, *Quantum Field Theory and Condensed Matter: An Introduction* (Cambridge University Press, Cambridge, UK, 2017).
[53] A. Caldeira and A. Leggett, Physica A **121**, 587 (1983).
[54] A. W. C. Lau and T. C. Lubensky, Phys. Rev. E **76**, 011123 (2007).
[55] K. Asheichyk, A. P. Solon, C. M. Rohwer, and M. Krüger, J. Chem. Phys. **150**, 144111 (2019).
[56] J. Wang, K. Zhang, H. Lu, and E. Wang, Biophys. J. **89**, 1612 (2005).
[57] J. Wang, K. Zhang, H. Lu, and E. Wang, Phys. Rev. Lett. **96**, 168101 (2006).
[58] S. Orioli, S. a Beccara, and P. Faccioli, J. Chem. Phys. **147**, 064108 (2017).
[59] V. Linetsky, Comput. Econ. **11**, 129 (1997).
[60] L. Ingber, Physica A **283**, 529 (2000).
[61] T. Spanio, J. Hidalgo, and M. A. Muñoz, Phys. Rev. E **96**, 042301 (2017).
[62] A. Kamenev, B. Meerson, and B. Shklovskii, Phys. Rev. Lett. **101**, 268103 (2008).
[63] U. C. Täuber, J. Phys. Conf. Ser. **319**, 012019 (2011).
[64] U. C. Täuber, J. Phys. A: Math. Theor. **45**, 405002 (2012).
[65] S. R. Serrao and U. C. Täuber, J. Phys. A: Math. Theor. **50**, 404005 (2017).
[66] G. K. Ocker, K. Josić, E. Shea-Brown, and M. A. Buice, PLoS Comput. Biol. **13**, e1005583 (2017).
[67] A. Iomin, Phys. Rev. E **100**, 010104(R) (2019).
[68] R. Baravalle, O. A. Rosso, and F. Montani, Physica A **486**, 986 (2017).
[69] P. C. Bressloff, J. Math. Neurosci. **5**, 4 (2015).
[70] P. C. Bressloff and J. M. Newby, Phys. Rev. E **89**, 042701 (2014).
[71] S.-W. Qiu and C. C. Chow, Phys. Rev. E **98**, 062414 (2018).
[72] M. Buice and C. Chow, Front. Comput. Neurosci. **7**, 162 (2013).
[73] C. Li and J. Wang, J. R. Soc., Interface **11**, 20140774 (2014).
[74] C. Li, T. Hong, and Q. Nie, Phys. Chem. Chem. Phys. **18**, 17949 (2016).
[75] C. Li and J. Wang, PLoS Comput. Biol. **9**, e1003165 (2013).
[76] J. Wang, K. Zhang, and E. Wang, J. Chem. Phys. **133**, 125103 (2010).
[77] J. Newby, arXiv:1412.8446 (2014).
[78] V. Elgart and A. Kamenev, Phys. Rev. E **70**, 041106 (2004).
[79] V. Elgart and A. Kamenev, Phys. Rev. E **74**, 041101 (2006).
[80] T. Li and F. Lin, J. Phys. A: Math. Theor. **49**, 135204 (2016).
[81] F. Benitez, C. Duclut, H. Chaté, B. Delamotte, I. Dornic, and M. A. Muñoz, Phys. Rev. Lett. **117**, 100601 (2016).
[82] P. Thomas, C. Fleck, R. Grima, and N. Popović, J. Phys. A: Math. Theor. **47**, 455007 (2014).
[83] A. Lazarescu, T. Cossetto, G. Falasco, and M. Esposito, J. Chem. Phys. **151**, 064117 (2019).
[84] M. Doi, J. Phys. A **9**, 1479 (1976).
[85] M. Doi, J. Phys. A **9**, 1465 (1976).
[86] L. Peliti, J. Phys. **46**, 1469 (1985).
[87] D. C. Mattis and M. L. Glasser, Rev. Mod. Phys. **70**, 979 (1998).
[88] P. C. Martin, E. D. Siggia, and H. A. Rose, Phys. Rev. A **8**, 423 (1973).
[89] H.-K. Janssen, Z. Physik B **23**, 377 (1976).
[90] C. De Dominicis, J. Phys. Colloques **37**, C1-247 (1976).

[91] C. De Dominicis and L. Peliti, Phys. Rev. B **18**, 353 (1978).

[92] J. A. Hertz, Y. Roudi, and P. Sollich, J. Phys. A: Math. Theor. **50**, 033001 (2016).

[93] M. Abramowitz, *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables* (Dover Publications, New York, NY, 1974).

[94] V. Lampret, Math. Mag. **74**, 109 (2001).

[95] R. Pathria and P. D. Beale, *Statistical Mechanics*, 3rd ed. (Academic Press, Boston, 2011).

[96] R. Kayser and J. E. Kilpatrick, J. Chem. Phys. **68**, 1511 (1978).

[97] Y. Liao and B. Miao, J. Chem. Phys. **142**, 164903 (2015).

[98] G. Rzadkowski and S. Lepkowski, J. Sci. Comput. **35**, 63 (2008).

[99] J. Cioslowski and J. Albin, J. Chem. Phys. **139**, 114109 (2013).

[100] Y. Karshon, S. Sternberg, and J. Weitsman, Proc. Natl. Acad. Sci. USA **100**, 426 (2003).

[101] V. Guillemin and S. Sternberg, Annales de l'Institut Fourier **57**, 2183 (2007).

[102] T. Tate, J. Funct. Anal. **260**, 501 (2011).

[103] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*, Springer Series in Synergetics (Springer, Berlin, 2009).

[104] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd ed., Springer series in synergetics No. 18 (Springer-Verlag, New York, 1996).

[105] D. T. Gillespie, J. Comput. Phys. **22**, 403 (1976).

[106] D. T. Gillespie, Physica A **188**, 404 (1992).

[107] D. T. Gillespie, *Markov Processes: An Introduction for Physical Scientists* (Elsevier Science, Amsterdam, 1992).

[108] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed., International Series in Pure and Applied Mathematics (McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976), pp. 325.

[109] R. Ellis, Trends Biochem. Sci. **26**, 597 (2001).

[110] A. P. Minton, J. Biol. Chem. **276**, 10577 (2001).

[111] F. Höfling and T. Franosch, Rep. Prog. Phys. **76**, 046602 (2013).

[112] G. Rivas and A. P. Minton, Trends Biochem. Sci. **41**, 970 (2016).

[113] Y. A. Komarova, I. A. Vorobjev, and G. G. Borisy, J. Cell Sci. **115**, 3527 (2002).

[114] J. Buceta, J. R. Soc., Interface **14**, 20170316 (2017).

[115] W. R. Holmes, Biophys. J. **116**, 1538 (2019).

[116] J. J. Vastola, arXiv:1910.09117 (2019).

[117] J. J. Vastola, arXiv:1910.10807 (2019).

[118] J. J. Vastola, arXiv:1911.00978 (2019).

[119] C. C. Chow and M. A. Buice, J. Math. Neurosci. **5**, 8 (2015).