Rigorous elimination of fast stochastic variables from the linear noise approximation using projection operators

Philipp Thomas[*]

Department of Physics, Humboldt University of Berlin, Newtonstr. 15, D-12489 Berlin, Germany
and School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JR, United Kingdom

Ramon Grima[†]

School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JR, United Kingdom

Arthur V. Straube[‡]

Department of Physics, Humboldt University of Berlin, Newtonstr. 15, D-12489 Berlin, Germany

The linear noise approximation (LNA) offers a simple means by which one can study intrinsic noise in monostable biochemical networks. Using simple physical arguments, we have recently introduced the slow-scale LNA (ssLNA), which is a reduced version of the LNA under conditions of timescale separation. In this paper we present the first rigorous derivation of the ssLNA using the projection operator technique and show that the ssLNA follows uniquely from the standard LNA under the same conditions of timescale separation as those required for the deterministic quasi-steady-state approximation. We also show that the large molecule number limit of several common stochastic model reduction techniques under timescale separation conditions constitutes a special case of the ssLNA.

　　　　　　PACS number(s): 05.40.−a

I. INTRODUCTION

In the study of complex systems, it is common to invoke assumptions under which the dimension (and hence complexity) of the system is reduced; such a strategy often leads to relatively simple theories capable of exact analytical predictions, which offer insights typically lost in numerical approaches. This approach is particularly useful in the study of biochemical pathway dynamics, which typically involves the interaction of hundreds or thousands of different chemical species [1]. Deterministic models of such systems are frequently simplified by invoking the presence of well-separated timescales [1,2], in particular by means of the quasi-steady-state approximation (QSSA) [3,4]. However, it is well appreciated nowadays that a discrete stochastic approach in terms of chemical master equations (CMEs) is more faithful in describing dynamics inside living cells since the number of molecules of many species is in the range of few tens to few thousands [5]. The development of reduced stochastic models consistently coarse-grained over timescales presents a significantly larger challenge than the reduction of deterministic models because of the much larger number of differential equations which need to be solved in the former compared to the latter.

Indeed, it has been shown that it is not possible to formulate a reduced CME for the whole region of parameter space in which the deterministic QSSA is valid but rather only over a subset of this space [6–8]. For example, consider the Michaelis-Menten mechanism in which substrate reversibly binds with enzyme to form an enzyme-substrate complex which then decays into product. We are interested in conditions such that the substrate species decays over a much longer timescale than the enzyme and complex species. For this case, a consistently reduced CME can be written provided only that the complex decay rate into product is much less than its decay rate into substrate and enzyme [8,9]. However, it is possible to write down reduced deterministic rate equations (REs) using the QSSA even when the aforementioned condition is not satisfied [8].

Recently we have shown that a reduced stochastic description with the same range of validity as the deterministic QSSA is possible. This reduction is achieved by first approximating the CME by linear Langevin equations, an approximation called the linear noise approximation (LNA) and then integrating out the fast fluctuations such that one obtains a reduced set of linear Langevin equations for the fluctuations in the slowly varying species only. The latter is the slow-scale LNA (ssLNA) [10]. The advantage of the ssLNA over the reduced CME approach is that the former is valid over the same range of parameter space as the QSSA, a claim which has been numerically verified for a number of biochemical circuits including cooperative and noncooperative enzyme reactions and gene networks with or without negative feedback regulation [10]. The ssLNA, as it currently stands, has been derived by means of simple and intuitively clear physical arguments but is lacking a formal and rigorous derivation. In this article we provide such a derivation and also show the relationship of the ssLNA with other common methods of stochastic model reduction based on timescale separation assumptions.

The paper is divided as follows. In Sec. II we consider the CME of a general system of elementary chemical reactions with two characteristic and clearly separated timescales and reformulate the conventional system size expansion of the CME under such conditions. In Sec. III we use the leading order terms of the expansion to show that the deterministic

---

[*]philippocampus@gmail.com
[†]ramon.grima@ed.ac.uk
[‡]straube@physik.hu-berlin.de

limit of the CME under timescale separation conditions is formally the same as the conventional reduced rate equations obtained from the QSSA. The next to leading order terms of the expansion provide us with a Fokker-Planck equation which we reduce to a simpler Fokker-Planck equation for the slow fluctuations only, by means of the projection operator formalism. The reduced Fokker-Planck equation is obtained in closed-form for all monostable reaction networks and is the same as the ssLNA obtained in Ref. [10]. Finally in Sec. IV we show how the large molecule number limit of conventional stochastic reduction methods leads to a special case of the ssLNA and that hence these methods provide results in a more restrictive range of parameter space than the ssLNA does. The methodology of our approach is summarized in Fig. 1.

II. THE SYSTEM SIZE EXPANSION UNDER TIMESCALE SEPARATION CONDITIONS

We consider a system that comprises N chemical species X_i $(i = 1, \ldots, N)$ confined in a compartment of volume Ω
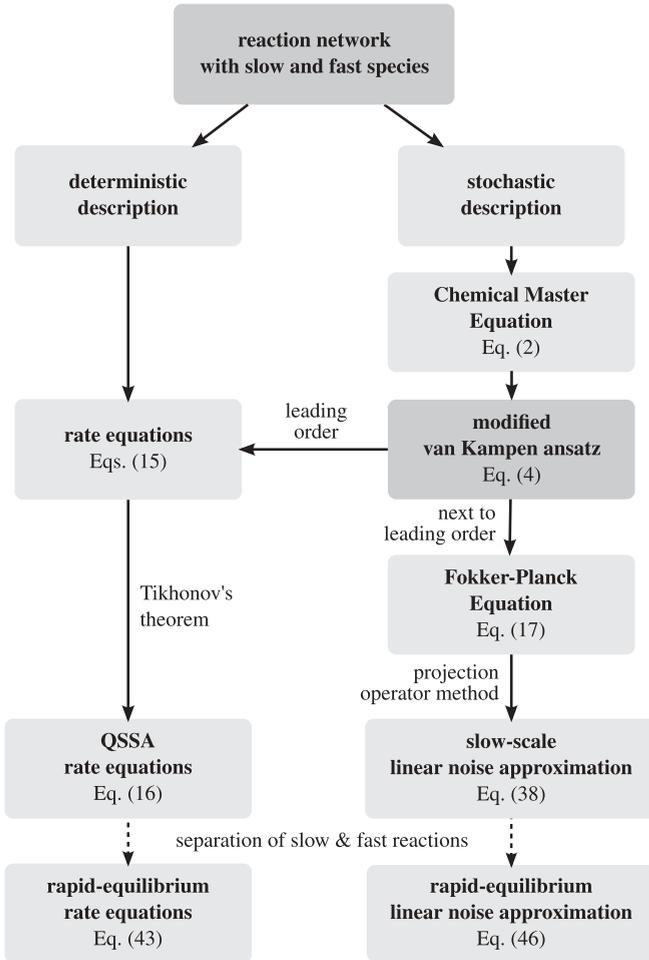

FIG. 1. Schematic illustrating the derivation of deterministic and stochastic quasi-steady-state approximations and rapid-equilibrium approximations for a reaction network with slow and fast species. The deterministic approach based on rate equations uses Tikhonov's theorem while the stochastic approach utilizes the projection operator method applied to the system size expansion derived from a modified van Kampen ansatz.

and assume that the species can interact via $j = 1, \ldots, R$ elementary chemical reactions

$$\sum_{i=1}^{N} s_{ij} X_i \xrightarrow{k_j} \sum_{i=1}^{N} r_{ij} X_i, \qquad (1)$$
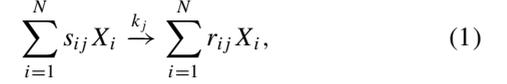
where s_{ij} and r_{ij} are the stoichiometric coefficients [11] and k_j is the macroscopic reaction rate. The constraint $\sum_i s_{ij} \leqslant 2$ for all j implies that the reactions are unimolecular or bimolecular and hence elementary. The total number of species N is assumed to comprise N_s slow and $N_f = N - N_s$ fast species, respectively. For convenience, we stick to the convention that X_1 to X_{N_s} denote the slow species, while X_{N_s+1} to X_N are reserved for the fast species.

In what follows, matrices are denoted by underlined quantities, and all vectors are column vectors. Let n_i denote the number of molecules of species i, then the probability $P(\vec{n}, t)$ to find the system in a particular state $\vec{n} = (n_1, \ldots, n_N)^T$ is determined by the CME [12,13]:

$$\frac{\partial P(\vec{n}, t)}{\partial t} = \Omega \sum_{j=1}^{R} \left(\prod_{i=1}^{N} E_i^{-S_{ij}} - 1 \right) \hat{f}_j(\vec{n}, \Omega) P(\vec{n}, t), \quad (2)$$

where $S_{ij} = (\underline{S})_{ij} = r_{ij} - s_{ij}$ is the stoichiometric matrix and E_i^z is a step operator, the action of which on some function of the absolute number of molecules results in the same function but with n_i replaced by $n_i + z$, and $\hat{f}_j(\vec{n}, \Omega)$ are the entries of the microscopic rate function vector. The product $\Omega \hat{f}_j(\vec{n}, \Omega) dt$ represents the propensity function, which has the meaning of the probability that reaction j takes place in a small time interval dt.

The CME is typically analytically intractable, and hence a systematic approximation method is needed. The system size expansion as developed by van Kampen is one such technique [12]. The heart of the method is the ansatz that the molecular concentration of the CME can be written as

$$\frac{\vec{n}}{\Omega} = \vec{\phi} + \Omega^{-1/2} \vec{\eta}, \qquad (3)$$

where $\vec{\phi}$ is vector of concentrations as given by the corresponding REs and the second term on the right-hand side is a stochastic term describing fluctuations about the concentrations of the REs.

We now impose timescale separation conditions, i.e., we assume that the transients in concentrations of a group of species decay much slower than those of the remaining group of species. The first group of species we label as the slow species while the second are the fast species. The characteristic timescales of these two are τ_s and τ_f, respectively. The vectors of molecular populations, of macroscopic concentrations, and of fluctuations can be divided into subpopulations of slow and fast species: $\vec{n} = (\vec{n}_s, \vec{n}_f)$, $\vec{\phi} = (\vec{\phi}_s, \vec{\phi}_f) = (\tau_s \vec{x}_s, \tau_f \vec{x}_f)$ and $\vec{\eta} = (\vec{\eta}_s, \vec{\eta}_f) = (\tau_s^{1/2} \vec{\epsilon}_s, \tau_f^{1/2} \vec{\epsilon}_f)$. The ansatz Eq. (3) can then be written as

$$\frac{1}{\tau_s} \frac{\vec{n}_s}{\Omega} = \vec{x}_s + (\Omega \tau_s)^{-1/2} \vec{\epsilon}_s, \quad \frac{1}{\tau_f} \frac{\vec{n}_f}{\Omega} = \vec{x}_f + (\Omega \tau_f)^{-1/2} \vec{\epsilon}_f.$$

$$(4)$$

We now use this ansatz to derive an expansion of the CME in powers of the small parameter $\Omega^{-1/2}$ valid in

the case of large volumes or, equivalently, for large copy numbers of molecules. For convenience the components of the various vectors will be denoted as follows: $\vec{x}_s = (x_1, \ldots, x_{N_s})^T$, $\vec{x}_f = (x_{N_s+1}, \ldots, x_N)^T$, $\vec{\epsilon}_s = (\epsilon_1, \ldots, \epsilon_{N_s})^T$ and $\vec{\epsilon}_f = (\epsilon_{N_s+1}, \ldots, \epsilon_N)^T$. The probability distribution of molecular populations $P(\vec{n}, t)$ is hence forth changed into the distribution of slow and fast fluctuations $\Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, t)$. In what follows we consider how the time derivative, the step operator, and the microscopic rate function vector which compose the CME Eq. (2) transform under the proposed ansatz Eq. (4).

A. Transformation of the time derivative

Using the change of variable theorem the time derivative can be written as

$$\frac{\partial}{\partial t} P(\vec{n}, t) = \left(\frac{\partial}{\partial t} + \nabla_s^T \frac{d\vec{\epsilon}_s}{dt}\bigg|_{\vec{n}} + \nabla_f^T \frac{d\vec{\epsilon}_f}{dt}\bigg|_{\vec{n}} \right) \Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, t),$$

(5)

where $\nabla_s = (\partial/\partial\epsilon_1, \ldots, \partial/\partial\epsilon_{N_s})^T$ and $\nabla_f = (\partial/\partial\epsilon_{N_s+1}, \ldots, \partial/\partial\epsilon_N)^T$. The time derivatives at constant $\vec{n}$, denoted by $\cdot|_{\vec{n}}$, can be computed from the ansatz Eq. (4), leading to

$$\frac{\partial}{\partial t} P(\vec{n}, t) = \left(\frac{\partial}{\partial t} - \sqrt{\Omega\tau_s} \nabla_s^T \frac{d\vec{x}_s}{dt} - \sqrt{\Omega\tau_f} \nabla_f^T \frac{d\vec{x}_f}{dt} \right) \times \Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, t).$$

(6)

B. Transformation of the step operator

By the definition of the step operator we have

$$\prod_{i=1}^{N} E_i^{-S_{ij}} g(n_1, \ldots, n_N)$$
$$= g(n_1 - (\underline{S}_s)_{1,j}, \ldots, n_{N_s} - (\underline{S}_s)_{N_s,j},$$
$$n_{N_s+1} - (\underline{S}_f)_{1,j}, \ldots, n_N - (\underline{S}_f)_{N_f,j}),$$

(7)

where we have partitioned the stoichiometric matrix into two parts: $(\underline{S}_s)_{ij} = (\underline{S})_{ij}$ for $1 \leqslant i \leqslant N_s$ (the stoichiometric matrix for the slow species) and $(\underline{S}_f)_{ij} = (\underline{S})_{i+N_s,j}$ for $1 \leqslant i \leqslant N_f$ (the stoichiometric matrix for the fast species). Note that g denotes some general function of the molecule numbers. Applying the ansatz Eq. (4) it follows that the above equation can be written as

$$\prod_{i=1}^{N} E_i^{-S_{ij}} g(\epsilon_1, \ldots, \epsilon_N)$$
$$= g(\epsilon_1 - (\Omega\tau_s)^{-1/2}(\underline{S}_s)_{1,j}, \ldots, \epsilon_{N_s} - (\Omega\tau_s)^{-1/2}(\underline{S}_s)_{N_s,j},$$
$$\epsilon_{N_s+1} - (\Omega\tau_f)^{-1/2}(\underline{S}_f)_{1,j}, \ldots, \epsilon_N - (\Omega\tau_f)^{-1/2}(\underline{S}_f)_{N,j}).$$

(8)

The right-hand side of the above equation can be Taylor expanded from which it follows that the step operator for the jth reaction has the following formal expansion in powers of the inverse square root of the system volume:

$$\prod_{i=1}^{N} E_i^{-S_{ij}} - 1$$
$$= \sum_{k=1}^{\infty} \frac{(-1)^k \Omega^{-k/2}}{k!} \left[\left(\tau_s^{-1/2} \nabla_s^T \underline{S}_s + \tau_f^{-1/2} \nabla_f^T \underline{S}_f \right)_j \right]^k. \quad (9)$$

C. Transformation of the microscopic rate function vector

We consider four general types of elementary reactions depending on the order j of the reaction, for which the microscopic rate functions have been rigorously derived from microscopic considerations [14,15]: (1) $\hat{f}_j(\vec{n}, \Omega) = k_j$, for a zeroth-order reaction by which a species is input into a compartment; (2) $\hat{f}_j(\vec{n}, \Omega) = k_j n_u \Omega^{-1}$, for a first-order unimolecular reaction describing the decay of some species u; (3) $\hat{f}_j(\vec{n}, \Omega) = k_j n_u (n_u - 1) \Omega^{-2}$ for a second-order bimolecular reaction between two molecules of the same species u; and (4) $\hat{f}_j(\vec{n}, \Omega) = k_j n_u n_v \Omega^{-2}$, for a second-order bimolecular reaction between two molecules of different species, u and v.

With each of these cases one can also associate a macroscopic rate function f, i.e., the rate of reaction as given by the corresponding rate equations. For the four cases discussed above these are (1) $f_j(\vec{x}_s, \vec{x}_f) = k_j$; (2) $f_j(\vec{x}_s, \vec{x}_f) = k_j \phi_u$; (3) $f_j(\vec{x}_s, \vec{x}_f) = k_j \phi_u^2$; and (4) $f_j(\vec{x}_s, \vec{x}_f) = k_j \phi_u \phi_v$, where $\phi_i = \tau_s x_i$ if $1 \leqslant i \leqslant N_s$ and $\phi_i = \tau_f x_i$ if $N_s + 1 \leqslant i \leqslant N$.

Given the microscopic and macroscopic rate functions for the four elementary reactions, one can write the former in terms of the latter, leading to the general result

$$\hat{f}_j(\vec{n}, \Omega) = f_j(\vec{x}_s, \vec{x}_f) + (\Omega\tau_s)^{-1/2} \nabla_{\vec{x}_s}^T f_j(\vec{x}_s, \vec{x}_f) \vec{\epsilon}_s$$
$$+ (\Omega\tau_f)^{-1/2} \nabla_{\vec{x}_f}^T f_j(\vec{x}_s, \vec{x}_f) \vec{\epsilon}_f + O(\Omega^{-1}), \quad (10)$$

where $\nabla_{\vec{x}_s}$ and $\nabla_{\vec{x}_f}$ denote the vectors of derivatives with respect to the components of the vectors $\vec{x}_s$ and $\vec{x}_f$, respectively.

This formula can be easily verified by considering each of the four elementary reactions discussed above. For example for case of a second-order jth reaction involving a slow species u and a fast species v, we have

$$\hat{f}_j(n_u, n_v, \Omega) = k_j \frac{n_u}{\Omega} \frac{n_v}{\Omega} = k_j [\tau_s \tau_f x_u x_v + (\Omega\tau_s)^{-1/2} \tau_s$$
$$\times \tau_f x_v \epsilon_u + (\Omega\tau_f)^{-1/2} \tau_s \tau_f x_u \epsilon_v] + O(\Omega^{-1})$$
$$= f_j(x_u, x_v) + (\Omega\tau_s)^{-1/2} \frac{\partial f_j(x_u, x_v)}{\partial x_u} \epsilon_u$$
$$+ (\Omega\tau_f)^{-1/2} \frac{\partial f_j(x_u, x_v)}{\partial x_v} \epsilon_v + O(\Omega^{-1}), \quad (11)$$

where we used the definitions of the microscopic and macroscopic rate functions given above and the ansatz Eq. (4).

D. The transformed CME

Substituting Eqs. (6), (9), and (10) in the CME [Eq. (2)] and rescaling time by the slow timescale $\tau = t/\tau_s$, we obtain the transformed CME:

$$\frac{\partial}{\partial \tau} \Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau) - (\Omega\tau_s)^{1/2} \nabla_s^T \left[\frac{\partial \vec{x}_s}{\partial \tau} - \underline{S}_s \vec{f}(\vec{x}_s, \vec{x}_f) \right] \Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau)$$

$$- (\Omega\tau_f)^{1/2} \nabla_f^T \left[\frac{\partial \vec{x}_f}{\partial \tau} - \gamma \underline{S}_f \vec{f}(\vec{x}_s, \vec{x}_f) \right] \Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau)$$

$$= \Omega^0 (\gamma \mathcal{L}_f + \gamma^{1/2} \mathcal{L}_{\text{int}} + \mathcal{L}_s) \Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau) + O(\Omega^{-1/2}),$$

(12)

where we have introduced the nondimensional ratio of slow and fast timescales

$$\gamma = \frac{\tau_s}{\tau_f}.$$

(13)

Note that the order Ω^0 is defined by the operators

$$\mathcal{L}_s = -\nabla_s^T \underline{\tilde{J}}_s \vec{\epsilon}_s + \tfrac{1}{2} \nabla_s^T \underline{\tilde{D}}_s \nabla_s, \tag{14a}$$

$$\mathcal{L}_f = -\nabla_f^T \underline{\tilde{J}}_f \vec{\epsilon}_f + \tfrac{1}{2} \nabla_f^T \underline{\tilde{D}}_f \nabla_f, \tag{14b}$$

$$\mathcal{L}_{\text{int}} = -\nabla_f^T \underline{\tilde{J}}_{fs} \vec{\epsilon}_s + \tfrac{1}{2} \nabla_s^T \underline{\tilde{D}}_{sf} \nabla_f$$
$$- \nabla_s^T \underline{\tilde{J}}_{sf} \vec{\epsilon}_f + \tfrac{1}{2} \nabla_f^T \underline{\tilde{D}}_{fs} \nabla_s, \tag{14c}$$

where $\underline{\tilde{J}}_s = \underline{S}_s (\nabla_{\tilde{x}_s} \vec{f}^T)^T$, $\underline{\tilde{J}}_f = \underline{S}_f (\nabla_{\tilde{x}_f} \vec{f}^T)^T$, $\underline{\tilde{J}}_{sf} = \underline{S}_s (\nabla_{\tilde{x}_f} \vec{f}^T)^T$, and $\underline{\tilde{J}}_{fs} = \underline{S}_f (\nabla_{\tilde{x}_s} \vec{f}^T)^T$ as well as $\underline{\tilde{D}}_s = \underline{S}_s \text{diag}(\vec{f}) \underline{S}_s^T$, $\underline{\tilde{D}}_f = \underline{S}_f \text{diag}(\vec{f}) \underline{S}_f^T$ and $\underline{\tilde{D}}_{sf} = \underline{\tilde{D}}_{fs}^T = \underline{S}_s \text{diag}(\vec{f}) \underline{S}_f^T$.

III. DERIVATION OF THE SLOW-SCALE LINEAR NOISE APPROXIMATION

A. Deterministic QSSA

The leading order terms, $O(\Omega^{1/2})$, of Eq. (12) describe the dynamics of macroscopic concentrations and is given by the coupled set of REs:

$$\frac{d\vec{x}_s}{d\tau} = \underline{S}_s \vec{f}(\vec{x}_s, \vec{x}_f), \tag{15a}$$

$$\frac{1}{\gamma} \frac{d\vec{x}_f}{d\tau} = \underline{S}_f \vec{f}(\vec{x}_s, \vec{x}_f). \tag{15b}$$

The presence of timescale separation is reflected by the large parameter γ diminishing the time derivative in Eq. (15b). Such a set of equations present a special case in singular perturbation theory, where Eqs. (15a) and (15b) for the slow and fast variables are typically referred to as the degenerate and adjoined systems, respectively [11]. Tikhonov's first theorem [16,17] states that a simplification of the above equations under timescale separation conditions is possible whenever certain requirements are met: (1) the solutions of both the degenerate and adjoined systems [Eqs. (15)] are unique and their right-hand sides are continuous functions; (2) the root $\vec{x}_f = h(\vec{x}_s, \tau)$ is the stable solution of the adjoined system; and (3) the initial values $\vec{x}_f(\tau = 0)$ are in the domain of influence of the solution as in (2). Whenever these prerequisites are met, the solution of the full system (15) for $\vec{x}_s$ tends to the solution of the reduced system

$$\frac{d\vec{x}_s}{d\tau} = \underline{S}_s \vec{f}(\vec{x}_s, h(\vec{x}_s)), \tag{16}$$

in the limit of timescale separation, i.e., $\gamma^{-1} \to 0$. Note that $\vec{x}_f = h(\vec{x}_s)$ is the solution of $\underline{S}_f \vec{f}(\vec{x}_s, \vec{x}_f) = 0$.

These requirements are typically fulfilled for the biochemical networks of interest. This is since the chemical transformation scheme (1) is formulated for elementary reactions, which are bimolecular or simpler, the right-hand sides of Eqs. (15) are continuous polynomial functions of the second order at most. For monostable networks, the rate equations admit a single steady state which is the same for the full and the reduced REs, i.e., Eqs. (15) and (16). It is therefore clear that all the solutions will tend to this state with time, quicker for fast variables and slower for the slow ones.

B. Adiabatic elimination of stochastic variables using the projection operator formalism

In the previous subsection, we reviewed how imposing timescale separation on the deterministic level leads to reduced time evolution equations for the concentrations of the slow species. A related question which is the main issue of this article is: What is the reduced Fokker-Planck equation describing the fluctuations about the concentrations predicted by the deterministic QSSA?

The Fokker-Planck equation describing the fluctuations in the concentrations of both fast and slow species is given by the $O(\Omega^0)$ terms in Eq. (12); this has the form

$$\frac{\partial}{\partial \tau} \Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau) = (\gamma \mathcal{L}_f + \gamma^{1/2} \mathcal{L}_{\text{int}} + \mathcal{L}_s) \Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau). \tag{17}$$

From the definitions of the operators Eq. (14), it can be seen that $\mathcal{L}_f$ acts only on the fast variables, $\mathcal{L}_s$ acts only on the slow variables and $\mathcal{L}_{\text{int}}$ acts only on both slow and fast variables. Hence the three operators describe processes evolving on fast, slow, and intermediate timescales, respectively. The dimensionless parameter γ here weights the degree of timescale separation in the system. The fact that γ is the same as for the deterministic QSSA implies that the conditions for timescale separation of the stochastic variables, in the limit of large Ω, are exactly the same as those for the REs.

For well-separated timescales, i.e., the case $\gamma \gg 1$, we are typically interested in the probability distribution of slow variables, $\Pi(\vec{\epsilon}_s, \tau) = \int d\vec{\epsilon}_f \Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau)$. Projection operator methods have been found useful in facilitating the adiabatic elimination of fast variables from stochastic descriptions [18,19]. Here we use one such method to rigorously obtain a reduced Fokker-Planck equation for the fluctuations in the concentrations of the slow variables.

The main idea behind the projection operator method is that one specifies the quasi-steady-state probability distribution $\pi(\vec{\epsilon}_f)$ of fast fluctuations, which is determined by the infinite γ-limit of Eq. (17):

$$\mathcal{L}_f \pi(\vec{\epsilon}_f) = 0. \tag{18}$$

The reduction is then carried out defining the operator

$$(\mathcal{P}\Pi)(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau) = \pi(\vec{\epsilon}_f) \int d\vec{\epsilon}_f \Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau) = \pi(\vec{\epsilon}_f) \Pi(\vec{\epsilon}_s, \tau), \tag{19}$$

projecting the probability distribution $\Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau)$ onto the distribution of fast fluctuations evaluated at steady state. Note that the above definition satisfies the relation $\mathcal{P}^2 = \mathcal{P}$ and hence $\mathcal{P}$ is indeed a projector.

In what follows we use the forms we have derived for the operators in the Fokker-Planck Eq. (17), i.e., those given by Eq. (14), to deduce three properties of the projection operator. Given these properties we then show how the projection operator applied to Eq. (17) leads to a reduced Fokker-Planck equation in the slow variables only.

1. Properties of the projection operator

In this subsection we will show that the following properties hold:

$$\mathcal{P}\mathcal{L}_s = \mathcal{L}_s \mathcal{P}, \tag{20a}$$

$$\mathcal{P}\mathcal{L}_f = \mathcal{L}_f\mathcal{P} = 0, \tag{20b}$$

$$\mathcal{P}\mathcal{L}_{\text{int}}\mathcal{P} = 0. \tag{20c}$$

First, property (20a) follows from the fact that the projection operator $\mathcal{P}$, as defined by Eq. (19), acts only on the fast variables $\vec{\epsilon}_f$, whereas $\mathcal{L}_s$, see Eq. (14a), acts only on the slow variables $\vec{\epsilon}_s$, and hence the two operators $\mathcal{P}$ and $\mathcal{L}_s$ commute.

Second, we show that both equalities of property (20b) are satisfied. Considering the left-hand side, we obtain

$$\mathcal{P}\mathcal{L}_f = \pi(\vec{\epsilon}_f)\int d\vec{\epsilon}_f\,\nabla_f{}^T(\cdots) = 0, \tag{21}$$

since $\mathcal{L}_f$, as given by Eq. (14b), has the form of a divergence in the fast variables and hence by the partial integration lemma, its integral vanishes in the absence of boundary terms. By considering the right-hand side, we have

$$\mathcal{L}_f\mathcal{P} = (\mathcal{L}_f\pi(\vec{\epsilon}_f))\int d\vec{\epsilon}_f = 0, \tag{22}$$

by the quasi-steady-state condition, $\mathcal{L}_f\pi = 0$.

The third property (20c) can be obtained as follows. The first, second, and fourth terms of $\mathcal{L}_{\text{int}}$ as given by Eq. (14c) have the form of a divergence in the fast variables, and hence by the partial integration lemma they give no contribution to $\mathcal{P}\mathcal{L}_{\text{int}}\mathcal{P}$. The third term of Eq. (14c) also does not contribute albeit for a different reason than for the three other terms just discussed:

$$\mathcal{P}\mathcal{L}_{\text{int}}\mathcal{P} = \pi(\vec{\epsilon}_f)\int d\vec{\epsilon}_f\big(-\nabla_s^T\,\underline{\tilde{J}}_{sf}\vec{\epsilon}_f\big)\pi(\vec{\epsilon}_f)\int d\vec{\epsilon}_f$$
$$= -\pi(\vec{\epsilon}_f)\nabla_s^T\,\underline{\tilde{J}}_{sf}\langle\vec{\epsilon}_f\rangle_\pi\int d\vec{\epsilon}_f = 0. \tag{23}$$

Here we have applied the partial integration lemma and then used the fact that $\langle\vec{\epsilon}_f\rangle_\pi = 0$. The latter follows from the explicit form of $\mathcal{L}_f$ [see Eq. (14)], which implies that the solution of the Fokker-Planck equation for the fast fluctuations $\mathcal{L}_f\pi(\vec{\epsilon}_f) = 0$ is a Gaussian distribution centered about zero.

2. Derivation of the projection operator method

We will now use the three properties of the projection operator just derived to obtain a reduced Fokker-Planck equation. Our approach in this subsection follows that of Gardiner [19,20].

We define $\mathcal{Q} = 1 - \mathcal{P}$ and the following two functions:

$$v(\tau) \equiv \mathcal{P}\Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau), \quad w(\tau) \equiv \mathcal{Q}\Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau), \tag{24}$$

together with their Laplace transforms:

$$\tilde{v}(s) = \int_0^\infty d\tau\,e^{-s\tau}v(\tau), \quad \tilde{w}(s) = \int_0^\infty d\tau\,e^{-s\tau}w(\tau). \tag{25}$$

The latter has the distinct advantage that instead of dealing with differential equations we obtain a set of algebraic equations. Using Eqs. (24) and (25) together with Eq. (17) we find

$$s\tilde{v}(s) - v(0) = \mathcal{L}_s\tilde{v}(s) + \gamma^{1/2}\mathcal{P}\mathcal{L}_{\text{int}}\tilde{w}(s), \tag{26a}$$

$$s\tilde{w}(s) - w(0) = (\mathcal{L}_s + \gamma\mathcal{L}_f + \gamma^{1/2}\mathcal{Q}\mathcal{L}_{\text{int}})\tilde{w}(s)$$
$$+ \gamma^{1/2}\mathcal{Q}\mathcal{L}_{\text{int}}\tilde{v}(s). \tag{26b}$$

Note that use has been made of the properties (20a) and (20b). Solving for $\tilde{v}(s)$, we obtain

$$s\tilde{v}(s) - v(0) - \gamma^{1/2}\mathcal{P}\mathcal{L}_{\text{int}}\mathcal{D}(\gamma)w(0)$$
$$= [\mathcal{L}_s + \gamma\mathcal{P}\mathcal{L}_{\text{int}}\mathcal{D}(\gamma)\mathcal{L}_{\text{int}}]\tilde{v}(s), \tag{27}$$

where we have used definition (24) and property (20c) and introduced $\mathcal{D}(\gamma) = (s - \mathcal{L}_s - \gamma\mathcal{L}_f - \gamma^{1/2}\mathcal{Q}\mathcal{L}_{\text{int}})^{-1}$. From the above equation one can draw the limit $\gamma \to \infty$ for which $\mathcal{D}(\gamma) \sim -(\gamma\mathcal{L}_f)^{-1}$:

$$\frac{\partial}{\partial\tau}v(\tau) = \big[\mathcal{L}_s - \mathcal{P}\mathcal{L}_{\text{int}}\mathcal{L}_f^{-1}\mathcal{L}_{\text{int}}\big]v(\tau), \tag{28}$$

where we have inverted the Laplace transform Eq. (25). Note that due to the vanishing of the third term on the left-hand side of Eq. (27) this asymptotic limit is Markovian and hence does not require the knowledge of the initial distribution $w(0)$ of the fast fluctuations. Using $v(\tau) = \pi(\vec{\epsilon}_f)\Pi(\vec{\epsilon}_s, \tau)$ and integrating over the fast fluctuations $\vec{\epsilon}_f$ we obtain

$$\frac{\partial\Pi(\vec{\epsilon}_s, \tau)}{\partial\tau} = \mathcal{L}'\Pi(\vec{\epsilon}_s, \tau), \tag{29a}$$

$$\mathcal{L}' = \mathcal{L}_s - \big\langle\mathcal{L}_{\text{int}}\mathcal{L}_f^{-1}\mathcal{L}_{\text{int}}\big\rangle_\pi, \tag{29b}$$

where the angled brackets with subscript π in Eq. (29b) denote the statistical average over the steady-state probability distribution $\pi(\vec{\epsilon}_f)$ of fast fluctuations.

3. Derivation of the slow-scale linear noise approximation

The above equation is a generic form for the Fokker-Planck equation for the slow fluctuations $\vec{\epsilon}_s$ under timescale separation conditions. What remains is to explicitly evaluate the average over $\pi(\vec{\epsilon}_f)$ such that we obtain a closed-form expression for the reduced Fokker-Planck equation. We now show these evaluation steps in detail.

Using $\mathcal{L}_{\text{int}}$ as given by Eq. (14) together with Eq. (29b) we can deduce the form of the reduced Fokker-Planck operator:

$$\mathcal{L}' = \mathcal{L}_s - \nabla_s^T\,\underline{\tilde{J}}_{sf}\langle\vec{\epsilon}_f\mathcal{L}_f^{-1}\nabla_f{}^T\rangle_\pi\,\underline{\tilde{J}}_{fs}\vec{\epsilon}_s$$
$$- \nabla_s^T\,\underline{\tilde{J}}_{sf}\langle\vec{\epsilon}_f\mathcal{L}_f^{-1}\vec{\epsilon}_f{}^T\rangle_\pi\,\underline{\tilde{J}}_{sf}^T\nabla_s$$
$$+ \tfrac{1}{2}\nabla_s^T\,\underline{\tilde{J}}_{sf}\langle\vec{\epsilon}_f\mathcal{L}_f^{-1}\nabla_f{}^T\rangle_\pi\,\underline{\tilde{D}}_{fs}\nabla_s$$
$$+ \tfrac{1}{2}\nabla_s^T\,\underline{\tilde{D}}_{sf}\langle\vec{\epsilon}_f\mathcal{L}_f^{-1}\nabla_f{}^T\rangle_\pi^T\,\underline{\tilde{J}}_{sf}^T\nabla_s. \tag{30}$$

Note that terms which have ∇_f^T to the left do not contribute to the reduced operator and hence are missing from the above equation.

We proceed by evaluating the two distinct correlators appearing in the above expression explicitly. We shall make use of the identity

$$\int_0^\infty du\,e^{\mathcal{L}_f u} = \mathcal{L}_f^{-1}e^{\mathcal{L}_f u}\big|_0^\infty = -\mathcal{L}_f^{-1}(1 - \mathcal{P}), \tag{31}$$

which can be verified from straightforward integration and the fact that $\mathcal{P}\Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau) = \lim_{u\to\infty}e^{\mathcal{L}_f u}\Pi(\vec{\epsilon}_s, \vec{\epsilon}_f, \tau)$ [19]. Using

the fact that $\mathcal{P}\vec{\epsilon}_f^T \pi = 0$, we can write

$$
\begin{aligned}
\langle \vec{\epsilon}_f \, \mathcal{L}_f^{-1} \, \vec{\epsilon}_f^{\,T} \rangle_\pi &= \int d\vec{\epsilon}_f \, \vec{\epsilon}_f \, \mathcal{L}_f^{-1} (1 - \mathcal{P}) \vec{\epsilon}_f^{\,T} \pi \\
&= - \int_0^\infty du \int d\vec{\epsilon}_f \, \vec{\epsilon}_f \, e^{\mathcal{L}_f u} \, \vec{\epsilon}_f^{\,T} \, \pi \\
&= - \int_0^\infty du \langle \vec{\epsilon}_f(u) \vec{\epsilon}_f^{\,T}(0) \rangle_\pi.
\end{aligned} \tag{32}
$$

Note that in the third step we have taken into account that $\vec{\epsilon}_f^T \pi(\vec{\epsilon}_f, u) = e^{\mathcal{L}_f u} \vec{\epsilon}_f^{\,T} \pi(\vec{\epsilon}_f, 0)$ is a solution to $\partial_u \pi = \mathcal{L}_f \pi$ with the initial condition $\vec{\epsilon}_f^{\,T} \pi(0)$. One can utilize the Fourier transform of the autocorrelation matrix $\langle \vec{\epsilon}_f(u) \vec{\epsilon}_f^{\,T}(0) \rangle_\pi = \int d\omega/(2\pi) e^{i\omega u} \underline{S}_f(\omega)$ to find that

$$
\langle \vec{\epsilon}_f \, \mathcal{L}_f^{-1} \, \vec{\epsilon}_f^{\,T} \rangle_\pi = -\tfrac{1}{2} \underline{S}_f(0) = -\tfrac{1}{2} \tilde{\underline{J}}_f^{-1} \tilde{\underline{D}}_f \tilde{\underline{J}}_f^{-T}. \tag{33}
$$

Similarly, one can show that

$$
\langle \vec{\epsilon}_f \, \mathcal{L}_f^{-1} \, \nabla_f^T \rangle_\pi = - \int_0^\infty du \, e^{\underline{J}_f u} \langle \vec{\epsilon}_f \nabla_f^T \rangle_\pi = -\tilde{\underline{J}}_f^{-1}. \tag{34}
$$

Here $\langle \vec{\epsilon}_f \nabla_f^T \rangle_\pi = -\underline{I}$ with identity matrix $\underline{I}$, is evaluated by partial integration. Plugging Eqs. (33) and (34) into Eq. (30) gives a closed-form expression for the Fokker-Planck operator of the reduced probability distribution $\Pi(\epsilon_s, \tau)$:

$$
\mathcal{L}' = -\nabla_s^T \tilde{\underline{J}} \vec{\epsilon}_s + \tfrac{1}{2} \nabla_s^T \tilde{\underline{D}}_{ss} \nabla_s, \tag{35a}
$$

$$
\tilde{\underline{J}} = \tilde{\underline{J}}_s - \tilde{\underline{J}}_{sf} \tilde{\underline{J}}_f^{-1} \tilde{\underline{J}}_{fs}, \tag{35b}
$$

$$
\tilde{\underline{D}}_{ss} = \tilde{\underline{D}}_s + \tilde{\underline{J}}_{sf} \big(\tilde{\underline{J}}_f^{-1} \tilde{\underline{D}}_f \tilde{\underline{J}}_f^{-T} \big) \tilde{\underline{J}}_{sf}^T
$$
$$
- \tilde{\underline{J}}_{sf} \tilde{\underline{J}}_f^{-1} \tilde{\underline{D}}_{fs} - \big(\tilde{\underline{J}}_{sf} \tilde{\underline{J}}_f^{-1} \tilde{\underline{D}}_{fs} \big)^T. \tag{35c}
$$

We complete our analysis by changing to natural variables of concentration fluctuations $\hat{\vec{\eta}}_s = (\tau_s/\Omega)^{1/2} \vec{\epsilon}_s$, of real time $t = \tau \tau_s$ and concentrations $\vec{\phi}_{s,f} = \tau_{s,f} \vec{x}_{s,f}$ to finally yield the reduced Fokker-Planck equation for the fluctuations in the slow species:

$$
\frac{\partial}{\partial t} P(\hat{\vec{\eta}}_s, t) = \big(-\hat{\mathbf{V}}_s^T \underline{J} \hat{\vec{\eta}}_s + \tfrac{1}{2} \hat{\mathbf{V}}_s^T \underline{D}_{ss} \hat{\mathbf{V}}_s \big) P(\hat{\vec{\eta}}_s, t), \tag{36a}
$$

$$
\underline{J} = \underline{J}_s - \underline{J}_{sf} \underline{J}_f^{-1} \underline{J}_{fs}, \tag{36b}
$$

$$
\underline{D}_{ss} = \underline{D}_s + \underline{J}_{sf} \big(\underline{J}_f^{-1} \underline{D}_f \underline{J}_f^{-T} \big) \underline{J}_{sf}^T
$$
$$
- \underline{J}_{sf} \underline{J}_f^{-1} \underline{D}_{fs} - \big(\underline{J}_{sf} \underline{J}_f^{-1} \underline{D}_{fs} \big)^T, \tag{36c}
$$

where $\hat{\mathbf{V}}_s$ denotes the derivative with respect to $\hat{\vec{\eta}}_s$. The coefficient matrices in the above expressions are given by $\underline{J}_s = \underline{S}_s (\nabla_{\vec{\phi}_s} \vec{f}^T)^T$, $\underline{J}_f = \underline{S}_f (\nabla_{\vec{\phi}_f} \vec{f}^T)^T$ and $\underline{J}_{sf} = \underline{S}_s (\nabla_{\vec{\phi}_f} \vec{f}^T)^T$, $\underline{J}_{fs} = \underline{S}_f (\nabla_{\vec{\phi}_s} \vec{f}^T)^T$ as well as $\underline{D}_s = \Omega^{-1} \underline{S}_s \underline{F} \, \underline{S}_s^T$, $\underline{D}_f = \Omega^{-1} \underline{S}_f \underline{F} \, \underline{S}_f^T$, $\underline{D}_{sf} = \underline{D}_{fs}^T = \Omega^{-1} \underline{S}_s \underline{F} \, \underline{S}_f^T$ and $\underline{F} = \text{diag}(\vec{f})$.

We note that the slow-scale Jacobian (36b) coincides with the reduced Jacobian as obtained from the macroscopic QSSA equations, i.e., Eqs. (16), as shown in Ref. [10]. It is also important to note that the reduced diffusion matrix $\underline{D}_{ss}$ admits the representation

$$
\underline{D}_{ss} = \Omega^{-1} (\underline{A} - \underline{B})(\underline{A} - \underline{B})^T, \tag{37}
$$

where $\underline{A} = \underline{S}_s \sqrt{\underline{F}}$ and $\underline{B} = \underline{J}_{sf} \underline{J}_f^{-1} \underline{S}_f \sqrt{\underline{F}}$. From this representation it can be immediately deduced that the reduced matrix $\underline{D}_{ss}$ is symmetric and positive semidefinite, two crucial properties of the diffusion matrices for all Fokker-Planck equations [12]. Using standard methods [19] one can also obtain the Langevin equations corresponding to the slow-scale Fokker-Planck equation (36a). These are given by

$$
\frac{d}{dt} \hat{\vec{\eta}}_s(t) = \underline{J} \hat{\vec{\eta}}_s(t) + \Omega^{-1/2} \big(\underline{S}_s - \underline{J}_{sf} \underline{J}_f^{-1} \underline{S}_f \big) \sqrt{\underline{F}} \vec{\Gamma}(t), \tag{38}
$$

where the R-dimensional vector $\vec{\Gamma}(t)$ is white Gaussian noise defined by $\langle \Gamma_i(t) \rangle = 0$ and $\langle \Gamma_i(t) \Gamma_j(t') \rangle = \delta_{i,j} \delta(t - t')$. Equations (36a) and (38) constitute equivalent forms of the LNA under timescale separation conditions, with the latter being particularly useful for Monte Carlo simulation purposes. This reduced LNA is precisely the ssLNA introduced in Ref. [10].

IV. REACTION NETWORKS WITH SLOW AND FAST REACTIONS

Besides the QSSA there exists another popular method to eliminate the fast variables from the deterministic REs; this is the rapid-equilibrium approximation [11,21,22]. While the QSSA considers the situation in which there exists slow and fast species whose transients decay on well-separated timescales, the rapid-equilibrium approximation divides the set of reactions into groups of slow and fast reactions where the latter determines the equilibrium of the fast species alone [11].

In the stochastic case there exist a variety of methods separating fast and slow reactions. Examples are the "nested stochastic simulation algorithm" [23], "slow-scale stochastic simulation algorithm" [24], the "stochastic partial equilibrium assumption" [25], and the "quasiequilibrium approximation" [26]. Confusingly, also the "stochastic quasi-steady-state assumption" of Rao and Arkin [27] is valid only in the limit of slow and fast reactions [8,9]. All of these approaches have in common that the microscopic rate functions, or equivalently the propensities are rearranged into two groups: those associated with R_s slow reactions which occur rarely over a long period of time, and those associated with $(R - R_s)$ fast reactions which occur frequently over a short period of time. Then we can define a constant μ such that

$$
\mu = \frac{\max(\hat{f}_1, \hat{f}_2, \ldots, \hat{f}_{R_s})}{\min(\hat{f}_{R_s+1}, \hat{f}_{R_s+2}, \ldots, \hat{f}_R)} \gg 1, \tag{39}
$$

holds. For the sake of this article we will refer to approximations utilizing the above criterion as "rapid equilibrium approximations." Since our present derivation of the ssLNA is based only on the assumption of the presence of slow and fast species it can be used to investigate the latter approximation as a partial case.

Using the size parameter μ we can write the vector of macroscopic rate functions as

$$
\vec{f}(\vec{x}_s, \vec{x}_f) = (\vec{f}_s(\vec{x}_s, \vec{x}_f), \mu \vec{f}_f(\vec{x}_s, \vec{x}_f)), \tag{40}
$$

where $\vec{f}_s(\vec{x}_s, \vec{x}_f) = (f_1, f_2, \ldots, f_{R_s})$ are the rates of the slow reactions and $\vec{f}_f(\vec{x}_s, \vec{x}_f) = \mu^{-1}(f_{R_s+1}, f_{R_s+2}, \ldots, f_R)$ are the rates of the fast reactions rescaled by the size parameter μ. Note that here μ is determined by the infinite volume limit of

Eq. (39). We can now partition the stoichiometric matrix into block matrices

$$\underline{S} = \begin{bmatrix} \underline{S}_s \\ \underline{S}_f \end{bmatrix} = \begin{bmatrix} \underline{S}_s^{(s)} & \underline{S}_s^{(f)} \\ \underline{S}_f^{(s)} & \underline{S}_f^{(f)} \end{bmatrix}, \tag{41}$$

discriminating slow and fast reactions (superscript) as well as slow and fast species (subscript). The matrices $\underline{S}_s^{(s)}$ and $\underline{S}_f^{(f)}$ denote the stoichiometries of the slow and fast species in the slow and fast reactions, respectively, while $\underline{S}_s^{(f)}$ and $\underline{S}_f^{(s)}$ represent the stoichiometry of slow species in fast reactions and the stoichiometry of fast species in slow reactions.

A. Deterministic rapid-equilibrium approximation

The macroscopic elimination starts from the conventional REs:

$$\frac{d\vec{x}_s}{d\tau} = \mu\,\underline{S}_s^{(f)}\vec{f}_f + \underline{S}_s^{(s)}\vec{f}_s, \tag{42a}$$

$$\frac{1}{\gamma}\frac{d\vec{x}_f}{d\tau} = \mu\,\underline{S}_f^{(f)}\vec{f}_f + \underline{S}_f^{(s)}\vec{f}_s, \tag{42b}$$

which are similar to Eqs. (15) but discriminate slow and fast reactions by the size parameter μ. It is clear that μ and γ must be of the same order since slow and fast timescales are determined by the size of the rate functions.

Further, we observe that for $\mu \to \infty$ the above equations are not immediately of the form required by Tikhonov's theorem [compare Eqs. (15)], and hence the adiabatic elimination is not immediately applicable unless we impose $\underline{S}_s^{(f)} = 0$. This condition implies that the populations of slow species are not changed in fast reactions and is also imposed throughout the literature in reducing stochastic slow-fast reaction networks [24–26,28]. Setting the time derivative of the second equation to zero we can solve $\underline{S}_f^{(f)}\vec{f}_f \approx 0$ for $\vec{x}_f = h(\vec{x}_s) + O(\mu^{-1})$ to obtain the reduced system

$$\frac{d\vec{x}_s}{dt} = \underline{S}_s^{(s)}\vec{f}_s(\vec{x}_s, h(\vec{x}_s)). \tag{43}$$

In the case where the equilibrium of the fast reactions $\underline{S}_f^{(f)}\vec{f}_f = 0$ is detailed balanced, the above approximation is called the deterministic rapid-equilibrium approximation [11], i.e., the case when the fast reactions are given by a set of reversible reactions for which the forward and backward rates of each reaction cancel each other.

B. Stochastic rapid-equilibrium approximation

We can now apply the ssLNA to obtain the contribution of the fluctuations using the same conditions as used above for the deterministic rapid-equilibrium approximation. First, we make use of the condition $\underline{S}_s^{(f)} = 0$ to obtain the coefficients $\underline{J}_s = \underline{S}_s^{(s)}(\nabla_{\vec{\phi}_s}\vec{f}_s^T)^T$, $\underline{J}_f = \mu\,\underline{S}_f^{(f)}(\nabla_{\vec{\phi}_f}\vec{f}_f^T)^T + \underline{S}_f^{(s)}(\nabla_{\vec{\phi}_f}\vec{f}_s^T)^T$, $\underline{J}_{sf} = \underline{S}_s^{(s)}(\nabla_{\vec{\phi}_f}\vec{f}_s^T)^T$, and $\underline{J}_{fs} = \underline{S}_f^{(s)}(\nabla_{\vec{\phi}_s}\vec{f}_s^T)^T + \mu\,\underline{S}_f^{(f)}(\nabla_{\vec{\phi}_s}\vec{f}_f^T)^T$, which distinguish contributions from slow and fast reactions. Second, we can use these together within Eq. (36b) to obtain

the reduced Jacobian by taking the limit $\mu \to \infty$:

$$\underline{J} = \underline{S}_s^{(s)}(\nabla_{\vec{\phi}_s}\vec{f}_s^T)^T - \underline{S}_s^{(s)}(\nabla_{\vec{\phi}_f}\vec{f}_s^T)^T\big[\underline{S}_f^{(f)}(\nabla_{\vec{\phi}_f}\vec{f}_f^T)^T\big]^{-1}$$
$$\times\,\underline{S}_f^{(f)}(\nabla_{\vec{\phi}_s}\vec{f}_f^T)^T. \tag{44}$$

It can be shown that the above expression coincides with the Jacobian obtained using Eq. (43). Third, we calculate the coefficients of the noise from Eq. (37) as

$$\underline{A} = \underline{S}_s^{(s)}\sqrt{\underline{F}_s}, \quad \underline{B} = 0, \tag{45}$$

where $\underline{F}_s = \text{diag}(\vec{f}_s)$ and $\underline{F}_f = \text{diag}(\vec{f}_f)$. Note that the second equation follows from inserting $\underline{S}_f\sqrt{\underline{F}} = \underline{S}_f^{(s)}\sqrt{\underline{F}_s} + \mu\,\underline{S}_f^{(f)}\sqrt{\underline{F}_f}$ into the definition of $\underline{B}$ after Eq. (37) together with the expression for $\underline{J}_f$ and $\underline{J}_{sf}$ and taking the limit $\mu \to \infty$. This implies that $\underline{B}$ is of order $\mu^{-1/2}$, and hence the noise stemming from the fast reactions can be neglected in the limit $\mu \to \infty$. Finally, we can formulate the Langevin equations:

$$\frac{d}{dt}\hat{\vec{\eta}}_s(t) = \underline{J}\hat{\vec{\eta}}_s(t) + \Omega^{-1/2}\underline{S}_s^{(s)}\sqrt{\underline{F}_s}\vec{\Gamma}(t). \tag{46}$$

Note that these equations are consistent with those obtained using stoichiometry $\underline{S}_s^{(s)}$ and propensity vector $\vec{f}_s(\vec{x}_s, h(\vec{x}_s))$ of the reduced macroscopic equations (43). Note that while in the ssLNA [Eq. (38)] (which is consistent with QSSA conditions) both the noise in the fast and slow reactions contribute to the noise of the slow species, in the stochastic rapid-equilibrium approximation [Eq. (46)], the noise in the slow reactions alone determines that in the slow species.

The latter Langevin equation has also been obtained by Pahlajani *et al.* [28] starting from the decomposition of reactions into slow and fast categories. However, our derivation is the first to clearly show that this Langevin equation is a partial case of the ssLNA and hence is valid only over a subset of the parameter space over which the QSSA holds.

V. DISCUSSION

In this article we have shown how to rigorously reduce the linear noise approximation of the CME by using the projection operator formalism to eliminate the fast fluctuation variables. The resulting Langevin equation, Eq. (38), is in agreement with the ssLNA as has been previously deduced from intuitive arguments only [10]. The present derivation provides a rigorous basis by deriving the LNA from the system size expansion using a modified van Kampen ansatz, Eq. (4), which is applicable under conditions of timescale separation. The resulting REs, Eq. (15), and the Fokker-Planck equation (17) obtained by this approach are of a particular form which allows direct application of Tikhonov's theorem and the projection operator method, respectively. Hence by this procedure it is guaranteed that the mesoscopic elimination of the fast fluctuations is valid in exactly the same limit as the macroscopic elimination of the concentrations of the fast species by the deterministic QSSA.

For reaction networks composed of slow and fast reactions, conditions considered by the majority of available stochastic model reduction methods employing timescale separation, we have shown that in the limit of large volumes or molecule

numbers the CME can be approximated by a Langevin equation, Eq. (46), which is a special case of the ssLNA. The main advantage of using the ssLNA over the various aforementioned methods [23–28] is that the ssLNA is valid over a larger parameter range than the latter methods. The path integral approach developed in Ref. [29] also enjoys this property. However, the ssLNA enjoys the further advantage that it is available in closed form for any monostable reaction network and hence can be readily constructed from knowledge of the stoichiometric matrix and deterministic rate equations. The disadvantage of the ssLNA is that for pathways composed of some second-order reactions, the ssLNA (and the LNA on which it is based) is valid only for large enough molecule numbers [30,31]. This limitation can be lifted by consideration of higher order terms in the system size expansion; such calculations present more formidable analytical challenges than encountered in the derivation of the ssLNA and are under current investigation.

For realistic biochemical networks there may be particular parameter ranges for which the stability of the dynamics is either monostable or bistable or even oscillatory states can occur. The applicability of the method therefore depends on the type of stability realized for a particular network under consideration. Oscillatory states are found for 10% of the transcriptome and 20% of the proteome in mouse liver [32,33] and similar fractions in the human metabolome [34]. Although bimodal distributions have been observed experimentally, as, for instance, in the *lac* operon of *E. coli* [35], a recent proteome-wide study suggests that such probability distributions (which potentially indicate bistability) are quite rare [36] and similarly for the human transcriptome [37]. Despite the fact that bistable

and oscillatory properties are important for specific cellular functions, it appears that monostable networks for which the present theory has been developed are common in living cells.

We note that the LNA has been applied also to networks with limit cycles [38]. The resulting equation is still a linear Fokker-Planck equation, and hence the elimination of fast variables can be performed along the same lines as in the present derivation. However, the analysis of the resulting equation has to be carried out by means of Floquet theory due to the inherent phase diffusion in these systems [39]. For bistable systems, the underlying distribution cannot be captured by a linear Fokker-Planck equation, and hence the LNA is not applicable in this case [12]. A commonly employed procedure to eliminate fast variables in such systems is the stochastic QSSA. A recent numerical study reported considerable discrepancies in the probability distributions of full and QSSA-reduced bistable systems [40]. We expect these discrepancies to be similar to the difference in the ssLNA and rapid-equilibrium approximations discussed in this article for the monostable case. Therefore, developing a technique to rigorously reduce the CME of bistable networks remains still an open question. Since our method is devised for monostable systems and can be extended to oscillatory ones under timescale separation conditions we expect it to be of broad applicability for the study of intracellular reaction networks.

[1] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (CRC Press, London, 2007).

[2] U. Alon, Nature (London) **446**, 497 (2007).

[3] L. A. Segel and M. Slemrod, SIAM Rev. **31**, 446 (1989).

[4] C. H. Lee and H. G. Othmer, J. Math. Biol. **60**, 387 (2010).

[5] R. Grima and S. Schnell, Essays Biochem. **45**, 41 (2008).

[6] J. A. M. Janssen, J. Stat. Phys. **57**, 187 (1989).

[7] E. A. Mastny, E. L. Haseltine, and J. B. Rawlings, J. Chem. Phys. **127**, 094106 (2007).

[8] P. Thomas, A. V. Straube, and R. Grima, J. Chem. Phys. **135**, 181103 (2011).

[9] K. R. Sanft, D. T. Gillespie, and L. Petzold, IET Syst. Biol. **5**, 58 (2011).

[10] P. Thomas, A. V. Straube, and R. Grima, BMC Syst. Biol. **6**, 39 (2012).

[11] R. Heinrich and S. Schuster, *The Regulation of Cellular Systems* (Chapman & Hall, New York, 1996).

[12] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 2007).

[13] D. T. Gillespie, Annu. Rev. Phys. Chem. **58**, 35 (2007).

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

[15] D. T. Gillespie, J. Chem. Phys. **131**, 164109 (2009).

[16] A. N. Tikhonov, Mat. Sb. **31**, 575 (1952) (in Russian).

[17] W. Klonowski, Biophys. Chem. **18**, 73 (1983).

[18] R. Zwanzig, J. Chem. Phys. **33**, 1338 (1960).

[19] C. W. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer, Berlin, 2010).

[20] C. W. Gardiner, Phys. Rev. A **29**, 2814 (1984).

[21] I. H. Segel, *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady State Enzyme Systems* (Wiley, Hoboken, 1993).

[22] E. Klipp, *Systems Biology in Practice: Concepts, Implementation and Application* (Wiley, Hoboken, 2005).

[23] E. Weinan, D. Liu, and E. Vanden-Eijnden, J. Comput. Phys. **221**, 158 (2007).

[24] Y. Cao, D. T. Gillespie, and L. R. Petzold, J. Chem. Phys. **122**, 014116 (2005).

[25] Y. Cao, D. T. Gillespie, and L. R. Petzold, J. Comput. Phys. **206**, 395 (2005).

[26] J. Goutsias, J. Chem. Phys. **122**, 184102 (2005).

[27] C. V. Rao and A. P. Arkin, J. Chem. Phys. **118**, 4999 (2003).

[28] C. D. Pahlajani, P. J. Atzberger, and M. Khammash, J. Theor. Biol. **272**, 96 (2011).

[29] N. A. Sinitsyn, N. Hengartner, and I. Nemenman, Proc. Natl. Acad. Sci. USA **106**, 10546 (2009).

[30] R. Grima, BMC Syst. Biol. **3**, 101 (2009).

[31] P. Thomas, H. Matuschek, and R. Grima, PLoS ONE **7**, e38518 (2012).

[32] S. Panda *et al*., Cell **109**, 307 (2002).

[33] A. B. Reddy *et al*., Curr. Biol. **16**, 1107 (2006).

[34] R. Dallmann *et al*., Proc. Natl. Acad. Sci. USA **109**, 2625 (2012).

[35] P. J. Choi, L. Cai, K. Frieda, and X. S. Xie, Science **322**, 442 (2008).

[36] Y. Taniguchi *et al*., Science **329**, 533 (2010).

[37] C. C. Mason *et al*., BMC Genomics **12**, 98 (2011).

[38] M. Scott, B. Ingalls, and M. Kærn, Chaos **16**, 026107 (2006).

[39] R. P. Boland, T. Galla, and A. J. McKane, Phys. Rev. E **79**, 051131 (2009).

[40] A. Agarwal, R. Adams, G. C. Castellani, and H. Z. Shouval, J. Chem. Phys. **137**, 044105 (2012).