

Strong diffusion formulation of Markov chain ensembles and its optimal weaker reductions

Marifi Güler*

Department of Computer Engineering, Eastern Mediterranean University, via Mersin 10, Famagusta, Turkey

(Received 1 July 2017; revised manuscript received 30 August 2017; published 16 October 2017)

Two self-contained diffusion formulations, in the form of coupled stochastic differential equations, are developed for the temporal evolution of state densities over an ensemble of Markov chains evolving independently under a common transition rate matrix. Our first formulation derives from Kurtz's strong approximation theorem of density-dependent Markov jump processes [Stoch. Process. Their Appl. **6**, 223 (1978)] and, therefore, strongly converges with an error bound of the order of $\ln N/N$ for ensemble size N . The second formulation eliminates some fluctuation variables, and correspondingly some noise terms, within the governing equations of the strong formulation, with the objective of achieving a simpler analytic formulation and a faster computation algorithm when the transition rates are constant or slowly varying. There, the reduction of the structural complexity is optimal in the sense that the elimination of any given set of variables takes place with the lowest attainable increase in the error bound. The resultant formulations are supported by numerical simulations.

DOI: [10.1103/PhysRevE.96.042136](https://doi.org/10.1103/PhysRevE.96.042136)**I. INTRODUCTION**

The most commonly used stochastic processes in science and engineering are of the Markovian type [1,2]. The temporal evolution of one Markov chain is comparatively easy to analyze [3,4], but the mathematics becomes quite complicated for the collective behavior of an ensemble of chains, dating back to Doob [5] and Spitzer [6]. The complication applies even in the absence of interaction among the chains of the ensemble due to the fact that the autocorrelation times of the state densities are finite but not 0. The *density of a state* here refers to the number of chains found in that state at a particular instant, in proportion to the total number of chains over the ensemble.

In a pioneering study, Kurtz [7] introduced a diffusion model of the class densities in a population of individuals belonging to certain classes; a density-dependent Markov jump process was assumed to take place there among the classes (for a simplified account, see Ref. [8]). Kurtz has also given a strong convergence bound, of the order of the error of his diffusion approximation. In an alternate approach, due to van Kampen, density-dependent processes were modeled by an expansion of the Kolmogorov (or master) equation (see Chap. X in Ref. [1]); the associated Langevin formulation contains a matrix square root operation in its noise term, which makes the use of this formulation a somewhat burdensome task. Density-dependent Markov models have appeared in a variety of contexts, including chemical kinetics [9,10], ecological modeling [11], epidemics [12], metapopulations [13], and telecommunications [14].

Using an approach similar to that of van Kampen, Fox and Lu [15] derived a multidimensional Langevin equation for the dynamics of ion channel clusters contained within excitable membranes, where the behavior of each ion channel type is specified by a continuous-time Markov chain. To bypass the inherent need for the matrix square root calculations in that channel-based model, Fox and Lu proposed another model based on the channel subunits, namely, the gates [15]. Numerical experiments, however, have shown that, although

the channel-based model accurately simulates channel noise [16], the subunit model fails to produce good enough statistics [17]. Analytic solutions to the matrix square roots in the channel-based model have been obtained for both potassium and sodium channel types [18,19]. These matrix decompositions are not unique, as different but equivalent Itô stochastic differential equation models of random dynamical systems can be constructed [20]. The work of Fox and Lu has also motivated the development of alternative diffusion models that can efficiently realize the excitability and spontaneous firing in finite-size membranes [18,19,21,22].

In this paper, we study ensembles of noninteracting continuous-time Markov chains near dynamical equilibrium. Our conduct is for general Markov chains with any state transition diagram having constant or slowly varying transition rates, hence the formalism developed here applies to ensembles of any Markov chain as well as to the ion channel clusters. We report two related diffusion formulations in this context. The first formulation is derived from Kurtz's strong approximation theorem; hence, it includes no matrix square roots and strongly converges with an error bound as $O(\ln N/N)$ for ensemble size N . When the formulation is used for clusters of potassium channels and sodium channels, it yields Orio and Soudry's stochastic differential equations [18] developed specifically for the ion channels using heuristics.

The second formulation we introduce serves the purpose of providing a simpler analytic formulation and a faster computation algorithm than the aforementioned strong diffusion formulation. That formulation reduces the structural complexity of the strong formulation by eliminating some selected state density variables, and correspondingly some noise terms, therein. It is a model for temporal evolution of the number of chains in a prescribed set of *relevant states*. For example, in the case of ion channels within excitable membranes, the open channel state is the only relevant state, while the multiple number of closed channel states is hidden or not of direct concern. Elimination of the state density variables is not done simply by setting them to their mean values. Instead, the error resulting from the removal of variables is minimized by introducing extra terms into the retained differential equations; the extra terms are functions of the

*marifi.guler@gmail.com

retained variables. Thus, the reduced formulation provides a weaker but optimal approximation to the strong diffusion formulation in order to facilitate a structurally simpler set of model equations. It also offers a computation algorithm faster than that of the strong formulation provided that the transition rates are constant or slowly varying. We also provide a procedure that specifies the set of variables to be eliminated for a given level of desired structural complexity. There, the variables for the relevant states are, of course, not allowed to be eliminated.

In a context similar to that above, our prior work [22] formulated a weak approximation scheme, called “the minimal diffusion formulation,” to deduce how the number of chains in a single relevant state evolves in time. It was developed under the assumption that marginal state density fluctuations are only weakly non-Markovian. The formulation consisted of two specifically coupled Ornstein-Uhlenbeck processes in a stochastic differential equation representation; its structure does not change with the state space size or the transition matrix density. More recently, two variants (namely, *2v1n* and *1v1n* formulations) that reduce the structural complexity of the minimal diffusion formulation even further, without a significant loss of accuracy, were also formulated [23]. These models are essentially special cases of the present reduced diffusion formulation at the level of minimal structural complexity.

The organization of this paper is as follows. Section II gives the essential background and the moment relations. Section III is dedicated to the development of the strong diffusion formulation. The reduced diffusion formulation is elaborated in Sec. IV and Appendix A. Section V includes comparative numerical simulation results that complement the theory. Finally, Sec. VI includes some concluding remarks.

II. THE MOMENTS

Consider an ensemble of ergodic (irreducible) continuous-time Markov chains evolving independently under a common transition rate matrix in some finite space of states. For physical visualization, each chain in the ensemble can be imagined as a particle. Let $\mathbb{S} = \{0, 1, \dots, L\}$ be the state space and N the number of Markov chains in the ensemble. The number of chains in state l at a particular time is $N\psi_l$, where ψ_l is the density of state l . Symbolize the fluctuation in ψ_l by ϕ_l , that is,

$$\psi_l := \langle \psi_l \rangle + \phi_l, \quad (1)$$

where $\langle \dots \rangle$ denotes the expectation value. The state density expectation value $\langle \psi_l \rangle$ corresponds to the probability of finding a chain in state l . By definition it reads that

$$\sum_{l \in \mathbb{S}} \psi_l = 1, \quad (2)$$

$$\sum_{l \in \mathbb{S}} \langle \psi_l \rangle = 1, \quad (3)$$

$$\sum_{l \in \mathbb{S}} \phi_l = 0, \quad (4)$$

and

$$\langle \phi_l \rangle = 0. \quad (5)$$

The expectation values $\langle \psi_l \rangle$ ($l \in \mathbb{S}$) can be computed from the master equation. Let z_{lm} denote the transition rate from state l to state m . Then the following coupled differential equations govern the expectation values:

$$\frac{d\langle \psi_l \rangle}{dt} = \sum_{m \in \mathbb{D}^l} (-z_{lm}\langle \psi_l \rangle + z_{ml}\langle \psi_m \rangle). \quad (6)$$

Here $\mathbb{D}^l = \{m \mid m \in \mathbb{S} - \{l\}, z_{lm} \neq 0 \text{ or } z_{ml} \neq 0\}$ is the set of states that directly interact with state l through a transition.

A fundamental property of the master equation in Markov processes is that, as $t \rightarrow \infty$, all solutions tend to a stationary solution if the state set contains strictly a finite number of discrete states and the transition rates are constant in time. There exists only one stationary solution if the transition rate matrix is not decomposable. Therefore, with constant transition rates, the state density expectation values reach the unique steady-state solution identified by

$$\frac{d\langle \psi_l \rangle}{dt} = 0 \quad (7)$$

in the long-time limit. Among the $|\mathbb{S}| = L + 1$ equations given by Eq. (7), L are linearly independent; nevertheless, using these L equations in conjunction with Eq. (3) uniquely solves the stationary-state density averages.

Also, the second moments of the fluctuations are of concern to us. The fluctuation variances can be concluded from the binomial distribution dispersion relation [24] as

$$\langle \phi_l^2 \rangle = \frac{\langle \psi_l \rangle (1 - \langle \psi_l \rangle)}{N}. \quad (8)$$

For the pair cross-correlation functions, it follows that [22]

$$\langle \phi_l \phi_m \rangle = -\frac{\langle \psi_l \rangle \langle \psi_m \rangle}{N}, \quad l \neq m. \quad (9)$$

III. STRONG DIFFUSION FORMULATION OF THE ENSEMBLE

Our diffusion formulation derives from Kurtz’s strong approximation theorem [7]. Therefore, we first give a brief account of that theorem.

A. Kurtz’s theorem

Consider a population of size N , in which an individual can be in one of c classes. Let vector $\mathbf{X} \in \mathbb{Z}^c$ represent the numbers of individuals in certain classes within the population. The collective state vector \mathbf{X} is assumed to evolve in time as a Markov process through a finite collection of possible jumps $\mathbf{r} \in \mathbb{Z}^c$. The jump rates $\Lambda(\mathbf{X}; \mathbf{r})$, at which the jumps $\mathbf{X} \rightarrow \mathbf{X} + \mathbf{r}$ take place, are assumed to scale according to

$$\Lambda(\mathbf{X}; \mathbf{r}) = Nf(\Psi; \mathbf{r}), \quad (10)$$

where the vector $\Psi := \mathbf{X}/N$ appearing in function f represents the densities of the different classes within the population. Markov processes with jump rates of that form are called density-dependent processes.

Kurtz represents the process Ψ as a sum of Poisson processes over the jumps; then he replaces each Poisson process with a scalar Wiener process (Brownian motion), where a

strong convergence error, of order $\ln N/N$, is introduced. The result is a stochastic differential equation of the form

$$d\Psi(t) = \sum_{\mathbf{r}} \mathbf{r} f(\Psi(t); \mathbf{r}) dt + \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} \mathbf{r} \sqrt{f(\Psi(t); \mathbf{r})} dW^{(\mathbf{r})}(t), \quad (11)$$

where $W^{(\mathbf{r})}$ are independent Wiener processes.

B. Use of the theorem for the Markov chain ensemble

In the case of Markov chain ensembles, classes are the states \mathbb{S} of an individual chain. The chains in state j make a jump to state m with the collective jump (transition) rate $Nz_{jm}\psi_j$; recall here that $N\psi_j$ corresponds to the number of chains in state j and that z_{jm} denotes the transition rate of a single chain from state j to state m . The collective jump rates therefore scale according to Eq. (10), and it is apparent that the collective jumps are Poisson processes. Thus, we can use Eq. (11) to formulate the density dynamics of the ensemble. The jump vector corresponding to the jump from state j to state m , denoted $\mathbf{r}^{(j;m)}$, reads

$$\mathbf{r}^{(j;m)} = \begin{pmatrix} \circ \\ -1 \\ \circ \\ +1 \\ \circ \end{pmatrix} \begin{matrix} \leftarrow j \\ \\ \leftarrow m \end{matrix}, \quad (12)$$

where all the elements are 0, except that $r_j^{(j;m)} = -1$ and $r_m^{(j;m)} = 1$. Note that at this setting the decrease caused by the jump $\mathbf{r}^{(j;m)}$ in the number of chains in state j is compensated by the same amount of increase in the number of chains in state m ; that is, the total number of chains remains invariant. Then Eq. (11) becomes

$$d\Psi(t) = \sum_{\substack{j,m \in \mathbb{S} \\ j \neq m}} \mathbf{r}^{(j;m)} z_{jm} \psi_j dt + \frac{1}{\sqrt{N}} \sum_{\substack{j,m \in \mathbb{S} \\ j \neq m}} \mathbf{r}^{(j;m)} \sqrt{z_{jm} \psi_j} dW^{(j;m)}(t). \quad (13)$$

It is more convenient to write Eq. (13) in terms of component derivatives. One can derive by means of Eq. (12) that

$$\dot{\psi}_l = \sum_{m \in \mathbb{D}^l} (-z_{lm} \psi_l + z_{ml} \psi_m) + \frac{1}{\sqrt{N}} \times \sum_{m \in \mathbb{D}^l} (-\sqrt{z_{lm} \psi_l} \hat{\xi}^{(l;m)} + \sqrt{z_{ml} \psi_m} \hat{\xi}^{(m;l)}); \quad (14)$$

recall that \mathbb{D}^l is the set of states that connects with the state l . $\hat{\xi}$ are independent zero-mean scalar Gaussian white noises of unit variance. Due to their independence, the two additive noise terms appearing within the sum for each state pair can be substituted by a single noise,

$$\dot{\psi}_l = \sum_{m \in \mathbb{D}^l} (-z_{lm} \psi_l + z_{ml} \psi_m + \xi_{lm}), \quad (15)$$

where the Gaussian noise ξ_{lm} is of mean 0 and has the mean square, denoted $\langle \dots \rangle$, as

$$\langle \xi_{lm}(t) \xi_{lm}(t') \rangle = \frac{z_{lm} \psi_l(t) + z_{ml} \psi_m(t)}{N} \delta(t - t'), \quad l \neq m. \quad (16)$$

Note from Eqs. (14) and (15) that ξ_{lm} are not independent: they must satisfy

$$\xi_{ml} = -\xi_{lm}. \quad (17)$$

Conveniently, the noises ξ_{lm} ($m > l$) can be generated independently and then the others set dependently by Eq. (17). In consequence, our strong diffusion formulation consists of $|\mathbb{S}| = L + 1$ coupled stochastic differential equations given by Eq. (15) subject to Eqs. (16) and (17). Note that the formulation satisfies $\sum_l \dot{\psi}_l = 0$, consistent with the conservation law given by Eq. (2); hence, there are L independent equations. The number of noise terms needed is decided by the transition rate matrix density of the chains. There exists a noise term corresponding to each pair of states having a nonzero transition rate between the two. There is no requirement for matrix square root operations. The deviation of the evaluated state densities from the actual densities is bounded by the order of $\ln N/N$.

In the cases of potassium channel and sodium channel clusters, our formulation regenerates Orio and Soudry's stochastic differential equations [18], developed specifically for ion channels using heuristics over Fox and Lu's channel-based model; in this context, see also the work of Dangerfield, Kay, and Burrage [19]. Channel-based Fox and Lu equations contain the matrix square root $S = \sqrt{D}$ of the diffusion matrix D (namely, $SS^T = D$) for both potassium and sodium channels. The decomposition of the diffusion matrix, however, is not unique. The decompositions put forward in both Ref. [18] and Ref. [19] are the same, and the matrix S obtained is of the smallest dimension that satisfies $SS^T = D$ [9]. There is no such ambiguity in the context of our formulation, and the strong formulation supports the above decomposition of Refs. [18] and [19], as Orio and Soudry's dynamics arises from our formulation. More importantly, beyond the case of ion channels, the strong diffusion formulation directly applies for any arbitrarily given state transition diagram.

IV. THE REDUCED FORMULATION

In many applications, not all, but only a subset of, the states—often just one, as in the case of ion channels—is relevant, the others being hidden or not of direct concern. The reduced formulation simplifies the strong formulation in this situation by eliminating some of the hidden state density variables and, correspondingly, some noise terms from its governing equations. This way the structural complexity of the governing equations can be reduced to the desired level but, of course, at the expense of some increase in the approximation error bound. The formulation is optimal in the sense that it reduces the complexity with the least possible increase in the error. In developing the formulation, we assume that the transition rates are constant (or slowly varying) in time and that the ensemble had enough time to reach dynamical equilibrium so that the steady-state condition, (7), applies.

A. Development of the formulation

First, by means of Eqs. (1) and (6), we write Eq. (15) in terms of the state density fluctuations rather than the state densities themselves:

$$\dot{\phi}_l = \sum_{m \in \mathbb{D}^l} (-z_{lm}\phi_l + z_{ml}\phi_m + \xi_{lm}). \quad (18)$$

By means of Eq. (1), the differentiation $\dot{\phi}_l$ has the apparent meaning

$$\dot{\phi}_l := \frac{d}{dt}(\psi_l - \langle \psi_l \rangle).$$

By splitting the sum $\sum_{m \in \mathbb{D}^l} z_{ml}\phi_m$ in it, Eq. (18) can be rewritten identically as

$$\begin{aligned} \dot{\phi}_l = & -\phi_l \sum_{m \in \mathbb{D}^l} z_{lm} + \sum_{j \in \mathbb{D}^l - \mathbb{E}^l} z_{jl}\phi_j - \sum_{p \in \mathbb{S} - \mathbb{E}} a_{lp}\phi_p \\ & + F_l + \sum_{m \in \mathbb{D}^l} \xi_{lm}, \end{aligned} \quad (19)$$

where F_l stands for

$$F_l := \sum_{k \in \mathbb{E}^l} z_{kl}\phi_k + \sum_{p \in \mathbb{S} - \mathbb{E}} a_{lp}\phi_p. \quad (20)$$

Here the parameters a_{lp} are to be determined. $\mathbb{E} \subset \mathbb{S}$ is the set of states where their corresponding fluctuation variables are desired to be eliminated from the governing equations, that is, if the variable ϕ_q is desired to be eliminated, then $q \in \mathbb{E}$. We refer to the states in \mathbb{E} as the *eliminated states* and the states in $\mathbb{S} - \mathbb{E}$ as the *retained states*. The set \mathbb{E}^l denotes states that are both in \mathbb{E} and in \mathbb{D}^l , that is, eliminated states in direct interaction with state l . Then $\mathbb{D}^l - \mathbb{E}^l$ is the set of retained states in \mathbb{D}^l . The rationale behind the split was to group the retained state variables and the eliminated state variables into separate summations, for reasons that become clear below.

Our reduced formulation sets the function F_l , for all $l \in \mathbb{S} - \mathbb{E}$, to 0, with the objective of removing the sum of the eliminated state variables, namely, $\sum_{k \in \mathbb{E}^l} z_{kl}\phi_k$, from the equation for $\dot{\phi}_l$. Note that, although not necessary in principle, F_l additionally contains the weighted sum of all the retained variables; this sum is compensated in Eq. (19). The presence of the second sum in Eq. (20) serves the purpose of confining F_l 's value as close to 0 as possible so that the error brought in by the removal of F_l from Eq. (19) is at its minimum. Due to Eqs. (4) and (9), the second sum in F_l lowers the variance $\langle F_l^2 \rangle$ provided that the parameters a_{lp} are set properly. It derives from Eq. (20) that $\langle F_l^2 \rangle$ attains its minimum with the parameter values given by

$$a_{lp} = \frac{\sum_{k \in \mathbb{E}^l} z_{kl} \langle \psi_k \rangle}{1 - \sum_{q \in \mathbb{S} - \mathbb{E}} \langle \psi_q \rangle}, \quad \forall p \in \mathbb{S} - \mathbb{E}. \quad (21)$$

With this parameter attribution, $\langle F_l^2 \rangle$ can be shown to read

$$\langle F_l^2 \rangle = \frac{1}{N} \sum_{k \in \mathbb{E}^l} z_{kl} \langle \psi_k \rangle \left(z_{kl} - \sum_{j \in \mathbb{E}^l} z_{jl} \langle \psi_j \rangle \right) / \sum_{s \in \mathbb{E}} \langle \psi_s \rangle. \quad (22)$$

If the second sum in Eq. (20) were not included, we would have Eq. (22) in the absence of the denominator $\sum_{s \in \mathbb{E}} \langle \psi_s \rangle$. Since the denominator is less than 1, it shows that the presence

of the second sum in Eq. (20) indeed facilitates a smaller $\langle F_l^2 \rangle$. Appendix A provides the detailed derivation of Eqs. (21) and (22).

It is useful to know not just $\langle F_l^2 \rangle$ but also its value with respect to the variance of the right-hand side of Eq. (18). Ignoring the noise term in it for the moment, the right-hand side, denoted RH_l , is

$$\text{RH}_l = \sum_{m \in \mathbb{D}^l} (-z_{lm}\phi_l + z_{ml}\phi_m). \quad (23)$$

Its variance can be evaluated down to

$$N \langle \text{RH}_l^2 \rangle = \sum_{m \in \mathbb{D}^l} z_{ml}^2 \langle \psi_m \rangle + \frac{1}{\langle \psi_l \rangle} \left(\sum_{m \in \mathbb{D}^l} z_{ml} \langle \psi_m \rangle \right)^2. \quad (24)$$

In obtaining this equation, we made use of the equality

$$\langle \psi_l \rangle \sum_{m \in \mathbb{D}^l} z_{lm} = \sum_{m \in \mathbb{D}^l} z_{ml} \langle \psi_m \rangle \quad (25)$$

derived from the simultaneous use of Eqs. (6) and (7). In addition, Eqs. (8) and (9) were utilized. Consequently, we obtain for the ratio that

$$\frac{\langle F_l^2 \rangle}{\langle \text{RH}_l^2 \rangle} = \frac{\sum_{k \in \mathbb{E}^l} z_{kl}^2 \langle \psi_k \rangle - (\sum_{k \in \mathbb{E}^l} z_{kl} \langle \psi_k \rangle)^2 / \sum_{s \in \mathbb{E}} \langle \psi_s \rangle}{\sum_{m \in \mathbb{D}^l} z_{ml}^2 \langle \psi_m \rangle + (\sum_{m \in \mathbb{D}^l} z_{ml} \langle \psi_m \rangle)^2 / \langle \psi_l \rangle}, \quad (26)$$

which indicates expectantly that the smaller the eliminated state sets \mathbb{E} and \mathbb{E}^l are, the better approximation the reduced formulation is. The error is in fact smaller than the one given by Eq. (26). This is because we ignored the noise terms ξ_{lm} from Eq. (18) in evaluating Eq. (26), and $\langle \text{RH}_l^2 \rangle$ increases with the inclusion of the noise terms.

We cannot use the noise mean squares as given by Eq. (16) in the reduced formulation since that equation accommodates the eliminated fluctuation variables as well as the retained variables. Therefore, we set the fluctuations for the eliminated variables to 0; that is, we modify the mean squares as

$$\langle \xi_{lm}(t) \xi_{lm}(t') \rangle = \frac{z_{lm} \bar{\psi}_l(t) + z_{ml} \bar{\psi}_m(t)}{N} \delta(t - t'), \quad l \neq m, \quad (27)$$

where the overbar is used to designate the identification

$$\bar{\psi}_j = \begin{cases} \psi_j & \text{if } j \notin \mathbb{E}, \\ \langle \psi_j \rangle & \text{if } j \in \mathbb{E}. \end{cases} \quad (28)$$

The modification does not have a significant effect, as the variances of the fluctuations are much smaller than 1. In this respect, one can set $\bar{\psi}_j = \langle \psi_j \rangle$ even for $j \notin \mathbb{E}$.

The governing equations for the reduced formulation are thus given by the $|\mathbb{S} - \mathbb{E}|$ coupled stochastic differential equations

$$\begin{aligned} \dot{\phi}_l = & -\phi_l \sum_{m \in \mathbb{D}^l} z_{lm} + \sum_{j \in \mathbb{D}^l - \mathbb{E}^l} z_{jl}\phi_j - \frac{\sum_{k \in \mathbb{E}^l} z_{kl} \langle \psi_k \rangle}{1 - \sum_{q \in \mathbb{S} - \mathbb{E}} \langle \psi_q \rangle} \\ & \times \sum_{p \in \mathbb{S} - \mathbb{E}} \phi_p + \sum_{m \in \mathbb{D}^l} \xi_{lm}, \quad l \in \mathbb{S} - \mathbb{E}, \end{aligned} \quad (29)$$

where not all the noises are independent, as they satisfy Eq. (17). Noise mean squares are given by Eq. (27). Following the solution of the governing equations, (29), the state densities easily result from Eq. (1). If desired, Eq. (29) can be rewritten, through Eqs. (1) and (6), in terms of the state density variables rather than the fluctuation variables. This, however, does not make the equation any simpler or shorter; in fact, the state density expectation values still appear in the equation and some extra terms arise. In the case of no eliminated states, i.e., $\mathbb{E} = \emptyset$, the reduced formulation is the same as the strong formulation, as it should be. With regard to the strong formulation, the reduced formulation not only reduces the number of variables from $|\mathbb{S}|$ to $|\mathbb{S} - \mathbb{E}|$ but also makes the noises $\{\xi_{ab} | a \in \mathbb{E} \text{ and } b \in \mathbb{E}\}$ redundant. Moreover, the sum of noises $\{\xi_{lb} | l \in \mathbb{S} - \mathbb{E} \text{ and } b \in \mathbb{E}'\}$ that may appear in Eq. (29) can be substituted by a single noise since these noises do not appear anywhere except in ϕ_l . Thus, the degree of simplification and, consequently, the reduction in the computational cost achieved by the reduced formulation depend on the number of eliminated states and the number of noises that have become redundant. We further address this issue in Sec. V and in Appendix B.

We assumed the transition rates to be constant or slowly varying, primarily for the following reason. Since the state density expectation values appear explicitly in the governing equations, (29), their calculation is required, which degrades the computational efficiency of the formulation. For constant transition rates, steady-state conditions, (7), nevertheless hold once equilibrium is reached. Hence, the expectation values need not be evaluated at each time step; it suffices to compute them, or solve them analytically, only once, at the beginning, from the system of linear equations (3), (6), and (7). The steady-state conditions can be safely presumed to be intact, to a high degree of accuracy, even when the rates are not constant but slowly varying. In this case, first, the expectation values are solved analytically, and then the solutions obtained as functions of the rates are used in the governing equations. To further clarify this point and to provide the reader with a simple case study that illustrates the implementation of the introduced formulations, we exploit the kinetic scheme of potassium channels in Appendix B.

B. A procedure for the specification of retained and eliminated states

In principle, any set of hidden (nonrelevant) states can be chosen as the eliminated states. However, forming the eliminated state set by giving higher precedence to states that are distant to the relevant states seems more effective: in this way, there exists no error originating from the variables of the states neighboring the relevant states; the error propagates from the states at a distance. Let us attribute a *level* to each state as follows. The relevant states are at level 0. The states at level n are such that they can be reached from at least one relevant state by n number of successive transitions, but it is not possible to reach them by less than n transitions from any relevant state. Our procedure for specifying the retained and eliminated state sets makes use of the levels. Let L_n denote the total number of states from level 0 up to and including level n . Suppose that d number of states are desired to be retained. Also suppose that

d falls into the range $L_n \leq d < L_{n+1}$. Then the procedure is such that all states from level 0 up to and including level n are retained states. In addition, $d - L_n$ number of states at level $n + 1$ are retained states which can be selected randomly. The remaining states at that level and the states at higher levels are eliminated states.

For demonstration, consider the state transition diagram in Fig. 1. The diagram contains 17 states, one of which—state 0—is the relevant state. States 1 and 2 are at level 1, states 3–5 are at level 2, states 6–8 are at level 3, states 9–11 are at level 4, states 12 and 13 are at level 5, and states 14–16 are at level 6. If, for example, four of the states are desired to be retained, then states $\{0, 1, 2\}$ and one randomly selected state among $\{3, 4, 5\}$ are the retained states.

At this point, we address a study [25] which argues that the computing time of the exact microscopic Markov simulations can be shortened considerably, without significant loss of accuracy, by regarding fluctuations only for transitions to and from states directly connected to the relevant state in every chain, setting all other fluctuations to 0. The method was named “stochastic shielding.” However, a diffusion approximation based on stochastic shielding cannot be the same as our

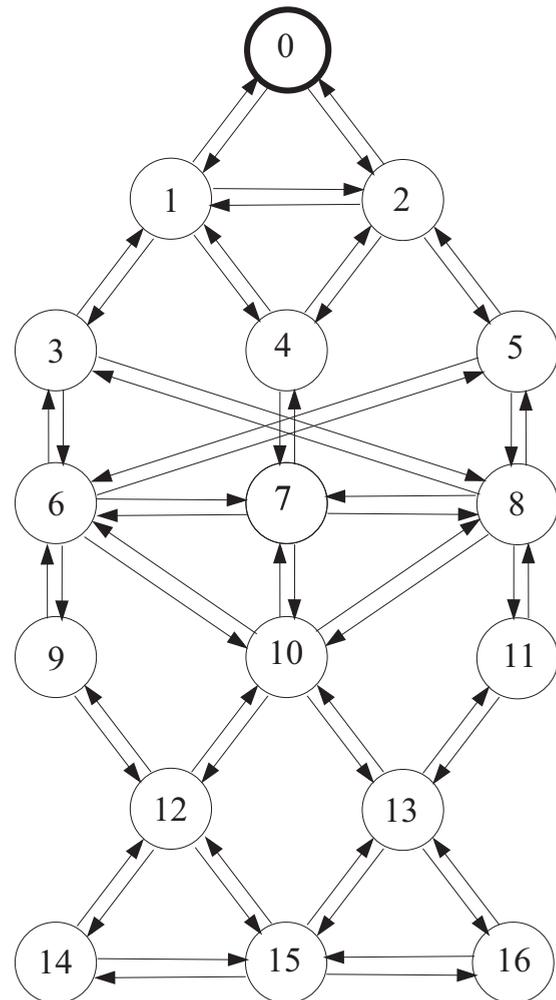


FIG. 1. State transition diagram used in the demonstrations and simulations. State 0 is the relevant state.

reduced formulation at level 1. This is because the reduced formulation does not completely neglect the fluctuations in those states that are not directly connected to the relevant state but, instead, incorporates their collective effect optimally into the governing equations. For further discussion of the issue, see Ref. [23].

C. The case of a single retained state

Consider the special case of having just one retained state, say state r . In this case, we have that $\mathbb{E}' = \mathbb{D}^r$ and $\mathbb{E} = \mathbb{S} - \{r\}$. Then, using Eq. (25), it derives from the governing equations, (29), that

$$\dot{\phi}_r = -\frac{\phi_r}{1 - \langle \psi_r \rangle} \sum_{m \in \mathbb{D}^r} z_{rm} + \xi, \quad (30)$$

in which we have substituted the noise ξ for the sum

$$\sum_{m \in \mathbb{D}^r} \xi_{rm}, \quad (31)$$

owing to the independence of the noises in the sum and their appearance only in that sum. By means of Eq. (27), ξ has the mean square

$$\langle \xi(t)\xi(t') \rangle = \frac{1}{N} \left(\langle \psi_r \rangle \sum_{m \in \mathbb{D}^r} z_{rm} + \sum_{m \in \mathbb{D}^r} z_{mr} \langle \psi_m \rangle \right), \quad (32)$$

where we have used the setting $\bar{\psi}_r = \langle \psi_r \rangle$. Employing the equality, (25), Eq. (32) can be written as

$$\langle \xi(t)\xi(t') \rangle = \frac{2}{N} \langle \psi_r \rangle \sum_{m \in \mathbb{D}^r} z_{rm}. \quad (33)$$

The governing equation, (30), together with Eq. (33), is in fact the same as the so-called $1v1n$ diffusion formulation introduced in our prior work [23]. The two formulations are thus in agreement with each other.

V. NUMERICAL EXPERIMENTS

In this section, we perform numerical experiments in order to assess the effectiveness of our models. The assessment develops computations and statistics from (a) the microscopic simulation of each Markov chain in the ensemble, (b) the strong formulation, and (c) the reduced formulation, in a comparative manner. The measurements are on the relevant state density, namely, on its time course, its expectation value, and the autocorrelation time.

In the numerical simulations of the strong formulation and the reduced formulation, the noise mean squares $\langle \xi_{lm}(t)\xi_{lm}(t') \rangle$ given by Eqs. (16) and (27), respectively, might numerically be negative if both $\langle \psi_l \rangle$ and $\langle \psi_m \rangle$ are very small; in such instances, the mean squares can simply be set to 0.

We perform two simulation experiments on ensembles of 300 Markov chains. In the first experiment, the Markov chains are characterized by the state transition diagram in Fig. 1, and in the second experiment, by that in Fig. 8.

A. Experiment 1

In this experiment, the ensemble we study consists of Markov chains identified by the state transition diagram in Fig. 1. Using state 0 as the relevant state, we obtained simulation results for the cases of

- (a) the strong diffusion formulation (abbreviated SF),
- (b) the reduced formulation with six retained states (abbreviated RF6),
- (c) the reduced formulation with three retained states (abbreviated RF3), and
- (d) the exact microscopic Markov simulation.

We have provided Supplemental Material that includes the C++ code for an implementation of SF, RF6, and RF3 [26].

For the state transition diagram under consideration, it reads from the governing equations, (29), that there are 17 variables and 29 independent noises in SF, 6 variables and 12 independent noises in RF6, and 3 variables and 7 independent noises in RF3. In fact, the numbers of noises for the cases of RF6 and RF3 can be brought down further, to 10 and 5, respectively. In the case of RF3, the retained states are 0, 1, and 2. Consequently, the noises ξ_{13} and ξ_{14} (or ξ_{31} and ξ_{41}) appearing in $\dot{\phi}_1$ do not appear in $\dot{\phi}_0$ or $\dot{\phi}_2$. It therefore follows that the sum $\xi_{13} + \xi_{14}$ in $\dot{\phi}_1$ can be substituted by a

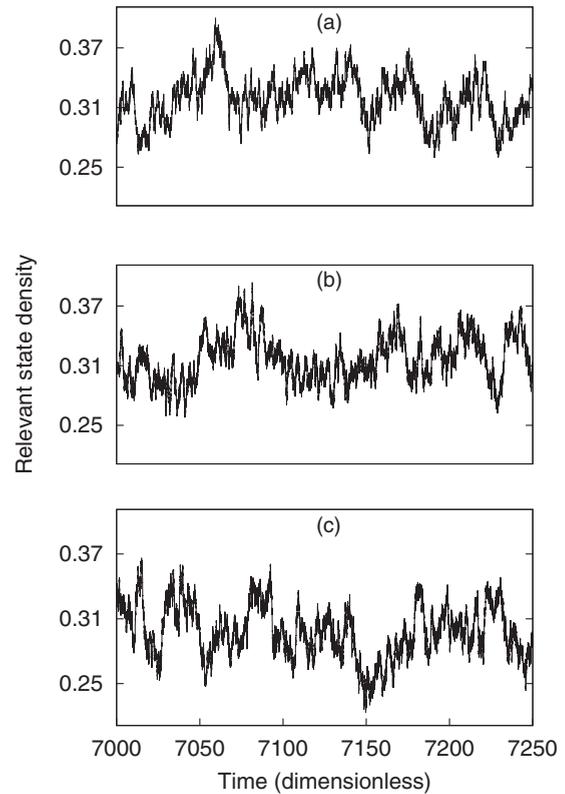


FIG. 2. A sample time course of the relevant state density for an ensemble of 300 Markov chains each evolving independently under the state transition diagram in Fig. 1. A randomly generated transition rate matrix was used. Time courses are for (a) the microscopic Markov simulation, (b) the strong formulation, and, (c) the reduced formulation with six retained states. The deterministic state density value in this particular case is 0.31.

single noise. Likewise, a single noise can substitute for the sum $\xi_{24} + \xi_{25}$ in ϕ_2 . A similar procedure applies for RF6.

The above numbers of variables and noises in SF, RF6, and RF3 facilitate the estimation of the relative computational costs of these formulations. RF6 and RF3 are about three and six times faster than SF, respectively. RF3 is about twice as fast as RF6. The measurement of computation times in a software implementation can be dependent on the coding and the random number generator used. In our coding, we found RF6 and RF3 to be 3.1 and 5.9 times faster than SF, respectively. As another two cases, consider RF2 (two retained states) and the minimal diffusion formulation [22]. There is no significant difference between the computational costs of the two formulations, as both formulations contain two variables, and for the noises the former includes three and the latter two. Suppose that 0 and 1 are the retained states for RF2. The three noises are then ξ_{01} , ξ_{02} , and another one as the substitute for $\xi_{12} + \xi_{13} + \xi_{14}$.

The sample time courses presented in Fig. 2 show the behavior of the relevant state density in the cases of the microscopic simulation, SF and RF6. The three time courses appear to behave in similar manner. For a more detailed quantitative investigation, we measure the expectation value, standard deviation, and autocorrelation time of the relevant state density. Results from the strong and reduced formulations are presented relative to the corresponding microscopic Markov simulation results. Measurements were taken over 100 different random transition rate matrices. For that, every matrix element was randomly and independently generated within the range of 0–0.5 using the uniform distribution. The autocorrelation time, denoted τ , satisfies the equality

$$\frac{\langle \phi_0(t)\phi_0(t + \tau) \rangle}{\langle \phi_0^2(t) \rangle} = e^{-1}, \tag{34}$$

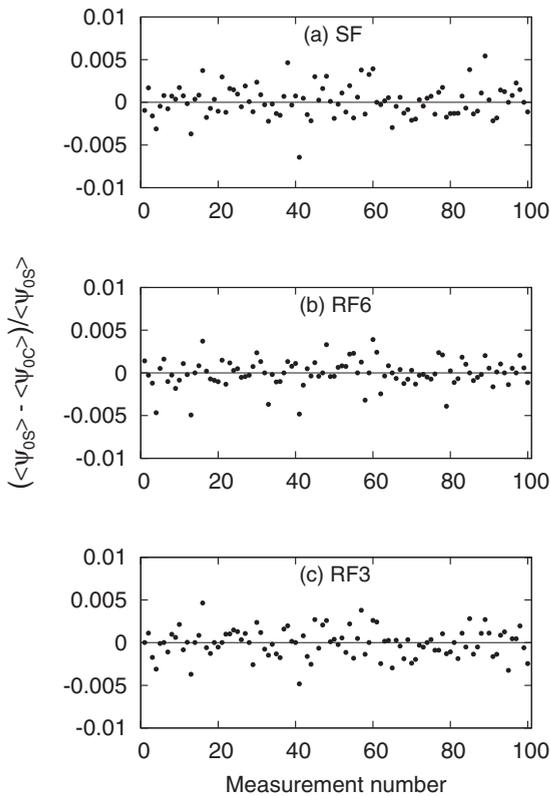


FIG. 3. Expectation value of the relevant state density (denoted $\langle \psi_{0C} \rangle$), measured using (a) the strong diffusion formulation, (b) the reduced formulation with six retained states, and (c) the reduced formulation with three retained states, relative to the standard deviation obtained from the exact microscopic Markov simulation (denoted $\langle \psi_{0S} \rangle$). An ensemble of 300 Markov chains, each evolving independently under the state transition diagram given in Fig. 1, was used. One hundred sets of standard deviation measurements were performed using a different transition rate matrix, with every matrix element being randomly and independently generated within the range of (0–0.5), in each set. The horizontal axis in each plot gives the measurement set number. The straight line, a guide for the eyes in each plot, indicates the situation of a perfect match between the formulation and the microscopic simulation.

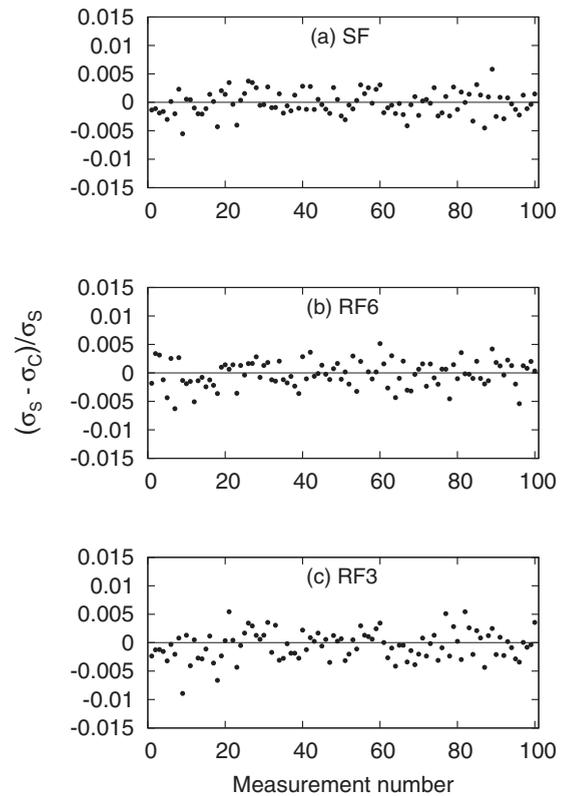


FIG. 4. Standard deviation of the relevant state density (denoted σ_C), measured using (a) the strong diffusion formulation, (b) the reduced formulation with six retained states, and (c) the reduced formulation with three retained states, relative to the standard deviation obtained from the exact microscopic Markov simulation (denoted σ_S). An ensemble of 300 Markov chains, each evolving independently under the state transition diagram given in Fig. 1, was used. One hundred sets of standard deviation measurements were performed using a different transition rate matrix, with every matrix element being randomly and independently generated within the range of 0–0.5, in each set. The horizontal axis in each plot gives the measurement set number. The straight line, a guide for the eyes in each plot, indicates the situation of a perfect match between the formulation and the microscopic simulation.

where the time t is large enough for equilibrium to be reached. It is seen from the results displayed in Figs. 3 and 4 that SF, RF6, and RF3 all predict the expectation values and standard deviations with virtually no error; less than 1% error materializes in all cases. Concerning the autocorrelation times, Fig. 5 shows that SF yields an excellent accuracy, whereas the RF6 findings and the RF3 findings can reveal errors of up to 8% and 24%, respectively. Having a higher accuracy in RF6 than in RF3 is expected. Of course, the error can be reduced to a desirable level by increasing the number of retained states. The worst-case autocorrelation time error of 24% in RF3 might seem high. However, consider the exponential decay function $A(x)$:

$$A(x) = e^{-x/\tau}. \tag{35}$$

We plotted the function in Fig. 6 using the worst-case autocorrelation time in RF3 and using the corresponding microscopic Markov simulation autocorrelation time. We see that two plots match quite closely; therefore, the state density

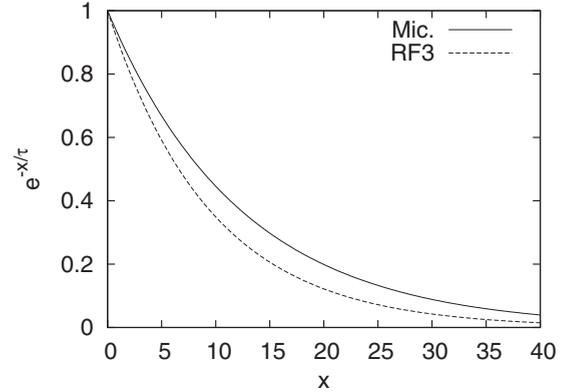


FIG. 6. Plot of the exponential decay function, given by Eq. (35), using (i) the worst-case autocorrelation time in the reduced formulation with three retained states (RF3) and (ii) the corresponding microscopic Markov simulation (Mic.) autocorrelation time.

autocorrelation function is reasonably good even in the case of RF3 and much better in the case of RF6.

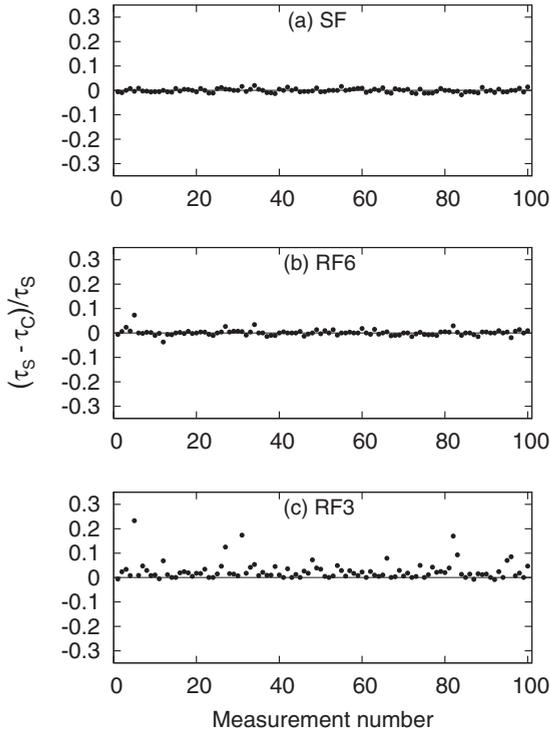


FIG. 5. Autocorrelation time of the relevant state density (denoted τ_C), measured using (a) the strong diffusion formulation, (b) the reduced formulation with six retained states, and (c) the reduced formulation with three retained states, relative to the standard deviation obtained from the exact microscopic Markov simulation (denoted τ_S). An ensemble of 300 Markov chains, each evolving independently under the state transition diagram given in Fig. 1, was used. One hundred sets of standard deviation measurements were performed using a different transition rate matrix, with every matrix element being randomly and independently generated within the range of 0–0.5, in each set. The horizontal axis in each plot gives the measurement set number. The straight line, a guide for the eyes in each plot, indicates the situation of a perfect match between the formulation and the microscopic simulation.

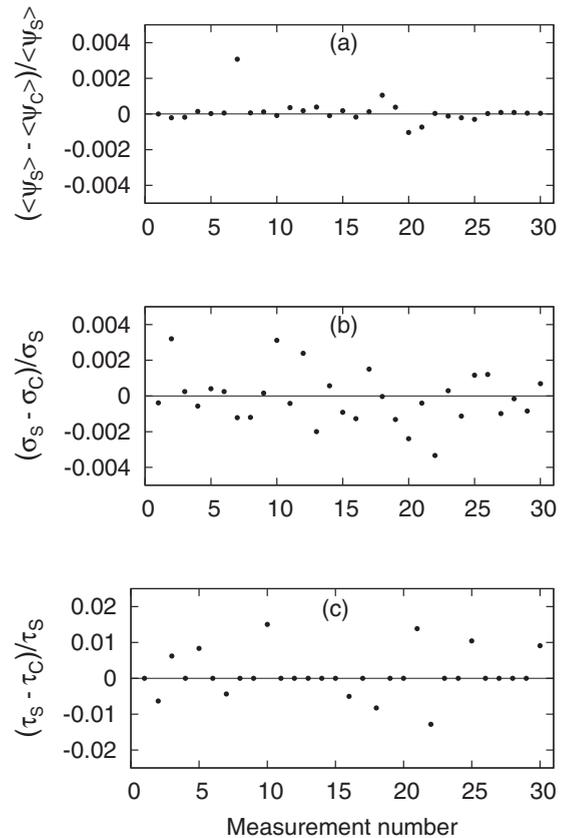


FIG. 7. Numerical experiments performed over an ensemble of 300 potassium channels, each characterized by the state transition diagram in Fig. 8. Measurements for the reduced formulation with two retained states (indicated by subscript “C”) are presented relative to the corresponding exact microscopic Markov simulation results (indicated by subscript “S”). (a) Expectation values, (b) standard deviations, and (c) autocorrelation times, over the density of the relevant state 4. A set of 30 randomly and independently generated rate values in the space of $\{\alpha \in (0:1); \beta \in (0:0.5)\}$ was used.

B. Experiment 2

Compared to the rather complicated state transition diagram used in experiment 1, we now employ a simpler transition diagram, namely, the kinetic scheme of potassium channels shown in Fig. 8 (Appendix B). We randomly and independently generated 30 pairs of rate values in the space of $\{\alpha \in (0 : 1); \beta \in (0 : 0.5)\}$ using the uniform distribution and measured the expectation value, standard deviation, and autocorrelation time of the relevant state density for each rate pair. The measurements obtained from the reduced formulation with two retained states are presented in Fig. 7. We observe that the formulation perfectly predicts the expectation values and standard deviations, essentially within the range of the numerical error limits. The prediction of autocorrelation times is excellent, with errors of less than 2%. When the strong formulation is adopted, even more accurate autocorrelation times are revealed (not plotted here).

The findings from these two experiments confirm the high accuracy of the strong formulation and support the adequacy of the reduced formulation as a simplified model. It is also indicated that taking even a small portion of the states as retained states, in the reduced formulation, suffices for good accuracy.

VI. CONCLUSION

In this paper, we have studied the temporal evolution of state density fluctuations in Markov chain ensembles evolving independently under a common transition rate matrix in some finite state space. We have in this context developed two related diffusion formulations in the form of coupled stochastic differential equations. The first formulation is referred to as the “strong diffusion formulation” since it was derived from Kurtz’s strong approximation theorem [7]; hence, it is a universal and consistent Langevin-type formulation that accurately represents the Markov channel dynamics. The numerical simulations we have provided support the high

accuracy of the formulation. Furthermore, simulations by other researchers [27] have shown the accuracy of Orio and Soudry’s stochastic differential equations [18]; these findings also support our formulation because the strong formulation generates Orio and Soudry equations for the ion channel clusters.

We have introduced the “reduced diffusion formulation” to obtain a simpler set of governing equations by reducing the numbers of variables and noise terms in use. Thereby, the reduced formulation facilitates a simpler analytic description and a faster computation algorithm over the strong formulation. It is important to emphasize that the simplification of the governing equations is realized with the least possible increase in the error. For the reduced formulation to apply, the set of relevant states must be a subset of the set of all states and the transition rates must be temporally constant (or slowly varying). The simulations indicate that using only a fraction of the states as the retained states suffices for good enough accuracy, which makes the reduced formulation an important model for practical use. In the case of a single retained state, the reduced formulation yields the so-called *1v1n* diffusion formulation introduced in our previous paper [23]; in the case of two retained states, it yields a result somewhat similar to that obtained with the so-called minimal diffusion formulation introduced in our earlier paper [22]. The reduced formulation, however, has the advantages that the retained state set there can be specified to be of any desired size and that it is a completely self-contained formulation.

APPENDIX A: DERIVATION OF EQS. (21) AND (22)

In this Appendix, we show that the minimum variance of the function F_l ($l \in \mathbb{S} - \mathbb{E}$) in Eq. (20) is obtained in the parameter space by setting a_{lp} as in Eq. (21). In addition, we derive Eq. (22) in that parameter regime.

The variance of F_l is

$$\langle F_l^2 \rangle = \sum_{p \in \mathbb{S} - \mathbb{E}} a_{lp}^2 \langle \phi_p^2 \rangle + \sum_{\substack{p, q \in \mathbb{S} - \mathbb{E} \\ q \neq p}} a_{lp} a_{lq} \langle \phi_p \phi_q \rangle + 2 \sum_{k \in \mathbb{E}^l} z_{kl} \sum_{p \in \mathbb{S} - \mathbb{E}} a_{lp} \langle \phi_k \phi_p \rangle + \sum_{k \in \mathbb{E}^l} z_{kl}^2 \langle \phi_k^2 \rangle + \sum_{\substack{j, k \in \mathbb{E}^l \\ j \neq k}} z_{jl} z_{kl} \langle \phi_j \phi_k \rangle, \quad (\text{A1})$$

where $l \in \mathbb{S} - \mathbb{E}$. Then making use of Eqs. (8) and (9) gives

$$\begin{aligned} N \langle F_l^2 \rangle &= \sum_{p \in \mathbb{S} - \mathbb{E}} a_{lp}^2 \langle \psi_p \rangle (1 - \langle \psi_p \rangle) - \sum_{\substack{p, q \in \mathbb{S} - \mathbb{E} \\ q \neq p}} a_{lp} a_{lq} \langle \psi_p \rangle \langle \psi_q \rangle - 2 \sum_{k \in \mathbb{E}^l} z_{kl} \sum_{p \in \mathbb{S} - \mathbb{E}} a_{lp} \langle \psi_k \rangle \langle \psi_p \rangle \\ &+ \sum_{k \in \mathbb{E}^l} z_{kl}^2 \langle \psi_k \rangle (1 - \langle \psi_k \rangle) - \sum_{\substack{j, k \in \mathbb{E}^l \\ j \neq k}} z_{jl} z_{kl} \langle \psi_j \rangle \langle \psi_k \rangle. \end{aligned} \quad (\text{A2})$$

Its derivative with respect to a_{lp} yields

$$N \frac{\partial \langle F_l^2 \rangle}{\partial a_{lp}} = 2 \langle \psi_p \rangle \left[(1 - \langle \psi_p \rangle) a_{lp} - \sum_{\substack{q \in \mathbb{S} - \mathbb{E} \\ q \neq p}} a_{lq} \langle \psi_q \rangle - \sum_{k \in \mathbb{E}^l} z_{kl} \langle \psi_k \rangle \right], \quad (\text{A3})$$

which can be rewritten as

$$N \frac{\partial \langle F_l^2 \rangle}{\partial a_{lp}} = 2 \langle \psi_p \rangle \left[a_{lp} - \sum_{q \in \mathbb{S} - \mathbb{E}} a_{lq} \langle \psi_q \rangle - \sum_{k \in \mathbb{E}^l} z_{kl} \langle \psi_k \rangle \right]. \quad (\text{A4})$$

The condition

$$\frac{\partial \langle F_l^2 \rangle}{\partial a_{lp}} = 0, \quad \forall p \in \mathbb{S} - \mathbb{E}$$

yields a unique minimum taking place at equal values of $\langle a_{lp} | \forall p \rangle$, and the value of a_{lp} at the minimum is as in Eq. (21) for each $l \in \mathbb{S} - \mathbb{E}$. Substituting Eq. (21) into Eq. (A2) and employing Eq. (3) results in Eq. (22).

APPENDIX B: A SIMPLE DEMONSTRATION USING POTASSIUM CHANNELS

The objective of this Appendix is to illustrate the implementation of our formulations through a simple example, namely, potassium channels in excitable membranes. The gating of ion channels is typically modeled by means of a continuous-time discrete state Markovian kinetic scheme, in which a channel can be open (the relevant state) or can be found in one of the multiple closed states. The state transition diagram of a potassium channel is as in Fig. 8. The transition rates α and β therein are time dependent through the transmembrane voltage. Nevertheless, it is the subthreshold activity that matters for the initiation of an action potential, and since the voltage does not vary much within that phase of activity, the rates can be supposed to be slowly varying.

In this case, Eq. (3) reads

$$\langle \psi_0 \rangle + \langle \psi_1 \rangle + \langle \psi_2 \rangle + \langle \psi_3 \rangle + \langle \psi_4 \rangle = 1, \quad (\text{B1})$$

and the master equation, (6), yields, at steady state, the coupled set of equations

$$\begin{aligned} \alpha \langle \psi_3 \rangle - 4\beta \langle \psi_4 \rangle &= 0, \\ 2\alpha \langle \psi_2 \rangle - (\alpha + 3\beta) \langle \psi_3 \rangle + 4\beta \langle \psi_4 \rangle &= 0, \\ 3\alpha \langle \psi_1 \rangle - 2(\alpha + \beta) \langle \psi_2 \rangle + 3\beta \langle \psi_3 \rangle &= 0, \\ 4\alpha \langle \psi_0 \rangle - (3\alpha + \beta) \langle \psi_1 \rangle + 2\beta \langle \psi_2 \rangle &= 0, \\ -4\alpha \langle \psi_0 \rangle + \beta \langle \psi_1 \rangle &= 0. \end{aligned} \quad (\text{B2})$$

Simultaneous solution of Eqs. (B1) and (B2) gives

$$\begin{aligned} \langle \psi_4 \rangle &= \bar{n}^4, \quad \langle \psi_3 \rangle = 4\bar{n}^3(1 - \bar{n}), \quad \langle \psi_2 \rangle = 6\bar{n}^2(1 - \bar{n})^2, \\ \langle \psi_1 \rangle &= 4\bar{n}(1 - \bar{n})^3, \quad \langle \psi_0 \rangle = (1 - \bar{n})^4, \end{aligned} \quad (\text{B3})$$

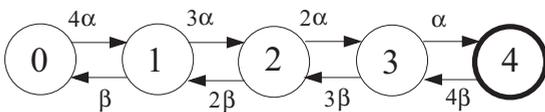


FIG. 8. Potassium channel state transition diagram. State 4 is the relevant state.

where

$$\bar{n} := \frac{\alpha}{\alpha + \beta}.$$

The strong formulation reads as follows. Equation (15) becomes

$$\begin{aligned} \dot{\psi}_4 &= -4\beta\psi_4 + \alpha\psi_3 - \xi_{34}, \\ \dot{\psi}_3 &= -(\alpha + 3\beta)\psi_3 + 2\alpha\psi_2 + 4\beta\psi_4 - \xi_{23} + \xi_{34}, \\ \dot{\psi}_2 &= -2(\alpha + \beta)\psi_2 + 3\alpha\psi_1 + 3\beta\psi_3 - \xi_{12} + \xi_{23}, \\ \dot{\psi}_1 &= -(3\alpha + \beta)\psi_1 + 4\alpha\psi_0 + 2\beta\psi_2 - \xi_{01} + \xi_{12}, \\ \dot{\psi}_0 &= -4\alpha\psi_0 + \beta\psi_1 + \xi_{01}, \end{aligned} \quad (\text{B4})$$

in which the noise variances conclude from Eq. (16)

$$\begin{aligned} \langle \xi_{34}^2 \rangle &= \frac{\alpha\psi_3 + 4\beta\psi_4}{N}, & \langle \xi_{23}^2 \rangle &= \frac{2\alpha\psi_2 + 3\beta\psi_3}{N}, \\ \langle \xi_{12}^2 \rangle &= \frac{3\alpha\psi_1 + 2\beta\psi_2}{N}, & \langle \xi_{01}^2 \rangle &= \frac{4\alpha\psi_0 + \beta\psi_1}{N}. \end{aligned} \quad (\text{B5})$$

For the reduced formulation, consider the case of two retained states—states 4 and 3—as an example. It reads through the governing equation, (29), that

$$\begin{aligned} \dot{\phi}_4 &= -4\beta\phi_4 + \alpha\phi_3 - \xi_{34}, \\ \dot{\phi}_3 &= -(\alpha + 3\beta)\phi_3 + 4\beta\phi_4 - \frac{2\alpha\langle \psi_2 \rangle}{1 - \langle \psi_3 \rangle - \langle \psi_4 \rangle} (\phi_3 + \phi_4) \\ &\quad - \xi_{23} + \xi_{34}, \end{aligned} \quad (\text{B6})$$

in which the noise variances conclude from Eqs. (27) and (28)

$$\begin{aligned} \langle \xi_{34}^2 \rangle &= \frac{\alpha\psi_3 + 4\beta\psi_4}{N}, \\ \langle \xi_{23}^2 \rangle &= \frac{2\alpha\langle \psi_2 \rangle + 3\beta\psi_3}{N}. \end{aligned} \quad (\text{B7})$$

Here, the dynamics is governed by two differential equations and two noises, whereas the corresponding strong formulation above consists of five differential equations and four noises. Hence, in this particular case, the reduced formulation is about twice as fast as the strong formulation.

- [1] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (North-Holland, Amsterdam, 2007).
- [2] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*, 4th ed. (Springer-Verlag, Berlin, 2009).
- [3] W. J. Anderson, *Continuous-Time Markov Chains: An Applications-Oriented Approach* (Springer-Verlag, New York, 1991).
- [4] J. R. Norris, *Markov Chains* (Cambridge University Press, Cambridge, UK, 1998).
- [5] J. L. Doob, *Stochastic Processes* (John Wiley & Sons, New York, 1953).
- [6] F. Spitzer, *Adv. Math.* **5**, 246 (1970).
- [7] T. G. Kurtz, *Stoch. Process. Their Appl.* **6**, 223 (1978).
- [8] P. H. Baxendale and P. E. Greenwood, *J. Math. Biol.* **63**, 433 (2011).
- [9] B. Mélykúti, K. Burrage, and K. C. Zygalakis, *J. Chem. Phys.* **132**, 164109 (2010).
- [10] P. K. Pollett and A. Vassallo, *Adv. Appl. Prob.* **24**, 875 (1992); D. T. Gillespie, *J. Chem. Phys.* **113**, 297 (2000); H. Ge and H. Qian, *J. Stat. Phys.* **166**, 190 (2017).
- [11] A. J. McKane and T. J. Newman, *Phys. Rev. Lett.* **94**, 218102 (2005).
- [12] I. Näsell, *Math. Biosci.* **179**, 1 (2002); D. Clancy, *J. Appl. Prob.* **42**, 726 (2005).
- [13] A. D. Barbour and A. Pugliese, *Ann. Appl. Prob.* **15**, 1306 (2005); J. V. Ross, *J. Math. Biol.* **52**, 788 (2006).
- [14] R. Serfozo, *Introduction to Stochastic Networks* (Springer-Verlag, New York, 1999); R. J. Gibbens, P. J. Hunt, and F. P. Kelly, in *Disorder in Physical Systems*, edited by G. R. Grimmett and D. J. A. Welsh (Oxford University Press, New York, 1998), pp. 113–127.
- [15] R. F. Fox and Y. N. Lu, *Phys. Rev. E* **49**, 3421 (1994).
- [16] J. H. Goldwyn, N. S. Imennov, M. Famulare, and E. Shear-Brown, *Phys. Rev. E* **83**, 041908 (2011).
- [17] H. Mino, J. T. Rubinstein, and J. A. White, *Ann. Biomed. Eng.* **30**, 578 (2002); S. Zeng and P. Jung, *Phys. Rev. E* **70**, 011903 (2004); I. C. Bruce, *Ann. Biomed. Eng.* **37**, 824 (2009); B. Sengupta, S. B. Laughlin, and J. E. Niven, *Phys. Rev. E* **81**, 011918 (2010); M. Güler, *J. Comput. Neurosci.* **31**, 713 (2011).
- [18] P. Orio and D. Soudry, *PLoS ONE* **7**, e36670 (2012).
- [19] C. E. Dangerfield, D. Kay, and K. Burrage, *Phys. Rev. E* **85**, 051907 (2012).
- [20] E. J. Allen, L. J. S. Allen, A. Arciniega, and P. E. Greenwood, *Stoch. Anal. Appl.* **26**, 274 (2008).
- [21] M. Güler, *Phys. Rev. E* **76**, 041918 (2007); D. Linaro, M. Storace, and M. Giugliano, *PLoS Comput. Biol.* **7**, e1001102 (2011); M. Güler, *Neural Comput.* **25**, 46 (2013); **25**, 2355 (2013).
- [22] M. Güler, *Phys. Rev. E* **91**, 062116 (2015).
- [23] M. Güler, *Phys. Rev. E* **93**, 022123 (2016).
- [24] M. C. Wang and G. E. Uhlenbeck, *Rev. Mod. Phys.* **17**, 323 (1945).
- [25] N. T. Schmandt and R. F. Galán, *Phys. Rev. Lett.* **109**, 118101 (2012).
- [26] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.96.042136> for C++ implementation of SF, RF6, and RF3 in the case of the transition diagram given in Fig. 1.
- [27] P. F. Rowat and P. E. Greenwood, *Front. Comput. Neurosci.* **8**, 111 (2014); Y. Huang, S. Rüdiger, and J. Shuai, *Phys. Biol.* **12**, 061001 (2015).