

## A study of the accuracy of moment-closure approximations for stochastic chemical kinetics

Ramon Grima

Citation: *J. Chem. Phys.* **136**, 154105 (2012); doi: 10.1063/1.3702848

View online: <http://dx.doi.org/10.1063/1.3702848>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v136/i15>

Published by the [American Institute of Physics](#).

---

### Additional information on *J. Chem. Phys.*

Journal Homepage: <http://jcp.aip.org/>

Journal Information: [http://jcp.aip.org/about/about\\_the\\_journal](http://jcp.aip.org/about/about_the_journal)

Top downloads: [http://jcp.aip.org/features/most\\_downloaded](http://jcp.aip.org/features/most_downloaded)

Information for Authors: <http://jcp.aip.org/authors>

## ADVERTISEMENT



**AIP Advances**

Special Topic Section:  
**PHYSICS OF CANCER**

Why cancer? Why physics? [View Articles Now](#)

# A study of the accuracy of moment-closure approximations for stochastic chemical kinetics

Ramon Grima

*School of Biological Sciences, University of Edinburgh, United Kingdom and Centre for Systems Biology, University of Edinburgh, United Kingdom*

(Received 26 February 2012; accepted 27 March 2012; published online 18 April 2012)

Moment-closure approximations have in recent years become a popular means to estimate the mean concentrations and the variances and covariances of the concentration fluctuations of species involved in stochastic chemical reactions, such as those inside cells. The typical assumption behind these methods is that all cumulants of the probability distribution function solution of the chemical master equation which are higher than a certain order are negligibly small and hence can be set to zero. These approximations are *ad hoc* and hence the reliability of the predictions of these class of methods is presently unclear. In this article, we study the accuracy of the two moment approximation (2MA) (third and higher order cumulants are zero) and of the three moment approximation (3MA) (fourth and higher order cumulants are zero) for chemical systems which are monostable and composed of unimolecular and bimolecular reactions. We use the system-size expansion, a systematic method of solving the chemical master equation for monostable reaction systems, to calculate in the limit of large reaction volumes, the first- and second-order corrections to the mean concentration prediction of the rate equations and the first-order correction to the variance and covariance predictions of the linear-noise approximation. We also compute these corrections using the 2MA and the 3MA. Comparison of the latter results with those of the system-size expansion shows that: (i) the 2MA accurately captures the first-order correction to the rate equations but its first-order correction to the linear-noise approximation exhibits the wrong dependence on the rate constants. (ii) the 3MA accurately captures the first- and second-order corrections to the rate equation predictions and the first-order correction to the linear-noise approximation. Hence while both the 2MA and the 3MA are more accurate than the rate equations, only the 3MA is more accurate than the linear-noise approximation across all of parameter space. The analytical results are numerically validated for dimerization and enzyme-catalyzed reactions. © 2012 American Institute of Physics. [<http://dx.doi.org/10.1063/1.3702848>]

## I. INTRODUCTION

It is well known that the chemical master equation (CME), the probabilistic description of chemical kinetics at the mesoscopic scale, can only be exactly solved for a handful of simple systems, e.g., those composed of only unimolecular reactions<sup>1–4</sup> and the (reversible or irreversible) bimolecular reaction involving two reactant species.<sup>5</sup> It can also be exactly solved for any chemical system which obeys the principle of detailed balance, which typically holds for closed systems.<sup>6</sup> However, naturally occurring chemical systems rarely fall into one of these exactly solvable cases. In particular, biochemical pathways inside cells are open chemical systems which typically involve the interaction of a large number of species via various types of elementary (unimolecular and bimolecular) interactions.

The difficulty in exactly solving the CME of a set of elementary chemical reactions stems from the fact that the time-evolution equation of the  $M$ th central moment of the probability density function solution of the CME is generally a function of the  $(M + 1)$ th central moment. This implies an infinite hierarchy of coupled equations which of course cannot be generally solved.<sup>7</sup> The only exception to this rule occurs when the CME describes a set of first-order (unimolecular) reactions, in which case the time-evolution equations for all

moments are closed. However, most biochemical systems involve a number of bimolecular reactions, and hence an approximation method to decouple some of the moment equations is needed. There are two such methods in popular use: (i) the linear-noise approximation (LNA) and (ii) moment-closure approximations. We now briefly describe each method.

The moments of the CME can be written as a series in powers of the inverse volume of the compartment in which the reactions occur. The method by which this is obtained is called the system-size expansion (SSE) of the CME.<sup>8</sup> The LNA refers to the first level of approximation, i.e., the first term of the moment expansion which is valid in the limit of very large volumes. In this case it is found that the mean concentrations (the first moments) are decoupled from the covariance of the fluctuations (the second moments) and are given by the rate equations (REs). This level of approximation also leads to time-evolution equations for the variances and covariances which are decoupled from the third moments; rather they are simply a function of the concentration solution of the REs. Hence the LNA provides an accurate means to approximately calculate the size of concentration fluctuations when compartment volumes are large or equivalently when the number of molecules is not too small. This accuracy is guaranteed for all times provided the system is monostable.

The advantage of the LNA is that it is derived from a systematic series expansion of the CME and is not based on *ad hoc* assumptions. Its disadvantage is that it may not be accurate for describing the stochastic kinetics of chemical systems characterized by low copy numbers of molecules (unless the system is composed of first-order reactions only in which case the covariance prediction of the LNA is exact for any number of molecules).

Moment-closure approximations are based on a completely different approach. Whittle<sup>9</sup> and McQuarrie<sup>10</sup> were some of the earliest proponents of the method, who applied it to stochastic non-linear models of population growth, plant competition, animal predation, and chemical kinetics. The infinite hierarchy of time-evolution equations for the moments is truncated by assuming that the  $(M + 1)$ th central moment is the same as that from some well-known distribution, for example, the normal distribution. It then follows that one can solve the time-evolution equations for the first  $M$  moments. The most popular of this class of approximations involves setting the third and all higher order cumulants to zero, i.e., that expected for a normal distribution, which leads to coupled equations for the mean concentrations and for the variances/covariances of the concentration fluctuations. In recent years, this method has appeared under the different guises of “Mass-fluctuation kinetics equations,”<sup>7</sup> “Statistical chemical kinetics equations”<sup>11</sup> and the “Two-moment approximation.”<sup>12,13</sup> In the rest of this article we shall refer to it by the latter name or its abbreviated form, 2MA. Software has also been written to facilitate the numerical solution of the 2MA and of higher order moment-closure approximations.<sup>14</sup> The foremost advantage of this method is its simplicity and the fact that unlike the LNA method, it does not appear to make implicit assumptions about the size of the molecular populations (or equivalently about the magnitude of the compartment volume). Its disadvantage is that it is based on *ad hoc* assumptions and hence it is not clear when it provides a good description and when it does not. What is known is that in the limit of large volumes, the 2MA’s prediction of mean concentrations and of the variances/covariances of concentration fluctuations matches exactly that of the REs and of the LNA, respectively.<sup>12</sup> Numerical integration of the 2MA equations for a handful of chemical systems suggests that for intermediate volumes, they may be more accurate than the RE and the LNA.<sup>12</sup> However, to date analytical proof of the postulated higher accuracy of the 2MA is missing. Indeed it is plausible that the 2MA’s accuracy is only evident for certain parameter sets but not a general feature of the method.

In this article we seek to clarify the accuracy of the predictions of the 2MA method. We utilize the SSE to calculate the mean concentrations and the variances of the concentration fluctuations according to the CME accurate to order  $\Omega^{-2}$  for two chemical systems: dimerization and an enzyme-catalyzed reaction. The mean concentration has the form  $\mu_0 + \mu_1\Omega^{-1} + \mu_2\Omega^{-2} + O(\Omega^{-3})$ , where  $\mu_0$  is the solution of the deterministic REs,  $\mu_1\Omega^{-1}$  is the first-order correction to the REs and  $\mu_2\Omega^{-2}$  is a second-order correction to the REs. The variance has the form  $\sigma_0^2\Omega^{-1} + \sigma_1^2\Omega^{-2} + O(\Omega^{-3})$ , where  $\sigma_0^2\Omega^{-1}$  is the LNA solution and  $\sigma_1^2\Omega^{-2}$  is the first-order

correction to the LNA. We also obtain series solutions of the 2MA equations. It is found that the mean concentration and variance predictions of the SSE and of the 2MA agree to order  $\Omega^{-1}$  but have different  $\Omega^{-2}$  terms. In particular the latter terms have a different dependence on the rate constants than the  $\Omega^{-2}$  terms of the series obtained using the SSE. Hence we conclude that the 2MA accurately captures the first-order correction to the mean concentration prediction of the REs but does not accurately predict the first-order correction to the variance and covariance predictions of the LNA. We also study the three-moment approximation (3MA) in which the fourth and all higher order cumulants are set to zero. Comparison of the series solutions of the 3MA with those of the SSE shows that the 3MA accurately captures the first- and second-order corrections to the RE prediction and the first-order correction to the LNA. We finally show using the multivariate SSE that our conclusions for the 2MA and the 3MA methods are correct for all chemical systems which are monostable and composed of elementary reactions.

## II. EXAMPLE A: DIMERIZATION

In this section we calculate the 2MA and the 3MA equations for a chemical species involved in a reversible dimerization reaction. The mean concentration and variance of concentration fluctuations solutions of these equations in steady-state conditions are then obtained as a series in powers of the inverse compartment volume. We also use the SSE of the CME to calculate the correct series solutions. A comparison of the two series enables us to infer the accuracy of the predictions of the 2MA and the 3MA equations.

### A. The 2MA and the 3MA approximations

We consider the reversible bimolecular association of monomer molecules (species  $A$ ) to form dimer molecules (species  $B$ ). The forward rate of reaction is  $k_1$  and the reverse rate of reaction is  $k_2$ . The reaction can be schematically represented as



The CME for this reaction is given by

$$\begin{aligned} \partial_t P(n_A, n_B, t) = & \frac{k_1}{\Omega} (E_A^{+2} E_B^{-1} - 1) n_A (n_A - 1) P(n_A, n_B, t) \\ & + k_2 (E_A^{-2} E_B^{+1} - 1) n_B P(n_A, n_B, t), \end{aligned} \quad (2)$$

where  $P(n_A, n_B, t)$  is the probability that the system at time  $t$  has  $n_A$  monomer molecules and  $n_B$  dimer molecules and  $\Omega$  is the volume of the compartment in which the reaction occurs. The step operator  $E_A^x$  when it acts on some function of  $n_A$  leads to the same function but with  $n_A$  replaced by  $n_A + x$ ; the operator  $E_B^x$  performs an equivalent operation when it acts on a function of  $n_B$ . The reader is referred to Appendix A for a concise review of CME and RE modeling of chemical reaction networks.

The reaction implies the conservation law  $n_A + 2n_B = n_T = \Omega\phi_T$ , where  $n_T$  is a constant at all times and it is equal to the

maximum number of monomers in the system. The associated constant concentration is  $\phi_T$ . By means of this relationship, we can reduce Eq. (2) to a univariate CME

$$\begin{aligned} \partial_t P(n_A, t) = & \frac{k_1}{\Omega} (E_A^{+2} - 1) n_A (n_A - 1) P(n_A, t) \\ & + \frac{k_2}{2} (E_A^{-2} - 1) (\Omega \phi_T - n_A) P(n_A, t). \end{aligned} \quad (3)$$

The mean monomer concentration, the second central moment (variance), third and fourth central moments of the monomer concentration fluctuations are formally defined as

$$\mu = \langle n_A / \Omega \rangle, \quad (4)$$

$$\sigma^2 = \langle (n_A / \Omega - \mu)^2 \rangle = \langle n_A^2 / \Omega^2 \rangle - \mu^2, \quad (5)$$

$$\eta = \langle (n_A / \Omega - \mu)^3 \rangle = \langle n_A^3 / \Omega^3 \rangle - 3\mu\sigma^2 - \mu^3, \quad (6)$$

$$\zeta = \langle (n_A / \Omega - \mu)^4 \rangle = \langle n_A^4 / \Omega^4 \rangle - 4\mu\eta - 6\sigma^2\mu^2 - \mu^4, \quad (7)$$

where the angled brackets denote the statistical average.

Multiplying both sides of Eq. (3) by  $n_A / \Omega$  and summing over  $n_A$ , after some algebra we obtain

$$\partial_t \mu = -2k_1\mu^2 + k_2(\phi_T - \mu) + \frac{2k_1\mu}{\Omega} - 2k_1\sigma^2. \quad (8)$$

Hence the time-evolution equation for the mean concentration is coupled to the variance of the concentration fluctuations. Next we obtain a time-evolution equation for the variance. We multiply both sides of Eq. (3) by  $n_A^2 / \Omega^2$  and sum over  $n_A$  which leads to

$$\begin{aligned} \partial_t \left\langle \frac{n_A^2}{\Omega^2} \right\rangle = & -\frac{4k_1}{\Omega^3} (\langle n_A \rangle - 2\langle n_A^2 \rangle + \langle n_A^3 \rangle) \\ & + \frac{2k_2}{\Omega^2} (\Omega \phi_T \langle n_A \rangle - \langle n_A^2 \rangle + \Omega \phi_T - \langle n_A \rangle). \end{aligned} \quad (9)$$

Given Eqs. (8) and (9), it follows from Eqs. (5) and (6) that we can write a time-evolution equation for the variance which reads

$$\begin{aligned} \partial_t \sigma^2 = & -2\sigma^2(4k_1\mu + k_2 - 4k_1\Omega^{-1}) \\ & - 2(2k_1\eta + \Omega^{-1}(k_2\mu - k_2\phi_T - 2k_1\mu^2) + 2k_1\mu\Omega^{-2}). \end{aligned} \quad (10)$$

In a similar manner we obtain a time-evolution equation for  $\eta$ , the third central moment. We multiply both sides of Eq. (3) by  $n_A^3 / \Omega^3$  and sum over  $n_A$  to obtain

$$\begin{aligned} \partial_t \left\langle \frac{n_A^3}{\Omega^3} \right\rangle = & \frac{k_1}{\Omega^4} (8\langle n_A \rangle - 20\langle n_A^2 \rangle + 18\langle n_A^3 \rangle - 6\langle n_A^4 \rangle) \\ & + \frac{k_2}{\Omega^3} (\Omega \phi_T (4 + 6\langle n_A \rangle + 3\langle n_A^2 \rangle) - 4\langle n_A \rangle \\ & - 6\langle n_A^2 \rangle - 3\langle n_A^3 \rangle). \end{aligned} \quad (11)$$

Given Eqs. (8), (10), and (11), it follows from Eqs. (6) and (7) that we can obtain the time-evolution equation for the third central moment

$$\begin{aligned} \partial_t \eta = & -3\eta(4k_1\mu + k_2 - 6k_1\Omega^{-1}) + \Omega^{-3}(8k_1\mu \\ & - 4\Omega(2k_1\mu^2 + k_2\mu + 5k_1\sigma^2 - k_2\phi_T) \\ & + 6\Omega^2\sigma^2(4k_1\mu - k_2) + 6k_1\Omega^3((\sigma^2)^2 - \zeta)). \end{aligned} \quad (12)$$

From Eqs. (8), (10), and (12), one notes that the time-evolution equation for the  $n$ th central moment depends on the  $(n + 1)$ th central moment which implies that it is not generally possible to solve these equations in closed form (as previously remarked in the Introduction). An approximative solution is possible by means of the 2MA and the 3MA methods, which are as follows.

The 2MA is obtained by setting the third central moment (the third-order cumulant) in Eq. (10) to zero (as expected for a Gaussian probability distribution) which leads to the following set of two coupled equations for the mean concentration and the variance

$$\begin{aligned} \partial_t \mu_{2M} = & -2k_1\mu_{2M}^2 + k_2(\phi_T - \mu_{2M}) + \frac{2k_1\mu_{2M}}{\Omega} - 2k_1\sigma_{2M}^2, \\ \partial_t \sigma_{2M}^2 = & -2\sigma_{2M}^2 \left( 4k_1\mu_{2M} + k_2 - \frac{4k_1}{\Omega} \right) \\ & - 2\Omega^{-1} \left( k_2\mu_{2M} - k_2\phi_T - 2k_1\mu_{2M}^2 + \frac{2k_1\mu_{2M}}{\Omega} \right). \end{aligned} \quad (13)$$

Note that we have replaced  $\mu$  by  $\mu_{2M}$  and  $\sigma^2$  by  $\sigma_{2M}^2$  to clearly identify that the solutions of these equations constitute the 2MA approximation.

In a similar way we can obtain the 3MA by setting the fourth central moment  $\zeta$  in Eq. (12) equal to  $3(\sigma^2)^2$  (as would be for a Gaussian probability distribution, i.e., the fourth-order cumulant is set to zero) which leads to a set of three coupled equations for the mean concentration, the variance and the third central moment

$$\begin{aligned} \partial_t \mu_{3M} = & -2k_1\mu_{3M}^2 + k_2(\phi_T - \mu_{3M}) + \frac{2k_1\mu_{3M}}{\Omega} - 2k_1\sigma_{3M}^2, \\ \partial_t \sigma_{3M}^2 = & -2\sigma_{3M}^2 \left( 4k_1\mu_{3M} + k_2 - \frac{4k_1}{\Omega} \right) \\ & - 2 \left( 2k_1\eta_{3M} + \frac{k_2\mu_{3M} - k_2\phi_T - 2k_1\mu_{3M}^2 + 2k_1\mu_{3M}}{\Omega} \right), \\ \partial_t \eta_{3M} = & -3\eta_{3M} \left( 4k_1\mu_{3M} + k_2 - \frac{6k_1}{\Omega} \right) + \frac{8k_1\mu_{3M}}{\Omega^3} \\ & - \frac{4(2k_1\mu_{3M}^2 + k_2\mu_{3M} + 5k_1\sigma_{3M}^2 - k_2\phi_T)}{\Omega^2} \\ & + \frac{6\sigma_{3M}^2(4k_1\mu_{3M} - k_2)}{\Omega} - 12k_1(\sigma_{3M}^2)^2. \end{aligned} \quad (14)$$

As previously, we have here replaced  $\mu$  by  $\mu_{3M}$ ,  $\sigma^2$  by  $\sigma_{3M}^2$ , and  $\eta$  by  $\eta_{3M}$  to clearly identify that the solutions of these equations are the 3MA approximation.

We now obtain a series solution of the 2MA equations in steady-state conditions. The form of the 2MA equations,

Eq. (13), suggests that the mean concentration and the variance can be written as a series in powers of the inverse volume  $\Omega$ . Furthermore, it is expected that in the limit of large volumes, the mean concentration approaches some non-zero value while the variance approaches zero (deterministic conditions). Hence we make the ansatz that the mean concentration and the variance can be written as  $\mu_{2M} = \sum_{i=0}^{\infty} a_i \Omega^{-i}$  and  $\sigma_{2M}^2 = \sum_{i=1}^{\infty} b_i \Omega^{-i}$ . Substitution of these series expansions in the 2MA equations, Eq. (13), setting the time derivative to zero and solving the equations simultaneously (for  $a_i$  and  $b_i$ ) we obtain

$$\begin{aligned} \mu_{2M}^* &= \phi^* + \frac{1}{2\Omega\sqrt{1+8\Lambda}} \left( 1 - \frac{1}{\sqrt{1+8\Lambda}} \right) \\ &+ \frac{2\Lambda(3+16\Lambda)(\sqrt{1+8\Lambda}-1-2\Lambda)}{\Omega^2\phi_T(1+8\Lambda)^{5/2}} + O(\Omega^{-3}), \end{aligned} \quad (15)$$

$$\begin{aligned} \sigma_{2M}^{2*} &= \frac{\phi_T}{2\Lambda\Omega} \left( \frac{1+4\Lambda}{\sqrt{1+8\Lambda}} - 1 \right) \\ &+ \frac{2((1+4\Lambda)^2 - (1+6\Lambda)\sqrt{1+8\Lambda})}{\Omega^2(1+8\Lambda)^2} + O(\Omega^{-3}), \end{aligned} \quad (16)$$

where the superscript \* indicates quantities evaluated in steady-state conditions,  $\Lambda = k_1\phi_T/k_2$  is a convenient non-dimensional parameter and  $\phi^* = \phi(t)|_{t \rightarrow \infty} = \phi_T(\sqrt{1+8\Lambda}-1)/4\Lambda$  is the steady-state solution for the monomer concentration obtained using the REs one conventionally writes down for the dimerization reaction, i.e.,  $\partial_t\phi(t) = -2k_1\phi(t)^2 + k_2(\phi_T - \phi(t))$ .

In a similar manner we can obtain series solutions for the 3MA equations by making the ansatz  $\mu_{3M} = \sum_{i=0}^{\infty} a'_i \Omega^{-i}$ ,  $\sigma_{3M}^2 = \sum_{i=1}^{\infty} b'_i \Omega^{-i}$ , and  $\eta_{3M} = \sum_{i=2}^{\infty} c'_i \Omega^{-i}$ , which leads to

$$\begin{aligned} \mu_{3M}^* &= \phi^* + \frac{1}{2\Omega\sqrt{1+8\Lambda}} \left( 1 - \frac{1}{\sqrt{1+8\Lambda}} \right) \\ &+ \frac{2\Lambda(2\Lambda + \sqrt{1+8\Lambda} - 1)}{\Omega^2\phi_T(1+8\Lambda)^{5/2}} + O(\Omega^{-3}), \end{aligned} \quad (17)$$

$$\begin{aligned} \sigma_{3M}^{2*} &= \frac{\phi_T}{2\Lambda\Omega} \left( \frac{1+4\Lambda}{\sqrt{1+8\Lambda}} - 1 \right) \\ &+ \frac{4\Lambda}{\Omega^2(1+8\Lambda)^2} (\sqrt{1+8\Lambda} - 2) + O(\Omega^{-3}). \end{aligned} \quad (18)$$

A comparison of the series solutions for the 2MA and the 3MA shows that the two agree to order  $\Omega^{-1}$  but their  $\Omega^{-2}$  terms are different. In the macroscopic limit of large volumes, the mean concentrations of both methods approach the RE solution, as expected. However, since both methods are based on an *ad hoc* assumption, we cannot deduce anything about the moment-closure approximations' accuracy from comparison of their series solutions. Rather we first need to derive the mean concentration and variance prediction of the CME correct to order  $\Omega^{-2}$  and then compare the latter to the 2MA and 3MA series solution. This is the topic of Section II B.

## B. System-size expansion and the accuracy of moment-closure approximations

We now use a systematic method, the SSE of the master equation, to obtain the steady-state mean concentration and variance accurate to order  $\Omega^{-2}$ . The SSE is not based on *ad hoc* moment-closure approximations and hence serves as an ideal method to benchmark the accuracy of the 2MA and 3MA results in Section II A. The reader is referred to Appendix B for a concise treatment of the multivariate SSE in the context of chemical reaction kinetics. In what follows we shall simply state the main results obtained by applying the SSE to the dimerization example at hand.

The SSE starts by making the change of variable  $n_A/\Omega = \phi(t) + \Omega^{-1/2}\epsilon_A$  in the master equation (3), where  $\phi(t)$  is the solution for the monomer concentration using the REs. This substitution causes the probability distribution  $P(n_A, t)$  to be transformed into a new one,  $\Pi(\epsilon_A, t)$ , whose time-evolution equation is given by

$$\begin{aligned} \frac{\partial \Pi}{\partial t} &= \Omega^0 \left( -J_{1(1)}^1 \partial_A (\epsilon_A \Pi) + \frac{1}{2} D_{1(2)} \partial_A^2 \Pi \right) \\ &+ \Omega^{-1/2} \left( -\frac{1}{2} J_{1(1)}^{11} \partial_A (\epsilon_A^2 \Pi) + \frac{1}{2} \phi J_{1(1)}^{11} \partial_A \Pi \right. \\ &+ \frac{1}{2} J_{1(2)}^1 \partial_A^2 (\epsilon_A \Pi) \left. \right) + \Omega^{-1} \left( \frac{1}{2} J_{1(1)}^{11} \partial_A (\epsilon_A \Pi) \right. \\ &+ \frac{1}{4} J_{1(2)}^{11} \partial_A^2 (\epsilon_A^2 \Pi) - \frac{1}{4} J_{1(2)}^{11} \phi \partial_A^2 \Pi - \frac{1}{6} J_{1(3)}^1 \partial_A^3 (\epsilon_A \Pi) \\ &+ \frac{1}{24} D_{1(4)} \partial_A^4 \Pi \left. \right) + \Omega^{-3/2} \left( -\frac{1}{4} J_{1(2)}^{11} \partial_A^2 (\epsilon_A \Pi) \right. \\ &+ \frac{1}{24} J_{1(4)}^1 \partial_A^4 (\epsilon_A \Pi) - \frac{1}{12} J_{1(3)}^{11} \partial_A^3 (\epsilon_A^2 \Pi) \\ &+ \left. \frac{1}{12} J_{1(3)}^{11} \phi \partial_A^3 \Pi \right) + O(\Omega^{-2}), \end{aligned} \quad (19)$$

where  $\partial_A$  denotes the partial derivative with respect to  $\epsilon_A$ . We have also used the following convenient abbreviations:

$$D_{1(m)} = 2^{(m-1)}(k_2(\phi_T - \phi) + 2(-1)^m k_1 \phi^2), \quad (20)$$

$$J_{1(m)}^1 = 2^{(m-1)}(4(-1)^m k_1 \phi - k_2), \quad (21)$$

$$J_{1(m)}^{11} = 2^{(m+1)}(-1)^m k_1. \quad (22)$$

Note that  $D_{1(m)} = 0$  for odd  $m$  since in steady-state conditions the deterministic rate equations imply  $k_2(\phi_T - \phi) - 2k_1\phi^2 = 0$ ; this is the reason why no odd  $m$  terms appear in Eq. (19). Note that  $D_{1(m)}$  is short-hand notation for  $D_{ij\dots r}$  as defined by Eq. (B21) with  $m$  subscript indices all equal to one;  $J_{1(m)}^1$  and  $J_{1(m)}^{11}$  are similarly related to  $J_{ij\dots r}^{i_1\dots i_m}$  defined by Eq. (B22).

The right-hand side of the time-evolution equation for  $\Pi(\epsilon_A, t)$  is an infinite series in powers of the inverse square root of the compartment volume which implies that it can be solved by making the substitution  $\Pi(\epsilon_A, t) = \sum_{j=0}^{\infty} \Pi_j(\epsilon_A, t) \Omega^{-j/2}$ . We also define the pseudo-moments as  $[(\epsilon_A)^i]_j = \int (\epsilon_A)^i \Pi_j(\epsilon_A, t) d\epsilon_A$ . We impose deterministic initial conditions, i.e., at time  $t = 0$ , the value of  $n_A/\Omega$  is taken

to be exactly equal to the deterministic solution  $\phi(0)$ . It can then be shown (see Appendix B) that the mean concentration and the variance of the concentration fluctuations are given by

$$\mu_{SSE} = \left\langle \frac{n_A}{\Omega} \right\rangle = \phi + [\epsilon_A]_1 \Omega^{-1} + [\epsilon_A]_3 \Omega^{-2} + O(\Omega^{-3}), \quad (23)$$

$$\sigma_{SSE}^2 = \left\langle \left( \frac{n_A}{\Omega} - \frac{\langle n_A \rangle}{\Omega} \right)^2 \right\rangle = \Omega^{-1} [(\epsilon_A)^2]_0 + \Omega^{-2} ([(\epsilon_A)^2]_2 - [\epsilon_A]_1^2) + O(\Omega^{-3}), \quad (24)$$

where the time-evolution equations for the pseudo-moments  $[(\epsilon_A)^j]_t$  are given by

$$\partial_t [(\epsilon_A)^2]_0 = 2J_{1(1)}^1 [(\epsilon_A)^2]_0 + D_{1(2)}, \quad (25)$$

$$\partial_t [(\epsilon_A)^4]_0 = 4J_{1(1)}^1 [(\epsilon_A)^4]_0 + 6D_{1(2)} [(\epsilon_A)^2]_0, \quad (26)$$

$$\partial_t [\epsilon_A]_1 = J_{1(1)}^1 [\epsilon_A]_1 + \frac{1}{2} J_{1(1)}^{11} ([(\epsilon_A)^2]_0 - \phi), \quad (27)$$

$$\begin{aligned} \partial_t [(\epsilon_A)^3]_1 &= 3J_{1(1)}^1 [(\epsilon_A)^3]_1 + \frac{3}{2} J_{1(1)}^{11} ([(\epsilon_A)^4]_0 - \phi [(\epsilon_A)^2]_0) \\ &\quad + 3D_{1(2)} [\epsilon_A]_1 + 3J_{1(2)}^1 [(\epsilon_A)^2]_0, \end{aligned} \quad (28)$$

$$\begin{aligned} \partial_t [(\epsilon_A)^2]_2 &= 2J_{1(1)}^1 [(\epsilon_A)^2]_2 + J_{1(1)}^{11} ([(\epsilon_A)^3]_1 - \phi [\epsilon_A]_1 - [(\epsilon_A)^2]_0) \\ &\quad + J_{1(2)}^1 [\epsilon_A]_1 + \frac{1}{2} J_{1(2)}^{11} ([(\epsilon_A)^2]_0 - \phi), \end{aligned} \quad (29)$$

$$\partial_t [\epsilon_A]_3 = J_{1(1)}^1 [\epsilon_A]_3 + \frac{1}{2} J_{1(1)}^{11} ([(\epsilon_A)^2]_2 - [\epsilon_A]_1). \quad (30)$$

The deterministic initial condition implies that all pseudo-moments are zero at  $t = 0$ .

Solving Eqs. (25)–(30) in steady-state conditions (setting the time derivative to zero) and substituting in Eqs. (23) and (24) we obtain the mean concentration and variance of the fluctuations accurate to order  $\Omega^{-2}$

$$\begin{aligned} \mu_{SSE}^* &= \phi^* + \frac{1}{2\Omega\sqrt{1+8\Lambda}} \left( 1 - \frac{1}{\sqrt{1+8\Lambda}} \right) \\ &\quad + \frac{2\Lambda(2\Lambda + \sqrt{1+8\Lambda} - 1)}{\Omega^2\phi_T(1+8\Lambda)^{5/2}} + O(\Omega^{-3}), \end{aligned} \quad (31)$$

$$\begin{aligned} \sigma_{SSE}^{2*} &= \frac{\phi_T}{2\Lambda\Omega} \left( \frac{1+4\Lambda}{\sqrt{1+8\Lambda}} - 1 \right) \\ &\quad + \frac{4\Lambda}{\Omega^2(1+8\Lambda)^2} (\sqrt{1+8\Lambda} - 2) + O(\Omega^{-3}). \end{aligned} \quad (32)$$

Note that the first term in Eq. (32) is the LNA. The first two terms in Eq. (31) are those given by the effective mesoscopic rate equation (EMRE) formalism.<sup>15</sup> A comparison of the above equations with the series solutions for the 2MA, Eqs. (15) and (16), and the 3MA, Eqs. (17) and (18), reveals that the 2MA and 3MA predictions are generally accurate to

orders  $\Omega^{-1}$  and  $\Omega^{-2}$ , respectively. The 2MA only achieves a higher accuracy of order  $\Omega^{-2}$  in both the mean concentration and the variance of fluctuations when  $\Lambda = (1 + \sqrt{2})/4 \simeq 0.6$ . Hence the 2MA accurately captures the first-order correction to the REs but does not accurately predict the first-order correction to the LNA. In particular the 2MA's prediction of the correction to the LNA has the wrong dependence on  $\Lambda$ . In contrast, the 3MA accurately captures the first- and second-order corrections to the RE predictions and the first-order correction to the LNA.

As we show now, a consequence of the 2MA's failure to predict the first-order correction to the LNA is that the 2MA's prediction of variance is only more accurate than that of the LNA in some regions of parameter space. We define the relative error in the variance prediction of a method as the solution from the SSE minus the solution from the method divided by the solution from the SSE. Using this definition, it is straightforward to show, using the series solutions of the 2MA, and that obtained using the SSE, that the relative errors in the variance predictions of the 2MA and of the LNA are given by

$$E_{var}^{LNA} = \frac{1 - 2y^{-3/2} - 3y^{-1}}{4\phi_T\Omega} + O(\Omega^{-2}), \quad (33)$$

$$E_{var}^{2MA} = -\frac{1 + 2(\sqrt{y} - y - y^{3/2}) + y^2}{4\phi_T y^{3/2}\Omega} + O(\Omega^{-2}), \quad (34)$$

where  $y = 1 + 8\Lambda$ . A comparison of these equations shows that in the limit  $\Lambda \rightarrow 0$  the absolute relative error of the LNA is larger than the absolute relative error of the 2MA while the reverse is true in the opposite limit of large  $\Lambda$ . In particular in the latter limit, the error of the LNA prediction becomes independent of  $\Lambda$  while the error of the 2MA scales as  $\sqrt{\Lambda}$ . One can also show that the relative error in the variance prediction of the 3MA scales as  $\Omega^{-2}$  and hence the 3MA is the lowest order moment-closure approximation which is more accurate than the LNA across all of parameter space.

### C. Numerical verification of the theoretical results

We have verified our theoretical predictions by calculating two quantities,  $\Delta$ , the ratio of the mean concentration of the CME to the RE solution and  $\Theta$ , the ratio of the variance of concentration fluctuations of the CME to the LNA prediction using four methods:

1. We numerically solve the 2MA Eqs. (13) to obtain the mean and variance estimates in steady-state conditions. We then divide the mean concentration solution by the steady-state solution of the deterministic RE,  $\phi^* = \phi_T(\sqrt{1+8\Lambda} - 1)/4\Lambda$ , to obtain the  $\Delta$  estimate. The variance estimate is divided by the steady-state LNA solution (the first term on the right-hand side of Eq. (32)) to obtain the  $\Theta$  estimate.
2. Same as above but using the 3MA Eqs. (14).
3. The SSE estimate of  $\Delta$  is obtained by dividing Eq. (31) by the steady-state solution of the deterministic RE while the estimate of  $\Theta$  is obtained by dividing Eq. (32)

by the steady-state solution of the LNA

$$\Delta_{SSE} = 1 + \frac{2\Lambda}{\phi_T \Omega (1 + 8\Lambda)} + O(\Omega^{-2}), \quad (35)$$

$$\Theta_{SSE} = 1 + \frac{1}{4\phi_T \Omega} \left( 1 - \frac{3}{1 + 8\Lambda} - \frac{2}{(1 + 8\Lambda)^{3/2}} \right) + O(\Omega^{-2}). \quad (36)$$

4. Exact values of  $\Delta$  and  $\Theta$  are calculated as follows. The exact solution of the master equation (3) in steady-state conditions has been previously found by the generating function method.<sup>16</sup> The steady-state generating function solution is

$$G(s) = \frac{{}_1F_1\left[-\frac{n_T}{2}, \frac{1}{2}, -\frac{n_T s^2}{4\Lambda}\right]}{{}_1F_1\left[-\frac{n_T}{2}, \frac{1}{2}, -\frac{n_T}{4\Lambda}\right]}, \quad (37)$$

where  $G(s) = \sum_{n_A} s^{n_A} P(n_A)$ ,  $n_T = \phi_T \Omega$  is the maximum number of monomer molecules, and  ${}_1F_1$  denotes the Kummer confluent hypergeometric function. The exact mean monomer concentration is then given by  $\Omega^{-1} \partial_s G(s)$  evaluated at  $s = 1$ . Dividing this concentration by the steady-state solution of the RE we obtain the exact  $\Delta$ . The exact variance of the monomer concentration fluctuations is  $\Omega^{-2} (\partial_s^2 G(s) + \partial_s G(s) - (\partial_s G(s))^2)$  with all the derivatives evaluated at  $s = 1$ . Dividing this

variance by the steady-state solution of the LNA we obtain the exact  $\Theta$ .

In Fig. 1 we fix  $\phi_T = 1$  and plot the values of  $\Delta$  and  $\Theta$  obtained from the above four methods as a function of the non-dimensional aggregate of rate constants  $\Lambda$  and for two different volumes: an intermediate volume of  $\Omega = 6$  (Figs. 1(a) and 1(c)) and a large volume of  $\Omega = 100$  (Figs. 1(b) and 1(d)). Note that given the chosen value of  $\phi_T$ , the intermediate and large volumes correspond to a maximum number of monomer molecules  $n_T$  equal to 6 and 100, respectively. The color and symbol coding is as follows: red triangle points (connected by solid red lines) show the 2MA predictions, open circle points show the 3MA predictions, blue dashed line shows the SSE predictions, Eqs. (35) and (36), and the solid black circles are computed using the exact approach.

Note that at intermediate volumes (Figs. 1(a) and 1(c)), the 2MA, 3MA, and SSE estimates are visibly different than the exact values, the smallest discrepancies being for the 3MA and SSE estimates. For large volumes (Figs. 1(b) and 1(d)), the 3MA and SSE estimates agree very well with the exact values, the 2MA estimate of  $\Delta$  also agrees well with the exact values but the 2MA estimate of  $\Theta$  has a completely different functional dependence on  $\Lambda$  than the exact one. Increasing  $\Omega$  larger than 100, we find that the 2MA  $\Theta$  always retains this incorrect dependence on  $\Lambda$ . These numerical results confirm the results of our large volume expansion method. In the limit of large volumes the 2MA accurately predicts the size

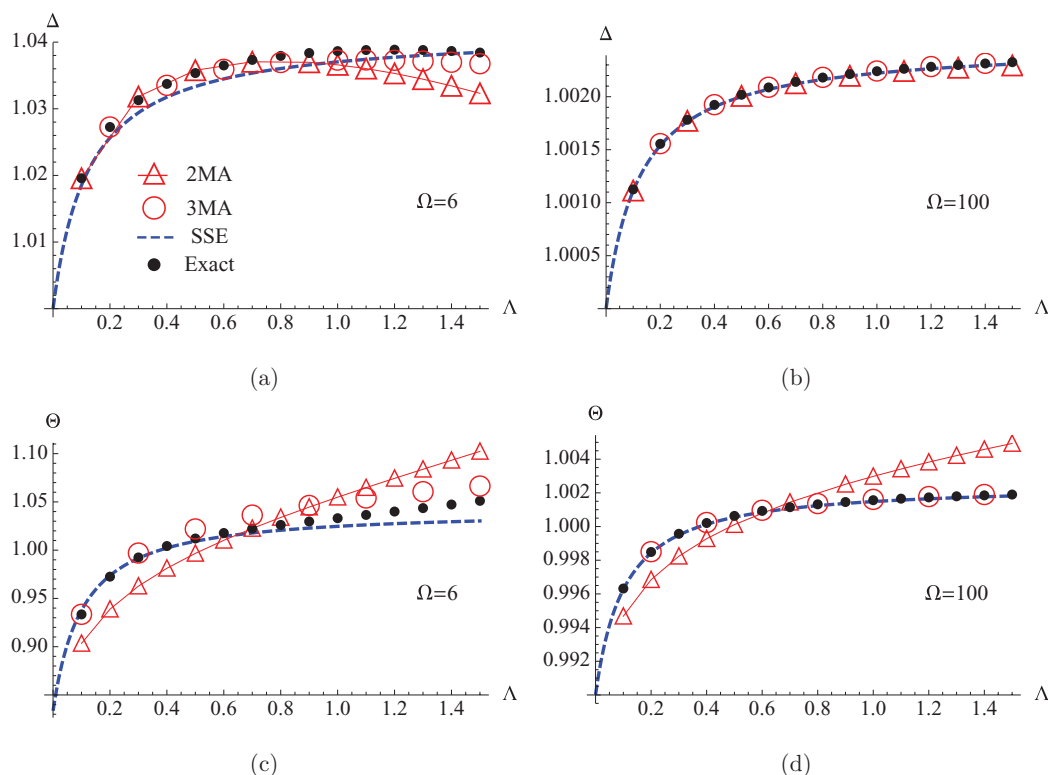


FIG. 1. Plots of  $\Delta$ , the ratio of the mean monomer concentration of the CME to the RE solution and  $\Theta$ , the ratio of the variance of monomer concentration fluctuations of the CME to the LNA, as a function of  $\Lambda$  for the reversible dimerization reaction in steady-state conditions. These are obtained using four different methods: 2MA, 3MA, SSE, and an exact approach (see text for details). The 2MA, 3MA, and SSE estimates are not very accurate for small volumes, panels (a) and (c). For large volumes, panels (b) and (d), the 3MA and SSE estimates of both  $\Delta$  and  $\Theta$  and the 2MA's estimate of  $\Delta$  match very well the exact values. However, the 2MA's estimate of  $\Theta$  is incorrect, in the sense that it predicts the wrong dependence of  $\Theta$  with  $\Lambda$ . The latter deficiency cannot be corrected by going to larger volumes than  $\Omega = 100$  and is a pathological feature of the 2MA method, namely, that the method cannot accurately capture corrections to the LNA.

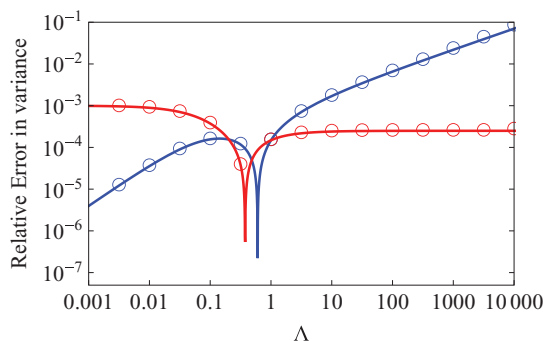


FIG. 2. Plots of the relative error in the variance predictions of the 2MA (blue) and the LNA (red) in steady-state conditions. Note that we have taken the absolute value of the relative errors. The solid lines are the theoretical predictions, Eqs. (33) and (34), and the open circles are computed using the numerical solution of the 2MA equations and the exact variance of the CME (see text for details). The theoretical and numerical relative errors for both the 2MA and the LNA match very well. The 2MA's estimate of variance is more accurate than that of the LNA for small values of  $\Lambda$  while the reverse occurs when  $\Lambda$  becomes larger than 1.

of the correction to the RE solution but it does not accurately capture corrections to the LNA solution except at the point  $\Lambda \simeq 0.6$ . The 3MA is the lowest order moment-closure approximation which accurately captures corrections to both the RE and LNA solution. Note that though the 2MA prediction for  $\Theta$  has an incorrect functional dependence on  $\Lambda$ , it is qualitatively correct in the sense that it is monotonically increasing with  $\Lambda$ .

We have also tested the accuracy of the expressions for the relative errors in the variance predictions of the LNA and the 2MA, Eqs. (33) and (34). The “true” relative error of the 2MA was calculated by substituting the numerical variance estimate of the 2MA (see point 1 above) and the exact variance (see point 4 above) in the equation: exact variance minus numerical 2MA variance divided by the exact variance. The “true” relative error of the LNA can be calculated in a similar way. In Fig. 2 we plot the absolute values of the “true” relative errors of the 2MA and the LNA and the theoretical Eqs. (33) and (34) as a function of  $\Lambda$  for the parameters  $\phi_T = 1$  and  $\Omega = 1000$ . We find excellent agreement between the “true” and theoretical values, in the process verifying the theoretical prediction that the variance estimate of the 2MA is only more accurate than that of the LNA in certain regions of parameter space.

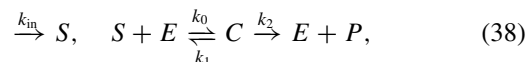
### III. EXAMPLE B: THE MICHAELIS-MENTEN REACTION WITH SUBSTRATE INPUT

In this section we repeat the same analysis as in Section II but for species involved in an enzyme-catalyzed reaction. We will show that the features of the 2MA and 3MA methods elucidated in the previous example also hold for the enzyme reaction, namely, that the 2MA accurately captures corrections to the RE solution but not those to the LNA's estimate of variance of concentration fluctuations while the 3MA accurately captures both corrections. There is however one notable and important difference between the two examples: the 2MA prediction of  $\Theta$ , the ratio of the

variance of fluctuations of the CME to the LNA, was at least qualitatively correct for the dimerization reaction but no such agreement will be found in the enzyme example.

#### A. The 2MA and 3MA equations

The Michaelis-Menten reaction with substrate input can be described in schematic form as



where the  $k$ 's denote the associated rate constants. The reaction can be described as follows. Substrate molecules (species  $S$ ) are pumped into some compartment at a constant rate, they bind to free enzyme molecules (species  $E$ ) to form substrate-enzyme complexes (species  $C$ ) which then either decay back to the original substrate and free enzyme molecules or else decay into free enzyme and product molecules (species  $P$ ). The master equation for these four elementary chemical processes is given by

$$\begin{aligned} \partial_t P(n_S, n_C, t) = & \left[ \frac{k_0}{\Omega} (E_S^{+1} E_C^{-1} - 1) n_S (n_T - n_C) \right. \\ & + \Omega k_{in} (E_S^{-1} - 1) + k_1 (E_S^{-1} E_C^{+1} - 1) n_C \\ & \left. + k_2 (E_C^{+1} - 1) n_C \right] P(n_S, n_C, t), \quad (39) \end{aligned}$$

where  $n_S$  and  $n_C$  are the absolute number of substrate and complex molecules, respectively, and  $n_T$  is the absolute total number of molecules of enzyme in free and complex form. Note that we have used the same concise step operator formalism utilized in the previous example. Note also that the number of free enzyme molecules  $n_E$  does not explicitly enter the CME description because of the conservation law  $n_E + n_C = n_T = \phi_T \Omega$ , where  $\phi_T$  is the total enzyme concentration, a time-independent quantity.

As in the previous example, we can obtain exact time-evolution equations for the first, second, and third moments by multiplying the master equation (39) by the appropriate quantity (e.g.,  $n_S/\Omega$  if we are interested in the equation for the mean substrate concentration,  $n_S n_C/\Omega^2$  if we are interested in the equation for  $\langle n_S n_C \rangle/\Omega^2$ , etc.) and summing over  $n_C$  and  $n_S$ . This leads us to the set of moment equations given by

$$\partial_t \mu_S = \sum_{j=1}^4 S_{1j} a_j, \quad (40)$$

$$\partial_t \mu_C = \sum_{j=1}^4 S_{2j} a_j, \quad (41)$$

$$\partial_t \mu_{CS} = \sum_{j=1}^4 S_{1j} c_j + S_{2j} b_j + \frac{S_{1j} S_{2j}}{\Omega} a_j, \quad (42)$$

$$\partial_t \mu_{SS} = 2 \sum_{j=1}^4 S_{1j} b_j + \frac{S_{1j}^2}{\Omega} a_j, \quad (43)$$



$$\partial_t \mu_{CC} = 2 \sum_{j=1}^4 S_{2j} c_j + \frac{S_{2j}^2}{\Omega} a_j, \quad (44)$$

$$\partial_t \mu_{SS} = \sum_{j=1}^4 3S_{1j} d_j + 3 \frac{S_{1j}^2}{\Omega} b_j + \frac{S_{1j}^3}{\Omega^2} a_j, \quad (45)$$

$$\partial_t \mu_{CC} = \sum_{j=1}^4 3S_{2j} e_j + 3 \frac{S_{2j}^2}{\Omega} c_j + \frac{S_{2j}^3}{\Omega^2} a_j, \quad (46)$$

$$\begin{aligned} \partial_t \mu_{CSS} &= \sum_{j=1}^4 S_{2j} d_j + 2S_{1j} g_j + 2 \frac{S_{1j} S_{2j}}{\Omega} b_j + \frac{S_{1j}^2}{\Omega} c_j \\ &+ \frac{S_{1j}^2 S_{2j}}{\Omega^2} a_j, \end{aligned} \quad (47)$$

$$\begin{aligned} \partial_t \mu_{CCS} &= \sum_{j=1}^4 2S_{2j} g_j + S_{1j} e_j + 2 \frac{S_{1j} S_{2j}}{\Omega} c_j + \frac{S_{2j}^2}{\Omega} b_j \\ &+ \frac{S_{1j} S_{2j}^2}{\Omega^2} a_j, \end{aligned} \quad (48)$$

where  $\mu_S = \Omega^{-1} \langle n_S \rangle$ ,  $\mu_C = \Omega^{-1} \langle n_C \rangle$ ,  $\mu_{CS} = \Omega^{-2} \langle n_S n_C \rangle$ ,  $\mu_{SS} = \Omega^{-2} \langle n_S^2 \rangle$ ,  $\mu_{CC} = \Omega^{-2} \langle n_C^2 \rangle$ ,  $\mu_{SSS} = \Omega^{-3} \langle n_S^3 \rangle$ ,  $\mu_{CCC} = \Omega^{-3} \langle n_C^3 \rangle$ ,  $\mu_{CSS} = \Omega^{-3} \langle n_S^2 n_C \rangle$ , and  $\mu_{CCS} = \Omega^{-3} \langle n_S n_C^2 \rangle$ . The notation  $S_{ij}$  refers to the  $i$ - $j$  element of the stoichiometric matrix  $S$  defined as

$$S = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & -1 \end{pmatrix}, \quad (49)$$

while  $a_j, b_j, c_j, d_j, e_j, g_j$  refer to the  $j$ th components of the column vectors

$$\vec{a} = (k_{in}, k_0(\phi_T \mu_S - \mu_{CS}), k_1 \mu_C, k_2 \mu_C)^T, \quad (50)$$

$$\vec{b} = (k_{in} \mu_S, k_0(\phi_T \mu_{SS} - \mu_{CSS}), k_1 \mu_{CS}, k_2 \mu_{CS})^T, \quad (51)$$

$$\vec{c} = (k_{in} \mu_C, k_0(\phi_T \mu_{CS} - \mu_{CCS}), k_1 \mu_{CC}, k_2 \mu_{CC})^T, \quad (52)$$

$$\vec{d} = (k_{in} \mu_{SS}, k_0(\phi_T \mu_{SSS} - \mu_{CCSS}), k_1 \mu_{CSS}, k_2 \mu_{CSS})^T, \quad (53)$$

$$\vec{e} = (k_{in} \mu_{CC}, k_0(\phi_T \mu_{CCS} - \mu_{CCCS}), k_1 \mu_{CCC}, k_2 \mu_{CCC})^T, \quad (54)$$

$$\vec{g} = (k_{in} \mu_{CS}, k_0(\phi_T \mu_{CSS} - \mu_{CCSS}), k_1 \mu_{CCS}, k_2 \mu_{CCS})^T, \quad (55)$$

where  $\mu_{CCCS} = \Omega^{-4} \langle n_C^3 n_S \rangle$ ,  $\mu_{CCSS} = \Omega^{-4} \langle n_S^3 n_C \rangle$ , and  $\mu_{CCSS} = \Omega^{-4} \langle n_S^2 n_C^2 \rangle$ . The set of exact equations, Eqs. (40)–(48), are not closed and hence the need of a moment-closure approximation method.

The 2MA involves solving Eqs. (40)–(44) with the third central moments (the third-order cumulants) set equal to those of a Gaussian, i.e.,  $\langle (n_i - \langle n_i \rangle)(n_j - \langle n_j \rangle)(n_k - \langle n_k \rangle) \rangle = 0$ , where  $i, j$ , and  $k$  are indices which can be either  $C$  or  $S$ . This is equivalent to setting the third (non-central) moments in Eqs. (42)–(44) equal to

$$\mu_{CSS} = \mu_C \mu_{SS} + 2(\mu_S \mu_{CS} - \mu_C \mu_S^2), \quad (56)$$

$$\mu_{CCS} = \mu_S \mu_{CC} + 2(\mu_C \mu_{CS} - \mu_S \mu_C^2). \quad (57)$$

The 3MA involves solving Eqs. (40)–(48) with the fourth central moments set equal to those of a Gaussian, i.e.,  $\langle (n_i - \langle n_i \rangle)(n_j - \langle n_j \rangle)(n_k - \langle n_k \rangle)(n_r - \langle n_r \rangle) \rangle = (\langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle)(\langle n_k n_r \rangle - \langle n_k \rangle \langle n_r \rangle) + (\langle n_i n_k \rangle - \langle n_i \rangle \langle n_k \rangle)(\langle n_j n_r \rangle - \langle n_j \rangle \langle n_r \rangle) + (\langle n_i n_r \rangle - \langle n_i \rangle \langle n_r \rangle)(\langle n_j n_k \rangle - \langle n_j \rangle \langle n_k \rangle)$ , where  $i, j, k$  and  $r$  are indices which can be either  $C$  or  $S$ . This condition forces the fourth-order cumulants to zero. It is also equivalent to setting the fourth moments in Eqs. (45)–(48) equal to

$$\begin{aligned} \mu_{CCSS} &= 6\mu_S^3 \mu_C - 6\mu_S^2 \mu_{CS} + 3\mu_{CS} \mu_{SS} \\ &+ 3\mu_S(-2\mu_C \mu_{SS} + \mu_{CSS}) + \mu_C \mu_{SSS}, \end{aligned} \quad (58)$$

$$\begin{aligned} \mu_{CCCS} &= 3\mu_{CS}(-2\mu_C^2 + \mu_{CC}) + \mu_S(\mu_{CCC} \\ &+ 6\mu_C^3 - 6\mu_C \mu_{CC}) + 3\mu_C \mu_{CCS}, \end{aligned} \quad (59)$$

$$\begin{aligned} \mu_{CCSS} &= 2(\mu_{CS}^2 - \mu_C^2 \mu_{SS} + \mu_C \mu_{CSS} + \mu_S(-4\mu_C \mu_{CS} + \mu_{CCS})) \\ &+ \mu_S^2(6\mu_C^2 - 2\mu_{CC}) + \mu_{SS} \mu_{CC}. \end{aligned} \quad (60)$$

## B. Series solutions of the 2MA and 3MA equations in steady-state conditions

All the moments in the 2MA and 3MA equations can be written as a series in powers of the inverse volume. The coupled differential equations can then be solved in steady-state leading to series solutions for the mean concentration of substrate and for the variance of its fluctuations

$$\mu_{S,2MA}^* = \phi_S^* + \frac{(1-\delta)^2 \theta}{\Omega \delta (\delta^2 + \theta)} + \frac{(1-\delta)^3 \theta (\delta + \theta)}{\phi_T \Omega^2 (\delta^2 + \theta)^3} + O(\Omega^{-3}), \quad (61)$$

$$\begin{aligned} \sigma_{S,2MA}^{2*} &= \frac{\phi_T (1-\delta) \theta l_0}{\delta^2 (\delta^2 + \theta) \Omega} - \frac{(1-\delta)^3 \theta (\delta^3 \eta + \delta^2 (\delta + \eta - 2) \theta - \theta^2)}{\delta^2 (\delta^2 + \theta)^3 \Omega^2} \\ &+ O(\Omega^{-3}), \end{aligned} \quad (62)$$

$$\begin{aligned} \mu_{S,3MA}^* &= \phi_S^* + \frac{(1-\delta)^2 \theta}{\Omega \delta (\delta^2 + \theta)} \\ &- \frac{(1-\delta) \delta \theta \sum_{i=0}^3 h_i \theta^i}{\phi_T \Omega^2 (\delta^2 + \theta)^3 (\delta^3 (2\delta + \eta) + 4\delta^2 \theta + 2\theta^2)} \\ &+ O(\Omega^{-3}), \end{aligned} \quad (63)$$

$$\begin{aligned} \sigma_{S,3MA}^{2*} &= \frac{\phi_T(1-\delta)\theta l_0}{\delta^2(\delta^2+\theta)\Omega} \\ &+ \frac{(1-\delta)^2 \sum_{i=1}^5 l_i \theta^i}{\delta^2(\delta^2+\theta)^3(\delta^3(2\delta+\eta)+4\delta^2\theta+2\theta^2)\Omega^2} \\ &+ O(\Omega^{-3}), \end{aligned} \quad (64)$$

where  $\phi_S^* = \phi_T(1-\delta)\theta/\delta$  is the steady-state solution of the REs for the enzyme system, and  $\delta = 1 - (k_{in}/k_2\phi_T)$ ,  $\theta = K_M/\phi_T$ ,  $\eta = k_2/k_0\phi_T$  are non-dimensional constants. Steady-state conditions require the substrate input rate  $k_{in}$  to be less than the maximum rate at which the enzyme can catalyze substrate into product  $k_2\phi_T$ , and hence  $\delta$  is a positive fraction; its value approaches zero as the enzyme becomes increasingly saturated with substrate. The constant  $\eta$  is always smaller than  $\theta$ . The functions  $h_i$  and  $l_i$  are defined as

$$h_0 = \delta^3(2\delta - 3)(2\delta + \eta), \quad (65)$$

$$h_1 = \delta(\delta(\delta(2\delta + \eta - 4) - 4) - 2\eta), \quad (66)$$

$$h_2 = 2(1 - 4\delta), \quad h_3 = -2, \quad (67)$$

$$l_0 = \delta^3 + (\delta - 1)\delta\eta + \theta, \quad (68)$$

$$l_1 = \delta^6(2\delta^3 + 11\delta^2\eta + 2\delta(\eta - 6)\eta - 3\eta^2), \quad (69)$$

$$l_2 = \delta^4(2\delta^2(8+(\delta-4)\delta)+\delta(\delta(2+5\delta)-10)\eta+(\delta^2-2)\eta^2), \quad (70)$$

$$l_3 = \delta^3(\delta(26+2\delta(\eta-4)-7\eta)+\eta), \quad (71)$$

$$l_4 = 2\delta^2(6-\delta(\delta+\eta-2)), \quad l_5 = 2(1+\delta). \quad (72)$$

The SSE of the CME can be used to calculate the mean substrate concentration and variance of fluctuations correct to order  $\Omega^{-2}$ . This is obtained as follows. According to the recipe in Appendix B, the only two pieces of information needed to compute the expansion are the macroscopic rate vector  $\vec{f}$  and the stoichiometric matrix  $S$ . Both of these are defined in Appendix A. If we label, the substrate as species 1, the complex as species 2, the input reaction as reaction 1, the enzyme-substrate association as reaction 2, its reverse as reaction 3, and the product forming step as reaction 4, then it follows that the stoichiometric matrix is given by Eq. (49) while the macroscopic rate function vector is given by  $\vec{f} = (k_{in}, k_0\phi_S(\phi_T - \phi_C), k_1\phi_C, k_2\phi_C)$ , where  $\phi_C$  and  $\phi_S$  are the RE solutions for the complex and substrate concentrations, respectively. Substituting these in Eqs. (B12), (B14), (B15), (B17), (B19) and (B20), solving these equations simultaneously and finally substituting the solutions in Eqs. (B23) and (B24), we obtain the mean substrate concentration and variance of fluctuations correct to order  $\Omega^{-2}$ . These are found to be precisely the same as those of the 3MA, namely, Eqs. (63) and (64). Hence as for the dimerization example, we can conclude that the 2MA's estimates are generally accurate to order  $\Omega^{-1}$ , except when  $\delta \rightarrow 0$  (enzyme saturation) in which case the variance exhibits a higher accuracy of  $\Omega^{-2}$ . This implies that while the 2MA captures small

corrections to the RE prediction of the mean concentration, it does not accurately predict corrections to the LNA prediction of the variance of concentration fluctuations.

These predictions are verified by a similar numerical procedure as to the one for dimerization. We numerically solve the 2MA equations (Eqs. (40)–(44) together with Eqs. (56) and (57)) and the 3MA equations (Eqs. (40)–(48) together with Eqs. (58)–(60)) for the mean substrate concentration and for the variance of the substrate concentration fluctuations in steady-state conditions. We then divide the means by the deterministic steady-state solution  $\phi_S^*$  to obtain the 2MA and 3MA estimate of  $\Delta$ , the ratio of the mean substrate concentration of the CME to the RE solution, while the variances are divided by the LNA prediction (the first term in Eq. (62)) to obtain the 2MA and 3MA estimates of  $\Theta$ , the ratio of the variance of substrate fluctuations of the CME to the LNA. We also obtain estimates of  $\Delta$  and  $\Theta$  using the SSE. Since the mean and variance results from this method are equal to order  $\Omega^{-2}$  with those of the 3MA, Eqs. (63) and (64), it then follows that the SSE estimates are given by

$$\Delta_{SSE} = 1 + \frac{1-\delta}{\phi_T\Omega(\delta^2+\theta)} + O(\Omega^{-2}), \quad (73)$$

$$\begin{aligned} \Theta_{SSE} &= 1 + \frac{(1-\delta) \sum_{i=1}^5 g_i \theta^{i-1}}{\phi_T g_0(\delta^2+\theta)^2(\delta^3(2\delta+\eta)+4\delta^2\theta+2\theta^2)\Omega} \\ &+ O(\Omega^{-2}). \end{aligned} \quad (74)$$

Finally we obtain  $\Delta$  and  $\Theta$  using an exact method. Stochastic simulation of the reaction scheme Eq. (38) using Gillespie's exact algorithm<sup>17</sup> is used to generate a set of 3 independent trajectories of the number of substrate molecules versus time. Estimates for the mean number of substrate molecules and for the mean squared number of substrate molecules in steady-state conditions are obtained from time averages of each trajectory over a period of time which is much longer than that necessary to achieve steady-state. The three values of the mean are then averaged to lead to a final mean estimate; the same procedure is used to obtain a final mean squared estimate. The mean is divided by the volume leading to the mean concentration which is then divided by the deterministic RE solution to obtain  $\Delta$ . Subtracting the square of the mean estimate from the mean square estimate and dividing by the volume squared leads us to the variance of fluctuations which upon division by the LNA prediction gives us  $\Theta$ . Note that  $\Delta$  and  $\Theta$  obtained from this method are exact, the only error being that due to sampling, which can be made negligible by time averaging over very long trajectories. The standard deviation in  $\Delta$  and  $\Theta$  can also be obtained from the standard deviation of the three values of the mean and of the mean squared estimates.

In Fig. 3 we show a comparison of the  $\Delta$  and  $\Theta$  estimates from the 2MA, 3MA, SSE, and exact methods for two different volumes: an intermediate volume,  $\Omega = 250$  (Figs. 3(a) and 3(c)) and a large volume,  $\Omega = 1000$  (Figs. 3(b) and 3(d)). The total enzyme concentration is fixed to  $\phi_T = 1$  and hence the total number of enzyme molecules is equal to  $\Omega$ . Note that the 2MA prediction of  $\Delta$  is accurate for both volumes. However, its  $\Theta$  prediction is qualitatively

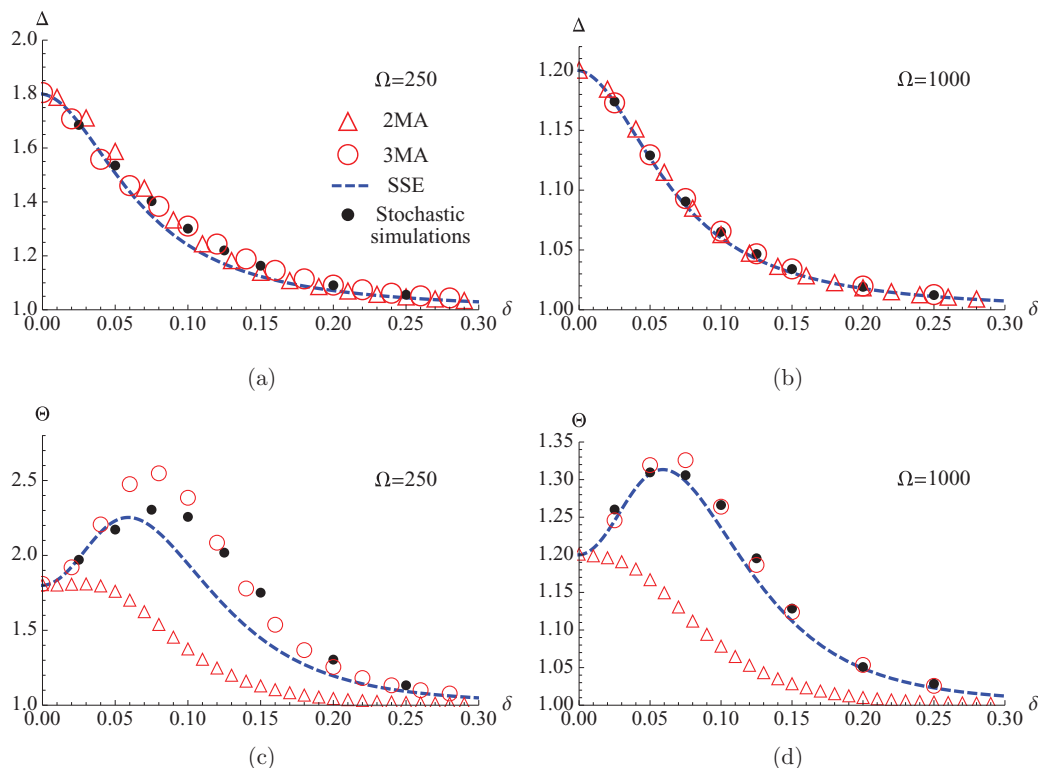


FIG. 3. Plots of  $\Delta$ , the ratio of the mean substrate concentration of the CME to the RE solution and  $\Theta$ , the ratio of the variance of substrate concentration fluctuations of the CME to the LNA, as a function of  $\delta$ , for the Michaelis-Menten reaction with substrate input in steady-state conditions. The plots are obtained using four different methods: 2MA, 3MA, SSE, and simulations performed using the stochastic simulation algorithm (see text for details). The standard error of the data points obtained from the latter method is less than 0.01 for the  $\Delta$  estimates in (a) and (b) and less than 0.05 and 0.02 for the  $\Theta$  estimates in (c) and (d), respectively. The total enzyme concentration is fixed to  $\phi_T = 1$  and the rate constants are  $k_0 = 200$  and  $k_1 = k_2 = 0.5$ . The 2MA accurately predicts the corrections to the REs ( $\Delta$ ) for both small volumes, panel (a), and large volumes, panel (b). However, its prediction for the corrections to the LNA ( $\Theta$ ) has an incorrect dependence on  $\delta$  for both volumes (see panels (c) and (d)). In contrast, as the volume is increased, the 3MA's estimates match those of the SSE and of the stochastic simulation algorithm (see panel (d)).

wrong since it does not capture the peak at a particular value of  $\delta$  predicted from the exact method. This qualitative disagreement cannot be resolved by going to larger volumes (as can be seen from a comparison of Figs. 3(c) and 3(d)). This is in agreement with our theoretical results which show that for any volume, the 2MA's  $\Omega^{-2}$  term has the wrong functional dependence on  $\delta$ . In contrast, the 3MA and the SSE estimates of both  $\Delta$  and  $\Theta$  approach the exact values as the volume becomes larger (compare Figs. 3(c) and 3(d)).

### C. Accuracy of the 2MA's and the 3MA's predictions for pre-steady-state kinetics

Thus far, we have exclusively studied the accuracy of the 2MA and the 3MA in steady-state conditions. Needless to say that one is generally interested in the kinetics leading to the steady state. We here numerically test whether the conclusions we obtained for steady-state conditions also hold in pre-steady-state conditions.

We have computed  $\Delta$  and  $\Theta$  as a function of time, using the 2MA, the 3MA, and the stochastic simulation algorithm. The details are as follows. We fix  $\delta$  to 0.07,  $\Omega$  to 1000, and the rest of the parameters are as in Fig. 3. The estimates are computed in the following way. The REs, the LNA equations (see Appendix B), the 2MA equations, and the 3MA equations are

numerically integrated to obtain estimates of the mean substrate concentration and of the variance of substrate concentration fluctuations as a function of time. The estimate of  $\Delta(t)$  from the 2MA and the 3MA methods is then obtained by dividing their mean concentrations at time  $t$  by the mean concentration of the REs at time  $t$ . Similarly  $\Theta(t)$  is obtained by dividing the variance predictions of the moment-closure approximations at time  $t$  by the variance prediction of the LNA at time  $t$ . We have also estimated  $\Delta(t)$  and  $\Theta(t)$  by means of the stochastic simulation algorithm. Thirty thousand trajectories were generated by the algorithm, from which we calculated the mean substrate concentration and the variance of the fluctuations at unit time intervals. Dividing the former by the mean concentration of the RE and the latter by the LNA prediction of variance, leads us to the desired values of  $\Delta(t)$  and  $\Theta(t)$ .

The initial conditions are deterministic, i.e., the probability distribution solution of the CME at time  $t = 0$  is a Dirac delta function centered on the concentration solution of the REs. Hence the initial concentrations of the 2MA and the 3MA equations are set equal to those of the REs, the second non-central moments are set equal to  $\mu_{ij} = \mu_i \mu_j$  and the third non-central moments are set equal to  $\mu_{ijk} = \mu_i \mu_j \mu_k$ , where  $i$ ,  $j$ , and  $k$  are indices equal to  $C$  or  $S$ . In particular, we chose  $\mu_S = 1$  and  $\mu_C = 0$  at time  $t = 0$ ; for the stochastic simulation

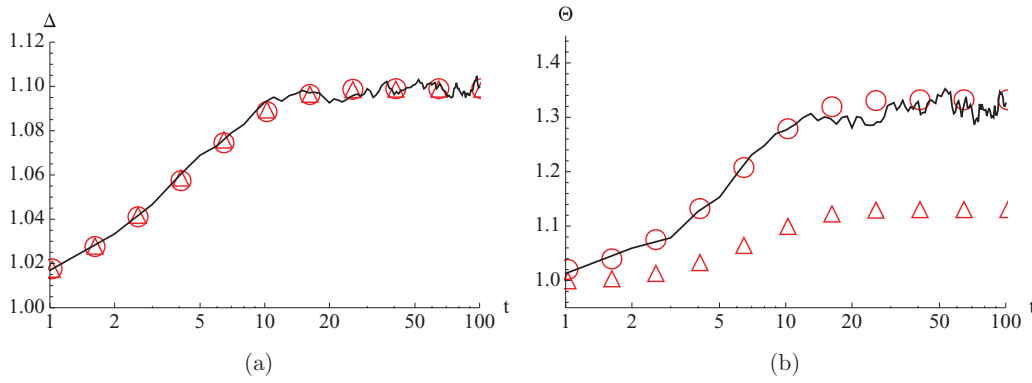


FIG. 4. Plots of  $\Delta$ , the ratio of the mean substrate concentration of the CME to the RE solution and  $\Theta$ , the ratio of the variance of substrate concentration fluctuations of the CME to the LNA, as a function of time  $t$ , for the Michaelis-Menten reaction with substrate input. The parameters are fixed to  $\phi_T = 1$ ,  $k_0 = 200$ ,  $k_1 = k_2 = 0.5$ ,  $\Omega = 1000$ , and  $\delta = 0.07$ . We impose deterministic initial conditions. The plots are obtained using four different methods: red triangle points show the 2MA estimates, the red open circle points show the 3MA estimates, and the solid black lines are computed from the output of the stochastic simulation algorithm sampled at unit time intervals (see text for details). The plots confirm that the 3MA is the lowest order moment-closure approximation which can accurately capture corrections to the REs and to the LNA for all times.

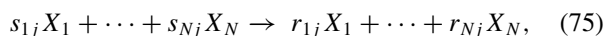
algorithm this means setting  $n_S = n_E = \Omega$  and  $n_C = 0$  at time  $t = 0$ . For the LNA, deterministic initial conditions imply setting all pseudo-moments to zero, initially (see Appendix B).

The results are shown in Fig. 4. It is evident from the plots that our previous conclusions derived from a steady-state analysis, hold for all times, i.e., the 3MA is the lowest order moment-closure approximation which can accurately capture corrections to the REs and to the LNA.

#### IV. A GENERAL PROOF OF THE ACCURACY OF THE 2MA AND THE 3MA

In Section II and III we analytically showed that for dimerization and enzyme-catalyzed reactions, the 2MA's and the 3MA's estimates of mean concentrations and of variance of fluctuations in steady-state conditions were, respectively, accurate to orders  $\Omega^{-1}$  and  $\Omega^{-2}$ . For the enzyme example, it was also numerically verified for one parameter set, that the accuracy of the two methods for all times is the same as that in steady-state conditions. In this section we sketch a proof which shows that these results are true for all monostable reaction networks and for all times. The notation and approach to the problem in this section bear considerable similarity to that used for the multivariate SSE and hence we encourage the reader to read through at least Appendix B before continuing with this section.

We consider a general molecular reaction network in a well-stirred volume  $\Omega$  involving the interaction of  $N$  distinct chemical species via  $R$  elementary reactions, i.e., at most bi-molecular reactions,



where  $j = 1, \dots, R$ ,  $X_i$  denotes species  $i$  ( $i = 1, \dots, N$ ), and  $s_{ij}$  and  $r_{ij}$  are the stoichiometric coefficients. In Appendix C, we show that the 2MA equations for this chemical system are given by

$$\partial_t \frac{\langle n_i \rangle}{\Omega} = \sum_{j=1}^R S_{ij} \langle \hat{f}_j(\vec{n}, \Omega) \rangle, \quad (76)$$

$$\begin{aligned} \partial_t \frac{\langle n_i n_k \rangle}{\Omega^2} &= \Omega^{-1} \sum_{j=1}^R (S_{kj} \langle n_i \hat{f}_j(\vec{n}, \Omega) \rangle + S_{ij} \langle n_k \hat{f}_j(\vec{n}, \Omega) \rangle \\ &\quad + S_{ij} S_{kj} \langle \hat{f}_j(\vec{n}, \Omega) \rangle), \end{aligned} \quad (77)$$

together with the condition

$$\begin{aligned} \langle n_i n_k n_m \rangle &= \langle n_i \rangle \langle n_k n_m \rangle + \langle n_k \rangle \langle n_i n_m \rangle + \langle n_m \rangle \langle n_i n_k \rangle \\ &\quad - 2 \langle n_i \rangle \langle n_k \rangle \langle n_m \rangle. \end{aligned} \quad (78)$$

The matrix  $S$  is the stoichiometric matrix of the chemical reaction system. When the  $j$ th chemical reaction occurs, the number of molecules of species  $i$  changes from  $n_i$  to  $n_i + S_{ij}$  where  $S_{ij} = r_{ij} - s_{ij}$  is the  $i$ - $j$  element of the stoichiometric matrix. The function  $\hat{f}_j(\vec{n}, \Omega)$  is the microscopic rate function of the  $j$ th reaction which is equal to the propensity function of the reaction divided by  $\Omega$ , the compartment volume (see Appendix A and Ref. 17 for a discussion of propensity functions). For elementary reactions, this function is a quadratic in the components of  $\vec{n}$ .

We now make the same ansatz as used for the derivation of the SSE

$$\frac{n_i}{\Omega} = \phi_i + \Omega^{-1/2} \epsilon_i, \quad (79)$$

where  $\phi_i$  is the macroscopic concentration of species  $i$  as determined by the REs. It has been previously shown<sup>15,18</sup> that this implies that the microscopic rate function can be written as

$$\begin{aligned} \hat{f}_j(\vec{n}, \Omega) &= \sum_{k=0}^2 \Omega^{-k/2} \frac{1}{k!} \left( \sum_{w=1}^N \epsilon_w \frac{\partial}{\partial \phi_w} \right)^k f_j(\vec{\phi}) \\ &\quad - \frac{1}{2} \Omega^{-1} \sum_{w=1}^N \phi_w \frac{\partial^2 f_j(\vec{\phi})}{\partial \phi_w^2} - \frac{1}{2} \Omega^{-3/2} \sum_{w=1}^N \epsilon_w \frac{\partial^2 f_j(\vec{\phi})}{\partial \phi_w^2}, \end{aligned} \quad (80)$$

where  $f_j(\vec{\phi})$  is the macroscopic rate function  $f_j(\vec{\phi}) = k_j \prod_{w=1}^N \phi_w^{s_{wj}}$  which encodes the law of the mass action for reaction  $j$ .

Substituting Eqs. (79) and (80) in Eqs. (76)–(78), after some algebraic manipulation one obtains

$$\begin{aligned} \partial_t \langle \epsilon_i \rangle &= J_i^w \langle \epsilon_w \rangle + \frac{\Omega^{-1/2}}{2} (J_i^{wz} \langle \epsilon_w \epsilon_z \rangle - J_i^{w(2)} \phi_w) \\ &\quad - \frac{\Omega^{-1}}{2} J_i^{w(2)} \langle \epsilon_w \rangle, \end{aligned} \quad (81)$$

$$\begin{aligned} \partial_t \langle \epsilon_i \epsilon_k \rangle &= J_k^w \langle \epsilon_i \epsilon_w \rangle + \frac{\Omega^{-1/2}}{2} (J_k^{wz} \langle \epsilon_i \epsilon_w \epsilon_z \rangle - \langle \epsilon_i \rangle J_k^{w(2)} \phi_w) \\ &\quad - \frac{\Omega^{-1}}{2} J_k^{w(2)} \langle \epsilon_i \epsilon_w \rangle + (i \leftrightarrow k) \\ &\quad + D_{ik} + \Omega^{-1/2} J_{ik}^w \langle \epsilon_w \rangle + \frac{\Omega^{-1}}{2} (J_{ik}^{wz} \langle \epsilon_w \epsilon_z \rangle - J_{ik}^{w(2)} \phi_w) \\ &\quad - \frac{\Omega^{-3/2}}{2} J_{ik}^{w(2)} \langle \epsilon_w \rangle, \end{aligned} \quad (82)$$

$$\begin{aligned} \langle \epsilon_i \epsilon_k \epsilon_m \rangle &= \langle \epsilon_i \rangle \langle \epsilon_k \epsilon_m \rangle + \langle \epsilon_k \rangle \langle \epsilon_i \epsilon_m \rangle + \langle \epsilon_m \rangle \langle \epsilon_i \epsilon_k \rangle \\ &\quad - 2 \langle \epsilon_i \rangle \langle \epsilon_k \rangle \langle \epsilon_m \rangle. \end{aligned} \quad (83)$$

Note that in the above three equations we have used the Einstein summation convention, the short-hand notation ( $i \leftrightarrow k$ ) and the symbols  $D_{ij\dots r}$  and  $J_{ij\dots r}^{st\dots z}$  which are all defined in Appendix B, just after Eq. (B20). We shall use this notation in the rest of this section since it allows expressions to be written compactly.

Now we want to obtain expressions for the mean concentrations and covariance of the concentration fluctuations of the 2MA to order  $\Omega^{-2}$ . To this end, we write the moments of the  $\epsilon$  variables as a series expansion in powers of the inverse square root of the volume

$$\langle \epsilon_k \epsilon_m \dots \epsilon_r \rangle = \sum_{j=0}^{\infty} [\epsilon_k \epsilon_m \dots \epsilon_r]_j \Omega^{-j/2}. \quad (84)$$

We refer to the square bracket quantities in the above equation as pseudo-moments (see Appendix B and Ref. 18 for a discussion of these quantities). Given this form and the ansatz Eq. (79), one can deduce that the mean concentration and the covariance of fluctuations are given by

$$\mu_i = \left\langle \frac{n_i}{\Omega} \right\rangle = \phi_i + \Omega^{-1/2} \langle \epsilon_i \rangle = \phi_i + \Omega^{-1/2} \sum_{j=0}^{\infty} [\epsilon_i]_j \Omega^{-j/2}, \quad (85)$$

$$\begin{aligned} \sigma_{ik}^2 &= \left\langle \left( \frac{n_i}{\Omega} - \left\langle \frac{n_i}{\Omega} \right\rangle \right) \left( \frac{n_k}{\Omega} - \left\langle \frac{n_k}{\Omega} \right\rangle \right) \right\rangle = \Omega^{-1} (\langle \epsilon_i \epsilon_k \rangle - \langle \epsilon_i \rangle \langle \epsilon_k \rangle) \\ &= \Omega^{-1} \left( \sum_{j=0}^{\infty} [\epsilon_i \epsilon_k]_j \Omega^{-j/2} - \sum_{j,m=0}^{\infty} [\epsilon_i]_j [\epsilon_k]_m \Omega^{-(j+m)/2} \right). \end{aligned} \quad (86)$$

Hence to calculate series solutions for the mean concentrations and the variances predicted by the 2MA, we first need to obtain time-evolution equations for the pseudo-moments. These are obtained by substituting Eq. (84) in Eqs. (81)–(83), and then equating powers of the inverse square root of the

volume, which leads to the following set of equations:

$$\partial_t [\epsilon_i]_0 = J_i^w [\epsilon_w]_0, \quad (87)$$

$$\partial_t [\epsilon_i \epsilon_k]_0 = J_k^w [\epsilon_i \epsilon_w]_0 + (i \leftrightarrow k) + D_{ik}, \quad (88)$$

$$\begin{aligned} [\epsilon_i \epsilon_k \epsilon_m]_0 &= [\epsilon_i]_0 [\epsilon_k \epsilon_m]_0 + [\epsilon_k]_0 [\epsilon_i \epsilon_m]_0 + [\epsilon_m]_0 [\epsilon_i \epsilon_k]_0 \\ &\quad - 2 [\epsilon_i]_0 [\epsilon_k]_0 [\epsilon_m]_0, \end{aligned} \quad (89)$$

$$\partial_t [\epsilon_i]_1 = J_i^w [\epsilon_w]_1 + \frac{1}{2} J_i^{wz} [\epsilon_w \epsilon_z]_0 - \frac{1}{2} J_i^{w(2)} \phi_w, \quad (90)$$

$$\begin{aligned} \partial_t [\epsilon_i \epsilon_k]_1 &= J_k^w [\epsilon_i \epsilon_w]_1 + \frac{1}{2} J_k^{wz} [\epsilon_i \epsilon_w \epsilon_z]_0 - \frac{1}{2} J_k^{w(2)} \phi_w [\epsilon_i]_0 \\ &\quad + (i \leftrightarrow k) + J_{ik}^w [\epsilon_w]_0, \end{aligned} \quad (91)$$

$$\begin{aligned} [\epsilon_i \epsilon_k \epsilon_m]_1 &= [\epsilon_i]_1 [\epsilon_k \epsilon_m]_0 + [\epsilon_k]_1 [\epsilon_i \epsilon_m]_0 + [\epsilon_m]_1 [\epsilon_i \epsilon_k]_0 \\ &\quad + [\epsilon_i]_0 [\epsilon_k \epsilon_m]_1 + [\epsilon_k]_0 [\epsilon_i \epsilon_m]_1 + [\epsilon_m]_0 [\epsilon_i \epsilon_k]_1 \\ &\quad - 2([\epsilon_i]_1 [\epsilon_k]_0 [\epsilon_m]_0 + [\epsilon_i]_0 [\epsilon_k]_1 [\epsilon_m]_0 \\ &\quad + [\epsilon_i]_0 [\epsilon_k]_0 [\epsilon_m]_1), \end{aligned} \quad (92)$$

$$\partial_t [\epsilon_i]_2 = J_i^w [\epsilon_w]_2 + \frac{1}{2} J_i^{wz} [\epsilon_w \epsilon_z]_1 - \frac{1}{2} J_i^{w(2)} [\epsilon_w]_0, \quad (93)$$

$$\begin{aligned} \partial_t [\epsilon_i \epsilon_k]_2 &= J_k^w [\epsilon_i \epsilon_w]_2 + \frac{1}{2} J_k^{wz} [\epsilon_i \epsilon_w \epsilon_z]_1 - \frac{1}{2} J_k^{w(2)} \phi_w [\epsilon_i]_1 \\ &\quad - \frac{1}{2} J_k^{w(2)} [\epsilon_i \epsilon_w]_0 + (i \leftrightarrow k) + J_{ik}^w [\epsilon_w]_1 \\ &\quad + \frac{1}{2} J_{ik}^{wz} [\epsilon_w \epsilon_z]_0 - \frac{1}{2} J_{ik}^{w(2)} \phi_w, \end{aligned} \quad (94)$$

$$\partial_t [\epsilon_i]_3 = J_i^w [\epsilon_w]_3 + \frac{1}{2} J_i^{wz} [\epsilon_w \epsilon_z]_2 - \frac{1}{2} J_i^{w(2)} [\epsilon_w]_1. \quad (95)$$

A comparison of the above equations with those obtained from the SSE, Eqs. (B11)–(B20), shows that the two sets of equations agree except for the pair of equations (89) and (B13) and the pair of equations (92) and (B17). The latter two pairs of equations are those of the third moments of the  $\epsilon$  variables, which of course are incorrectly predicted by the 2MA due to the *ad hoc* setting of the central third-order moments to zero.

Now deterministic initial conditions imply that all pseudo-moments are initially zero which in turn implies by the above equations that  $[\epsilon_i]_0 = [\epsilon_i \epsilon_j \epsilon_k]_0 = [\epsilon_i \epsilon_k]_1 = [\epsilon_i]_2 = 0$  for all times (see Appendix B for a discussion). Hence from Eq. (85), we can deduce that the first correction to the RE prediction of the mean concentration of species  $i$  is given by the term  $\Omega^{-1} [\epsilon_i]_1$ . The latter is obtained by solving the coupled time-evolution equations, Eqs. (88) and (90). Note that these equations do not depend on the third moments of the  $\epsilon$  variables and hence the 2MA estimate of the correction

to the mean concentration agrees with that predicted by the SSE which implies that it is accurate. The pair of equations, Eqs. (88) and (90), has been previously derived by Grima,<sup>15</sup> they have been solved for several monostable systems, including dimerization, trimerization, and enzyme-catalyzed reactions and their accuracy has been verified by comparison with the mean concentrations obtained from ensemble averaging the trajectories produced by the stochastic simulation algorithm.<sup>15,19–22</sup> By the deterministic initial conditions, it also follows that the second correction to the RE prediction of the mean concentration of species  $i$  is given by the term  $\Omega^{-2}[\epsilon_i]_3$ . The time-evolution equation for this quantity, Eq. (95), is dependent on the third moments of the  $\epsilon$  variables and hence the 2MA's prediction of this second correction cannot be accurate. Thus, the 2MA accurately predicts the first-order correction to the REs but cannot capture the second-order correction to the same.

We finally turn our attention to the 2MA prediction of the covariance. Deterministic initial conditions imply that the first two non-zero terms of the series expansion for the covariance of species  $i$  and  $k$ , Eq. (86), are given by  $\Omega^{-1}[\epsilon_i\epsilon_k]_0$  and  $\Omega^{-2}([\epsilon_i\epsilon_k]_2 - [\epsilon_i]_1[\epsilon_k]_1)$ . Now the equation for  $[\epsilon_i\epsilon_k]_0$ , Eq. (88), is exactly the same as that given by the LNA, Eq. (B12) in Appendix B. Hence in the limit of large volumes the 2MA agrees with the LNA, a fact which has been previously established by Ferm *et al.*<sup>12</sup> The correction to the LNA is provided by the term proportional to  $\Omega^{-2}$  in Eq. (86). This term depends on  $[\epsilon_i\epsilon_k]_2$  which by Eq. (94) depends on the knowledge of the third moment  $[\epsilon_i\epsilon_w\epsilon_z]_1$ , a quantity which we have already established to be incorrectly predicted by the 2MA. Hence the 2MA cannot accurately capture the correction to the LNA's prediction of covariance.

A similar calculation can be performed for the 3MA whose general form is derived in Appendix C and is given by Eqs. (C1)–(C3) together with the condition Eq. (C7). Substitution of the ansatz Eq. (79) and the microscopic rate function Eq. (80) leads to Eqs. (81) and (82), a new time-evolution for  $\langle\epsilon_i\epsilon_k\epsilon_m\rangle$  and an expression for  $\langle\epsilon_i\epsilon_k\epsilon_m\epsilon_r\rangle$ . Upon making the series expansion Eq. (84), one obtains equations for the pseudo-moments. These are found to be precisely the same as those obtained from the SSE, Eqs. (B11)–(B20). Hence it follows by Eqs. (85) and (86), that the 3MA equations lead to mean concentration and variance predictions accurate to order  $\Omega^{-2}$ : two orders higher than the REs and one order higher than the LNA.

## V. CONCLUSION

Our results confirm and provide theoretical support for the observation, by a number of numerical studies,<sup>7,12</sup> that the 2MA's mean concentration prediction is a better approximation to that of the CME than the prediction of REs. We have shown that the underlying reason for this higher accuracy of the 2MA is simply that its first-order correction to the RE prediction agrees with that from the SSE of the CME which is systematic and not based on the *ad hoc* setting of some cumulants to zero. This also implies that in the limit of large reaction volumes, the mean concentration solution of the 2MA approaches that of EMREs,<sup>15</sup> effective mesoscopic rate equa-

tions which have been shown to be more accurate than REs but which are not based on a moment-closure approximation technique.

Our results also show, for the first time, that the 2MA is not the ideal moment-closure approximation if one needs to obtain information about the variances and covariances of concentration fluctuations at intermediate volumes. The reason is that at these volumes, corrections to the LNA cannot be neglected and the 2MA predicts the wrong first-order correction to the LNA. The 3MA is the method of choice in such cases. We emphasize that our results are based on a large volume expansion and hence the warranty behind the use of the 3MA as an accurate alternative to both the REs and the LNA is only valid for large enough volumes. We note that the approach developed in this paper is also applicable to study moment-closure approximations other than the ones considered in this paper, e.g., the separable derivative-matching moment closure of Singh and Hespanha.<sup>23</sup> The major limitation of our approach is that results can only be obtained for monostable systems since the ansatz leading to the SSE is only meaningful for such systems. Biological systems can display bistable behavior (see, for example, Refs. 24 and 25) and hence the extension of our results to these systems is an open and interesting research problem.

We finish by noting that our results also enable us to obtain the dependence of the leading order relative errors in the predictions of moment-closure approximation methods, of the REs, and of the LNA on the reaction volume. The relative error in the mean or variance/covariance prediction of a method is defined as the solution from the SSE minus the solution from the method divided by the solution from the SSE. Since the mean concentrations and the variances and covariances of fluctuations from the 2MA and the 3MA agree to order  $\Omega^{-2}$ , respectively, with those from the SSE, then it follows by the above definition, that to leading order: (i) the relative errors in the mean concentration predictions of the REs, the 2MA, and the 3MA scale as  $\Omega^{-1}$ ,  $\Omega^{-2}$ , and  $\Omega^{-3}$ , respectively; (ii) the relative errors in the variance and covariance of concentration fluctuations prediction of the LNA, the 2MA, and the 3MA scale as  $\Omega^{-1}$ ,  $\Omega^{-1}$ , and  $\Omega^{-2}$ , respectively. Recently relative errors have also been obtained for the chemical Langevin/Fokker-Planck approximation. For non-detailed balance conditions, these scale as  $\Omega^{-2}$  for the mean concentrations and  $\Omega^{-1}$  for the variances and covariances of fluctuations.<sup>18</sup> Hence the 2MA and the chemical Langevin equation approximations are expected to typically give results of comparable accuracy, while the 3MA is expected to fare better than both.

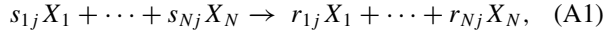
## ACKNOWLEDGMENTS

R.G. acknowledges support by SULSA (Scottish Universities Life Science Alliance).

## APPENDIX A: PRELIMINARIES: CME AND REMODELING

We consider a general molecular reaction network in a well-stirred volume  $\Omega$  involving the interaction of  $N$  distinct

chemical species via  $R$  reactions of the type



where  $j = 1, \dots, R$ ,  $X_i$  denotes species  $i$  ( $i = 1, \dots, N$ ), and  $s_{ij}$  and  $r_{ij}$  are the stoichiometric coefficients. The  $j$ th reaction is characterized by the rate constant  $k_j$ . The constraint  $\sum_i s_{ij} \leq 2$  ensures that each reaction is at most bimolecular and hence elementary.

The macroscopic description of this system is deterministic and is given by REs of the form

$$\frac{\partial \phi_i(t)}{\partial t} = \sum_{j=1}^R (r_{ij} - s_{ij}) f_j(\vec{\phi}(t)), \quad i = 1, \dots, N, \quad (\text{A2})$$

where  $\vec{\phi}(t) = (\phi_1(t), \dots, \phi_N(t))^T$  is the column vector of macroscopic concentrations at time  $t$ , and  $f_j(\vec{\phi}(t))$  is the macroscopic rate function which has a form dictated by the law of the mass action:  $f_j(\vec{\phi}(t)) = k_j \prod_{w=1}^N \phi_w^{s_{wj}}(t)$ . This framework provides only the mean concentrations and is suitable as an accurate modeling approach in cases where intrinsic noise can be safely ignored, e.g., when the population numbers of all species are considerably large.

On mesoscopic length scales, intrinsic noise becomes important and a new approach is needed to realistically capture the dynamics of the chemical system. The formalism which provides a physically correct picture is that of CMEs which have the general form

$$\frac{\partial P(\vec{n}, t)}{\partial t} = \Omega \sum_{j=1}^R \left( \prod_{i=1}^N E_i^{-S_{ij}} - 1 \right) \hat{f}_j(\vec{n}, \Omega) P(\vec{n}, t), \quad (\text{A3})$$

where  $S_{ij} = r_{ij} - s_{ij}$  (the  $i - j$  element of the stoichiometric matrix  $S$ ). The CME is a time-evolution equation for the probability  $P(\vec{n}, t)$  of the system being in a particular mesoscopic state  $\vec{n} = (n_1, \dots, n_N)^T$ , where  $n_i$  is the discrete number of molecules of species  $i$ . The CME form is made considerably compact by the use of the step operator  $E_i^{-S_{ij}}$  which is defined by its action on a general function  $g$  of the absolute number of molecules

$$E_i^{-S_{ij}} g(n_1, \dots, n_i, \dots, n_N) = g(n_1, \dots, n_i - S_{ij}, \dots, n_N). \quad (\text{A4})$$

The probability that the  $j$ th reaction occurs in the time interval  $[t, t + dt)$  is given by  $\Omega \hat{f}_j(\vec{n}, \Omega) dt$ , where  $\hat{f}_j(\vec{n}, \Omega)$  is the microscopic rate function. The function  $\Omega \hat{f}_j(\vec{n}, \Omega)$  is also commonly referred to as the propensity function.<sup>17</sup>

For elementary reactions, the microscopic rate function takes one of four different forms, depending on the order of the  $j$ th reaction: (i) a zeroth-order reaction by which a species is input into a compartment gives  $\hat{f}_j(\vec{n}, \Omega) = k_j$ ; (ii) a first-order unimolecular reaction involving the decay of some species  $h$  gives  $\hat{f}_j(\vec{n}, \Omega) = k_j n_h \Omega^{-1}$ ; (iii) a second-order bimolecular reaction between two molecules of the same species  $h$  gives  $\hat{f}_j(\vec{n}, \Omega) = k_j n_h (n_h - 1) \Omega^{-2}$ ; (iv) a second-order bimolecular reaction between two molecules of different species,  $h$  and  $v$ , gives  $\hat{f}_j(\vec{n}, \Omega) = k_j n_h n_v \Omega^{-2}$ .

## APPENDIX B: THE MULTIVARIATE SYSTEM-SIZE EXPANSION

### 1. Ansatz and the transformed master equation

The starting point of the SSE is to write the absolute number of molecules of species  $i$  as

$$\frac{n_i}{\Omega} = \phi_i + \Omega^{-1/2} \epsilon_i, \quad (\text{B1})$$

where  $\phi_i$  is the macroscopic concentration of species  $i$  as determined by the REs. Substitution of this ansatz in the CME, Eq. (A3), leads to a new time-evolution equation for the probability density function  $\Pi(\vec{\epsilon}, t)$  which reads (see Ref. 26 for a derivation)

$$\frac{\partial \Pi(\vec{\epsilon}, t)}{\partial t} = \sum_{j=1}^R \sum_{k=1}^{\infty} (-1)^k a_j^k \Omega^{-k/2} \left( (1 - \delta_{k,1} \delta_{m,0}) \sum_{m=0}^2 b_j^m \Omega^{1-m/2} + (c_j^2 + c_j^3 \Omega^{-1/2}) \right) \Pi(\vec{\epsilon}, t), \quad (\text{B2})$$

where

$$a_j^k = \frac{1}{k!} \left( \sum_{i=1}^N S_{ij} \frac{\partial}{\partial \epsilon_i} \right)^k, \quad (\text{B3})$$

$$b_j^k = \frac{1}{k!} \left( \sum_{w=1}^N \epsilon_w \frac{\partial}{\partial \phi_w} \right)^k f_j(\vec{\phi}), \quad (\text{B4})$$

$$c_j^2 = -\frac{1}{2} \sum_{w=1}^N \phi_w \frac{\partial^2 f_j(\vec{\phi})}{\partial \phi_w^2}, \quad (\text{B5})$$

$$c_j^3 = -\frac{1}{2} \sum_{w=1}^N \epsilon_w \frac{\partial^2 f_j(\vec{\phi})}{\partial \phi_w^2}. \quad (\text{B6})$$

### 2. Perturbative solution of the transformed master equation

This partial differential equation admits a power series solution of the form  $\Pi(\vec{\epsilon}, t) = \sum_{j=0}^{\infty} \Pi_j(\vec{\epsilon}, t) \Omega^{-j/2}$ . Note that although  $\Pi(\vec{\epsilon}, t)$  is a true probability density (positive and its integral over  $\vec{\epsilon}$  equal to one), the functions  $\Pi_j(\vec{\epsilon}, t)$  have the freedom to take positive or negative values and hence these functions (with the exception of  $\Pi_0(\vec{\epsilon}, t)$ ) cannot be interpreted as probability densities themselves. It also follows that the moments have a corresponding expansion

$$\langle \epsilon_k \epsilon_m \dots \epsilon_r \rangle = \sum_{j=0}^{\infty} [\epsilon_k \epsilon_m \dots \epsilon_r]_j \Omega^{-j/2}, \quad (\text{B7})$$

where

$$[\epsilon_k \epsilon_m \dots \epsilon_r]_j = \int \epsilon_k \epsilon_m \dots \epsilon_r \Pi_j(\vec{\epsilon}, t) d\vec{\epsilon}. \quad (\text{B8})$$

We refer to the latter quantities as pseudo-moments since these cannot be interpreted as moments because generally  $\Pi_j(\vec{\epsilon}, t)$  are not probability densities themselves.

It can be shown using Eqs. (B1) and (B7) that the mean concentration and the covariance of the concentration fluctuations are related to the pseudo-moments by the equations

$$\mu_r = \left\langle \frac{n_r}{\Omega} \right\rangle = \phi_r + \Omega^{-1/2} \langle \epsilon_r \rangle = \phi_r + \Omega^{-1/2} \sum_{j=0}^{\infty} [\epsilon_r]_j \Omega^{-j/2}, \quad (\text{B9})$$

$$\begin{aligned} \sigma_{rk}^2 &= \left\langle \left( \frac{n_r}{\Omega} - \left\langle \frac{n_r}{\Omega} \right\rangle \right) \left( \frac{n_k}{\Omega} - \left\langle \frac{n_k}{\Omega} \right\rangle \right) \right\rangle = \Omega^{-1} (\langle \epsilon_r \epsilon_k \rangle - \langle \epsilon_r \rangle \langle \epsilon_k \rangle) \\ &= \Omega^{-1} \left( \sum_{j=0}^{\infty} [\epsilon_r \epsilon_k]_j \Omega^{-j/2} - \sum_{j,m=0}^{\infty} [\epsilon_r]_j [\epsilon_k]_m \Omega^{-(j+m)/2} \right). \end{aligned} \quad (\text{B10})$$

In what follows we will determine the mean concentrations and covariance of concentration fluctuations accurate to order  $\Omega^{-2}$ . The procedure involves obtaining time-evolution equations for the pseudo-moments, solving them and then substituting into Eqs. (B9) and (B10). The time-evolution equation for the pseudo-moment  $[\epsilon_k \epsilon_m \dots \epsilon_r]_j$  is obtained as follows. One substitutes  $\Pi(\vec{\epsilon}, t) = \sum_{j=0}^{\infty} \Pi_j(\vec{\epsilon}, t) \Omega^{-j/2}$  in Eq. (B2), multiplies the resulting equation on both sides by  $\epsilon_k \epsilon_m \dots \epsilon_r$ , integrates over  $\vec{\epsilon}$  and finally equates terms of order  $\Omega^{-j/2}$  on both sides of the equation. This procedure leads to the following set of time-evolution equations (see Ref. 18 for details of the derivation)

$$\partial_t [\epsilon_r]_0 = J_r^w [\epsilon_w]_0, \quad (\text{B11})$$

$$\partial_t [\epsilon_r \epsilon_k]_0 = J_r^w [\epsilon_w \epsilon_k]_0 + (r \leftrightarrow k) + D_{rk}, \quad (\text{B12})$$

$$\begin{aligned} \partial_t [\epsilon_r \epsilon_k \epsilon_l]_0 &= J_l^w [\epsilon_w \epsilon_k \epsilon_r]_0 + (l \leftrightarrow k) + (k \leftrightarrow r) \\ &+ D_{rl} [\epsilon_k]_0 + (k \leftrightarrow l) + (r \leftrightarrow l), \end{aligned} \quad (\text{B13})$$

$$\begin{aligned} \partial_t [\epsilon_r \epsilon_k \epsilon_l \epsilon_m]_0 &= J_r^w [\epsilon_w \epsilon_k \epsilon_l \epsilon_m]_0 + (r \leftrightarrow m) + (m \leftrightarrow k) \\ &+ (k \leftrightarrow l) + D_{rm} [\epsilon_k \epsilon_l]_0 \\ &+ (m \leftrightarrow l) + (l \leftrightarrow k) + (r \leftrightarrow m) \\ &+ (m \leftrightarrow l) + (k \leftrightarrow m), \end{aligned} \quad (\text{B14})$$

$$\partial_t [\epsilon_r]_1 = J_r^w [\epsilon_w]_1 + \frac{1}{2} J_r^{wp} [\epsilon_w \epsilon_p]_0 - \frac{1}{2} J_r^{w(2)} \phi_w, \quad (\text{B15})$$

$$\begin{aligned} \partial_t [\epsilon_r \epsilon_k]_1 &= J_r^w [\epsilon_w \epsilon_k]_1 + \frac{1}{2} J_r^{wp} [\epsilon_w \epsilon_p \epsilon_k]_0 - \frac{1}{2} J_r^{w(2)} \phi_w [\epsilon_k]_0 \\ &+ (r \leftrightarrow k) + J_{kr}^w [\epsilon_w]_0, \end{aligned} \quad (\text{B16})$$

$$\begin{aligned} \partial_t [\epsilon_r \epsilon_k \epsilon_l]_1 &= J_l^w [\epsilon_w \epsilon_k \epsilon_r]_1 + \frac{1}{2} J_l^{wp} [\epsilon_w \epsilon_p \epsilon_r \epsilon_k]_0 \\ &- \frac{1}{2} J_l^{w(2)} \phi_w [\epsilon_r \epsilon_k]_0 + (l \leftrightarrow k) + (k \leftrightarrow r) \\ &+ D_{rl} [\epsilon_k]_1 + J_{rl}^w [\epsilon_w \epsilon_k]_0 + (k \leftrightarrow l) \\ &+ (r \leftrightarrow l) + D_{rkl}, \end{aligned} \quad (\text{B17})$$

$$\partial_t [\epsilon_r]_2 = J_r^w [\epsilon_w]_2 + \frac{1}{2} J_r^{wp} [\epsilon_w \epsilon_p]_1 - \frac{1}{2} J_r^{w(2)} [\epsilon_w]_0, \quad (\text{B18})$$

$$\begin{aligned} \partial_t [\epsilon_r \epsilon_k]_2 &= J_r^w [\epsilon_w \epsilon_k]_2 + \frac{1}{2} J_r^{wp} [\epsilon_w \epsilon_p \epsilon_k]_1 - \frac{1}{2} J_r^{w(2)} \phi_w [\epsilon_k]_1 \\ &- \frac{1}{2} J_r^{w(2)} [\epsilon_w \epsilon_k]_0 + (r \leftrightarrow k) + J_{kr}^w [\epsilon_w]_1 \\ &+ \frac{1}{2} J_{rk}^{wm} [\epsilon_w \epsilon_m]_0 - \frac{1}{2} J_{rk}^{w(2)} \phi_w, \end{aligned} \quad (\text{B19})$$

$$\partial_t [\epsilon_r]_3 = J_r^w [\epsilon_w]_3 + \frac{1}{2} J_r^{wp} [\epsilon_w \epsilon_p]_2 - \frac{1}{2} J_r^{w(2)} [\epsilon_w]_1. \quad (\text{B20})$$

Note that in the above equations, we have used the Einstein summation convention where all twice repeated indices are understood to be summed over 1 to  $N$ . Note also that the short-hand notation  $(r \leftrightarrow k)$  stands for all the expressions of the same form as the ones preceding the notation but with  $r$  and  $k$  interchanged. For example, in Eq. (B16),  $(r \leftrightarrow k)$  stands for  $J_k^w [\epsilon_w \epsilon_r]_1 + \frac{1}{2} J_k^{wp} [\epsilon_w \epsilon_p \epsilon_r]_0 - \frac{1}{2} J_k^{w(2)} \phi_w [\epsilon_r]_0$ . We have also used the following two convenient definitions

$$D_{ij..r} = \sum_{k=1}^R S_{ik} S_{jk} \dots S_{rk} f_k(\vec{\phi}), \quad (\text{B21})$$

$$J_{ij..r}^{st..z} = \frac{\partial}{\partial \phi_s} \frac{\partial}{\partial \phi_t} \dots \frac{\partial}{\partial \phi_z} D_{ij..r}, \quad J_{ij..r}^{s(2)} = J_{ij..r}^{ss}. \quad (\text{B22})$$

Note that these quantities can be directly constructed from the REs and from knowledge of the stoichiometry of the chemical system. In particular it is easy to deduce that  $D_i = \partial \phi_i / \partial t$  and  $J_i^s$  represents the  $i$ - $s$  element of the Jacobian matrix associated with the REs of the system.

### 3. Initial conditions

We impose the simplest and most commonly used initial conditions: deterministic initial conditions, i.e., at time  $t = 0$ , the vector  $\vec{n} / \Omega$  is set to be exactly equal to the macroscopic concentration vector  $\vec{\phi}$ . This implies that there is no noise initially and that consequently all pseudo-moments at  $t = 0$  are zero. This condition implies by Eqs. (B11), (B13), (B16), and (B18) that  $[\epsilon_r]_0 = [\epsilon_r \epsilon_k \epsilon_l]_0 = [\epsilon_r \epsilon_k]_1 = [\epsilon_r]_2 = 0$  for all times. Hence Eqs. (B9) and (B10) take the simplified form

$$\mu_r = \phi_r + [\epsilon_r]_1 \Omega^{-1} + [\epsilon_r]_3 \Omega^{-2} + O(\Omega^{-3}), \quad (\text{B23})$$

$$\sigma_{rk}^2 = \Omega^{-1} [\epsilon_r \epsilon_k]_0 + \Omega^{-2} ([\epsilon_r \epsilon_k]_2 - [\epsilon_r]_1 [\epsilon_k]_1) + O(\Omega^{-3}). \quad (\text{B24})$$

### 4. The SSE based methods

All SSE methods are based on the concentration solution of the REs; this is necessarily the case due to the form of the ansatz Eq. (B1) upon which the expansion is based. Having solved the REs, one can then proceed to one of the three approximations.

The linear-noise approximation (LNA) involves calculating the covariances to lowest order, i.e., accurate to order  $\Omega^{-1}$ . This implies solving Eq. (B12) for all the pseudo-moments  $[\epsilon_w \epsilon_k]_0$ , where  $w$  and  $k$  take values between 1 and  $N$ . The covariance of the concentration fluctuations of species  $r$  and  $k$  is



then obtained by substituting the latter in Eq. (B24), truncated to order  $\Omega^{-1}$ .

A second approximation is called the effective mesoscopic rate equation approximation (EMRE).<sup>15</sup> This involves solving equations (B12) and (B15) and substituting the solutions in Eq. (B23) (truncated to order  $\Omega^{-1}$ ). The EMRE provides the lowest order correction to the RE prediction of concentration.

A more accurate approximation is obtained by computing both the mean and variance accurate to order  $\Omega^{-2}$ . This involves simultaneously solving the set of equations (B12), (B14), (B15), (B17), (B19), and (B20) and finally substituting the solutions in Eqs. (B23) and (B24). This method provides a second-order correction to the mean concentration of the RE and a first-order correction to the covariance predictions of the LNA.

### APPENDIX C: GENERAL EXPRESSIONS FOR THE 2MA AND THE 3MA

In this section, we derive the 2MA and 3MA equations for the general chemical reaction network of Appendix A. Multiplying both sides of Eq. (A3) by  $n_i/\Omega$  and summing over  $\vec{n}$  we obtain an exact time-evolution equation for the mean concentration of species  $i$

$$\partial_t \frac{\langle n_i \rangle}{\Omega} = \sum_{j=1}^R S_{ij} \langle \hat{f}_j(\vec{n}, \Omega) \rangle. \quad (\text{C1})$$

As mentioned in Appendix A, for a system of elementary chemical reactions, the microscopic rate function is generally a quadratic in the components of  $\vec{n}$  and hence from Eq. (C1) we can see that the determination of the mean concentrations requires knowledge of the second moments. We can obtain a time-evolution equation for these moments by multiplying both sides of Eq. (A3) by  $n_i n_k / \Omega^2$  and then summing over  $\vec{n}$ , which leads to

$$\begin{aligned} \partial_t \frac{\langle n_i n_k \rangle}{\Omega^2} &= \Omega^{-1} \sum_{j=1}^R (S_{kj} \langle n_i \hat{f}_j(\vec{n}, \Omega) \rangle + S_{ij} \langle n_k \hat{f}_j(\vec{n}, \Omega) \rangle \\ &+ S_{ij} S_{kj} \langle \hat{f}_j(\vec{n}, \Omega) \rangle). \end{aligned} \quad (\text{C2})$$

Examination of this equation shows that its solution requires knowledge of the third moments. A time-evolution equation for the latter moments can be obtained by multiplying both sides of Eq. (A3) by  $n_i n_k n_m / \Omega^3$  and summing over  $\vec{n}$  which leads to

$$\begin{aligned} \partial_t \frac{\langle n_i n_k n_m \rangle}{\Omega^3} &= \Omega^{-2} \sum_{j=1}^R (S_{mj} \langle n_i n_k \hat{f}_j(\vec{n}, \Omega) \rangle \\ &+ S_{kj} \langle n_i n_m \hat{f}_j(\vec{n}, \Omega) \rangle + S_{ij} \langle n_k n_m \hat{f}_j(\vec{n}, \Omega) \rangle \\ &+ S_{kj} S_{mj} \langle n_i \hat{f}_j(\vec{n}, \Omega) \rangle + S_{ij} S_{mj} \langle n_k \hat{f}_j(\vec{n}, \Omega) \rangle \\ &+ S_{ij} S_{kj} \langle n_m \hat{f}_j(\vec{n}, \Omega) \rangle \\ &+ S_{ij} S_{kj} S_{mj} \langle \hat{f}_j(\vec{n}, \Omega) \rangle). \end{aligned} \quad (\text{C3})$$

Since the microscopic rate function is a quadratic in the components of  $\vec{n}$  for elementary reaction systems, it follows that

the solution of the above equation requires knowledge of the fourth moments. Hence we can deduce a general pattern: the time-evolution equation for the  $M$ th moment is a function of the  $(M + 1)$ th moment. This means we have an infinite hierarchy of exact equations which cannot be solved.

A way out of this impasse is to use the M-moment approximation scheme in which the  $(M + 1)$ th central moment is assumed to be same as that from a normal distribution. This assumption breaks the infinite hierarchy of equations and leads to a finite number of simultaneous equations which can be solved to obtain an approximate solution for the first  $M$  moments.

The simplest of these class of approximations is the 2MA in which one solves Eqs. (C1) and (C2) with the third central moments (the third-order cumulants) set equal to zero

$$\langle (n_i - \langle n_i \rangle)(n_k - \langle n_k \rangle)(n_m - \langle n_m \rangle) \rangle = 0. \quad (\text{C4})$$

Note that this is equivalent to assuming that the (non-centered) third moments are given by

$$\begin{aligned} \langle n_i n_k n_m \rangle &= \langle n_i \rangle \langle n_k n_m \rangle + \langle n_k \rangle \langle n_i n_m \rangle + \langle n_m \rangle \langle n_i n_k \rangle \\ &- 2 \langle n_i \rangle \langle n_k \rangle \langle n_m \rangle. \end{aligned} \quad (\text{C5})$$

The next order approximation is the 3MA in which one solves Eqs. (C1)–(C3) with the fourth central moments set equal to that of a Gaussian (the fourth-order cumulants set to zero)

$$\begin{aligned} \langle (n_i - \langle n_i \rangle)(n_k - \langle n_k \rangle)(n_m - \langle n_m \rangle)(n_r - \langle n_r \rangle) \rangle \\ &= (\langle n_i n_k \rangle - \langle n_i \rangle \langle n_k \rangle)(\langle n_m n_r \rangle - \langle n_m \rangle \langle n_r \rangle) \\ &+ (\langle n_i n_m \rangle - \langle n_i \rangle \langle n_m \rangle)(\langle n_k n_r \rangle - \langle n_k \rangle \langle n_r \rangle) \\ &+ (\langle n_i n_r \rangle - \langle n_i \rangle \langle n_r \rangle)(\langle n_k n_m \rangle - \langle n_k \rangle \langle n_m \rangle). \end{aligned} \quad (\text{C6})$$

This result was first proved by Isserlis.<sup>27</sup> The above is equivalent to assuming that the (non-centered) fourth moments are given by

$$\begin{aligned} \langle n_i n_k n_m n_r \rangle &= \langle n_i n_m n_r \rangle \langle n_k \rangle + \langle n_i n_r \rangle \langle n_k n_m \rangle + \langle n_i \rangle \langle n_k n_m n_r \rangle \\ &+ \langle n_i n_m \rangle \langle n_k n_r \rangle + \langle n_i n_k n_r \rangle \langle n_m \rangle \\ &+ \langle n_i n_k \rangle \langle n_m n_r \rangle + \langle n_i n_k n_m \rangle \langle n_r \rangle \\ &- 2(\langle n_i n_r \rangle \langle n_k \rangle \langle n_m \rangle + \langle n_k n_r \rangle \langle n_i \rangle \langle n_m \rangle + \langle n_m n_r \rangle \\ &\times \langle n_k \rangle \langle n_i \rangle + (\langle n_i n_m \rangle \langle n_k \rangle + \langle n_i n_k \rangle \langle n_m \rangle \\ &+ \langle n_i \rangle (\langle n_k n_m \rangle - 3 \langle n_k \rangle \langle n_m \rangle)) \langle n_r \rangle). \end{aligned} \quad (\text{C7})$$

<sup>1</sup>I. G. Darvey and P. J. Staff, *J. Chem. Phys.* **44**, 990 (1966).

<sup>2</sup>C. Gadgil, C. H. Lee, and H. G. Othmer, *Bull. Math. Biol.* **67**, 901 (2005).

<sup>3</sup>W. J. Heuett and H. Qian, *J. Chem. Phys.* **124**, 044110 (2006).

<sup>4</sup>T. Jahnke and W. Huisinga, *J. Math. Biol.* **54**, 1 (2007).

<sup>5</sup>I. J. Laurenzi, *J. Chem. Phys.* **113**, 3315 (2000).

<sup>6</sup>H. Haken, *Synergetics* (Springer-Verlag, 1978).

<sup>7</sup>C. A. Gomez-Urbe and G. C. Verghese, *J. Chem. Phys.* **126**, 024109 (2007).

<sup>8</sup>N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, 2007).

<sup>9</sup>P. Whittle, *J. R. Stat. Soc. Ser. B (Methodol.)* **19**, 268 (1957).

<sup>10</sup>D. A. McQuarrie, C. J. Jachimowski, and M. E. Russell, *J. Chem. Phys.* **40**, 2914 (1964).

- <sup>11</sup>J. Goutsias, *Biophys. J.* **92**, 2350 (2007).
- <sup>12</sup>L. Ferm, P. Lotstedt, and A. Hellander, *J. Sci. Comput.* **34**, 127 (2008).
- <sup>13</sup>M. Ullah and O. Wolkenhauer, *J. Theor. Biol.* **260**, 340 (2009).
- <sup>14</sup>C. S. Gillespie, *IET Syst. Biol.* **3**, 52 (2009).
- <sup>15</sup>R. Grima, *J. Chem. Phys.* **133**, 035101 (2010).
- <sup>16</sup>I. G. Darvey, B. W. Ninham, and P. J. Staff, *J. Chem. Phys.* **45**, 2145 (1966).
- <sup>17</sup>D. T. Gillespie, *Annu. Rev. Phys. Chem.* **58**, 35 (2007).
- <sup>18</sup>R. Grima, P. Thomas, and A. Straube, *J. Chem. Phys.* **135**, 084103 (2011).
- <sup>19</sup>R. Grima, *Phys. Rev. Lett.* **102**, 218103 (2009).
- <sup>20</sup>R. Grima, *BMC Syst. Biol.* **3**, 101 (2009).
- <sup>21</sup>P. Thomas, A. V. Straube, and R. Grima, *J. Chem. Phys.* **133**, 195101 (2010).
- <sup>22</sup>R. Ramaswamy, N. Gonzalez-Segredo, I. F. Sbalzarini, and R. Grima, *Nat. Comm.* **3**, 779 (2012).
- <sup>23</sup>A. Singh and J. P. Hespanha, *IEEE Trans. Automat. Contr.* **56**, 414 (2010).
- <sup>24</sup>J. Hasty, J. Pradines, M. Dolnik, and J. J. Collins, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2075 (2000).
- <sup>25</sup>K. Lai, M. J. Robertson, and D. V. Schaffer, *Biophys. J.* **86**, 2748 (2004).
- <sup>26</sup>R. Grima, *Phys. Rev. E* **84**, 056109 (2011).
- <sup>27</sup>L. Isserlis, *Biometrika* **12**, 134 (1918).