Chemical Master versus Chemical Langevin for First-Order Reaction Networks

Desmond J. Higham* Raya Khanin†

Abstract

Markov jump processes are widely used to model interacting species in circumstances where discreteness and stochasticity are relevant. Such models have been particularly successful in computational cell biology, and in this case, the interactions are typically first-order. The Chemical Langevin Equation is a stochastic differential equation that can be regarded as an approximation to the underlying jump process. In particular, the Chemical Langevin Equation allows simulations to be performed more effectively. In this work, we obtain expressions for the first and second moments of the Chemical Langevin Equation for a generic first-order reaction network. Moreover, we show that these moments exactly match those of the underlying jump process. Hence, in terms of means, variances and correlations, the Chemical Langevin Equation is an excellent proxy for the Chemical Master Equation. Our work assumes that a unique solution exists for the Chemical Langevin Equation. We also show that the moment matching result extends to the case where a gene regulation model of Raser and O'Shea (Science, 2004) is replaced by a hybrid model that mixes elements of the Master and Langevin equations. We finish with numerical experiments on a dimerization model that involves second order reactions, showing that the two regimes continue to give similar results.

Keywords: Birth-and-death process, Chemical Master Equation, chemical kinetics, correlation matrix, Euler–Maruyama, gene regulation network, Gillespie, Ito Lemma, stochastic simulation algorithm.

*Corresponding author: Department of Mathematics, University of Strathclyde, Glasgow G1 1XH, UK. email:djh"at"maths.strath.ac.uk

†Department of Statistics, University of Glasgow, Glasgow G12 8QQ, U.K.

1874-1142/082008 Bentham Open

particular, for the hybrid model, we extend the work in [8] by showing that (a) Ito's lemma allows us to study moments for all time, not just at steady state, and (b) including a diffusion approximation allows us to recover the exact first and second moments. Overall, we regard this work as initial evidence that the CLE regime is useful, and we point out that further analysis to compare the CME and CLE in the presence of bi-molecular reactions would be very desirable.

Section 2 very briefly sets up the CME and CLE models. In section 3 we describe the four reaction types that make up the first-order networks defined by Gadgil *et al.* in [7] and we state the ODEs that those authors derived for the means, variances and correlations. The corresponding first-order network CLE is studied in section 4. By applying Ito's lemma we obtain generic ODEs for the means, variances and correlations and find that they match those for the CLE precisely. In section 5 we show that this moment matching extends to a gene regulation model where a typically small collection of genes that can switch between active and inactive states, are modeled in the CME regime whereas mRNA and proteins, typically more abundant, are modeled via the CLE. The resulting switching SDE remains amenable to a generalized version of Ito's lemma. In section 6 we give simulation results for a dimerization model that involves second-order reactions and we compare the CME and CLE behaviour.

2 Stoichiometry

This section very briefly introduces the CME and CLE models. For derivations and further details we refer the reader to [1, 2, 12, 15, 4, 5, 6].

Suppose we have chemical species $S_1, S_2, \ldots, S_N$ taking part in M different types of chemical reaction, or *reaction channels*. We will let $\mathbf{X}(t) \in \mathbb{R}^N$ denote the state vector, so $X_i(t)$ records the number of molecules of species i present at time t. This value will clearly be a non-negative integer. We assume that $\mathbf{X}(0)$ is known. Associated with each of the M possible reactions is a *stoichiometric* vector, $\boldsymbol{\nu}_j \in \mathbb{R}^N$, whose ith component is the change in the number of S_i molecules caused by the jth reaction. So one reaction of type j has the effect of changing the state vector from $\mathbf{X}(t)$ to $\mathbf{X}(t) + \boldsymbol{\nu}_j$. Each reaction also has a *propensity function*, $a_j(\mathbf{X}(t))$. Here, the probability that the jth reaction takes place in the infinitesimal time interval $[t, t + dt)$ is taken to be $a_j(\mathbf{X}(t))dt$. Letting $P(\mathbf{x}, t)$ denote the probability that $\mathbf{X}(t) = \mathbf{x}$, the CME then takes the form

$$\frac{dP(\mathbf{x},t)}{dt} = \sum_{j=1}^{M} \left(a_j(\mathbf{x} - \boldsymbol{\nu}_j) P(\mathbf{x} - \boldsymbol{\nu}_j, t) - a_j(\mathbf{x}) P(\mathbf{x}, t) \right). \tag{1}$$

tionality constant $\mathrm{k}_{ij}^{\mathrm{cat}}$. The stoichiometric vector takes the form $\boldsymbol{\nu} = \mathbf{e}_i$, with propensity function is $\mathrm{k}_{ij}^{\mathrm{cat}} X_j(t)$.

Overall, there are N possible reactions involving production from a source, N possible degradation reactions, $N(N-1)$ possible conversion reactions and N^2 possible catalytic production reactions. If any reaction is not present, then, of course, we simply set the corresponding rate constant to zero. Conversion from i to i does not make sense in this context, but it is convenient to define rate constants $\mathrm{k}_{ii}^{\mathrm{con}}$, all of which are zero. In this way, we have $M := N+N+N^2+N^2 = 2N(1+N)$ reactions.

Gadgil *et al.* used a moment generating function approach to analyse the CME (1) for this class of reactions. Letting $\mathbf{M}(t) = \mathbb{E}\left[\mathbf{X}(t)\right] \in \mathbb{R}^N$ denote the mean of the state vector $\mathbf{X}(t)$ and $\mathbf{V}(t) \in \mathbb{R}^{N\times N}$ denote a matrix of correlations defined by

$$V_{kl}(t) = \begin{cases} \mathbb{E}\left[X_k(t)X_l(t)\right], & l \neq k, \\ \mathbb{E}\left[X_k(t)^2\right] - \mathbb{E}\left[X_k(t)\right], & l = k, \end{cases}$$

they showed that [7, equation (28)]

$$\frac{d\mathbf{M}(t)}{dt} = \mathcal{K}\mathbf{M}(t) + K^{\mathrm{s}}\mathbf{1} \tag{3}$$

and [7, equation (29)]

$$\frac{d\mathbf{V}(t)}{dt} = \mathcal{K}\mathbf{V}(t) + \left(\mathcal{K}\mathbf{V}(t)\right)^T + \Gamma(t) + \Gamma^T(t). \tag{4}$$

Here, $\mathcal{K} = K^{\mathrm{con}} + K^{\mathrm{cat}} - K^{\mathrm{d}}$, with $K^{\mathrm{d}} = \mathrm{diag}(\mathrm{k}_i^{\mathrm{d}})$, $(K^{\mathrm{cat}})_{ij} = \mathrm{k}_{ij}^{\mathrm{cat}}$, and

$$(K^{\mathrm{con}})_{ij} = \begin{cases} \mathrm{k}_{ij}^{\mathrm{con}}, & i \neq j, \\ -\sum_{k=1}^{N} \mathrm{k}_{kj}^{\mathrm{con}}, & i = j, \end{cases}$$

$K^{\mathrm{s}} = \mathrm{diag}(\mathrm{k}_i^{\mathrm{s}})$, $\mathbf{1} \in \mathbb{R}^N$ denotes the vector with all components equal to 1, and

$$\Gamma_{ij}(t) = \left((K^{\mathrm{cat}})_{ij} + (K^{\mathrm{s}})_{ii}\right) M_j(t).$$

4 Moments for the Chemical Langevin Equation

For the general first-order network defined in section 3, we will write the CLE (2) in the form

$$d\mathbf{Y}(t) = \mathbf{b}\left(\mathbf{Y}(t)\right) dt + \boldsymbol{\sigma}\left(\mathbf{Y}(t)\right) d\mathbf{W}(t), \tag{5}$$

It follows that the drift coefficient $\mathbf{b}(\mathbf{Y}(t))$ in (5) has kth component given by

$$b_k(\mathbf{Y}(t)) = \mathrm{k}_k^{\mathrm{s}} - \mathrm{k}_k^{\mathrm{d}} Y_k(t) + \sum_{r=1}^{N} \mathrm{k}_{kr}^{\mathrm{con}} Y_r(t) - \sum_{r=1}^{N} \mathrm{k}_{rk}^{\mathrm{con}} Y_k(t) + \sum_{r=1}^{N} \mathrm{k}_{kr}^{\mathrm{cat}} Y_r(t). \quad (8)$$

In order to apply Ito's lemma, we must characterise the matrix $\mathbf{a}(\mathbf{Y}(t)) := \boldsymbol{\sigma}(\mathbf{Y}(t))\boldsymbol{\sigma}^T(\mathbf{Y}(t)) \in \mathbb{R}^{N\times N}$. The diffusion coefficient $\boldsymbol{\sigma}(\mathbf{Y}(t)) \in \mathbb{R}^{N\times M}$ in (5) has the general form

$$\left[\begin{array}{c|c|c|c|c|c|c|c|c} T^{\mathrm{s}} & T^{\mathrm{d}} & T^{\mathrm{con}[1]} & \cdots & T^{\mathrm{con}[N]} & T^{\mathrm{cat}[1]} & \cdots & T^{\mathrm{cat}[N]} \end{array}\right],$$

and hence

$$\mathbf{a}(\mathbf{Y}(t)) = T^{\mathrm{s}}(T^{\mathrm{s}})^T + T^{\mathrm{d}}\left(T^{\mathrm{d}}\right)^T + \sum_{r=1}^{N} T^{\mathrm{con}[r]}\left(T^{\mathrm{con}[r]}\right)^T + \sum_{r=1}^{N} T^{\mathrm{cat}[r]}\left(T^{\mathrm{cat}[r]}\right)^T, \quad (9)$$

where each submatrix is of dimension $N \times N$. More specifically, the source production and degradation reactions give rise to diagonal submatrices of the form $T^{\mathrm{s}} = \mathrm{diag}\left(\sqrt{\mathrm{k}_k^{\mathrm{s}}}\right)$ and $T^{\mathrm{d}} = \mathrm{diag}\left(\sqrt{\mathrm{k}_k^{\mathrm{d}} Y_k(t)}\right)$. Hence, $T^{\mathrm{s}}(T^{\mathrm{s}})^T = \mathrm{diag}\left(\mathrm{k}_k^{\mathrm{s}}\right)$ and $T^{\mathrm{d}}\left(T^{\mathrm{d}}\right)^T = \mathrm{diag}\left(\mathrm{k}_k^{\mathrm{d}} Y_k(t)\right)$. For the conversion reactions, the block $T^{\mathrm{con}[1]}$ relates to reactions where species S_1 is converted to another species, and has the form

$$\left[\begin{array}{ccccccc} 0 & -\sqrt{\mathrm{k}_{21}^{\mathrm{con}} Y_1(t)} & -\sqrt{\mathrm{k}_{31}^{\mathrm{con}} Y_1(t)} & \cdots & \cdots & -\sqrt{\mathrm{k}_{N1}^{\mathrm{con}} Y_1(t)} \\ & \sqrt{\mathrm{k}_{21}^{\mathrm{con}} Y_1(t)} & & & & \\ & & \sqrt{\mathrm{k}_{31}^{\mathrm{con}} Y_1(t)} & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & \sqrt{\mathrm{k}_{N1}^{\mathrm{con}} Y_1(t)} \end{array}\right].$$

Generally, the block $T^{\mathrm{con}[r]}$ relates to conversion reactions that remove species S_r, and has the form

$$\left[\begin{array}{ccccccccc} \sqrt{\mathrm{k}_{1r}^{\mathrm{con}} Y_r(t)} & & & & & & & & \\ & \sqrt{\mathrm{k}_{2r}^{\mathrm{con}} Y_r(t)} & & & & & & & \\ & & \ddots & & & & & & \\ & & & \sqrt{\mathrm{k}_{r-1,r}^{\mathrm{con}} Y_r(t)} & & & & & \\ -\sqrt{\mathrm{k}_{1r}^{\mathrm{con}} Y_r(t)} & -\sqrt{\mathrm{k}_{2r}^{\mathrm{con}} Y_r(t)} & \cdots & -\sqrt{\mathrm{k}_{r-1,r}^{\mathrm{con}} Y_r(t)} & 0 & -\sqrt{\mathrm{k}_{r+1,r}^{\mathrm{con}} Y_r(t)} & \cdots & -\sqrt{\mathrm{k}_{N,r}^{\mathrm{con}} Y_r(t)} \\ & & & & & \sqrt{\mathrm{k}_{r+1,r}^{\mathrm{con}} Y_r(t)} & & & \\ & & & & & & \ddots & & \\ & & & & & & & \sqrt{\mathrm{k}_{Nr}^{\mathrm{con}} Y_r(t)} \end{array}\right].$$

Comparing (3) and (8) we see that $\mathbb{E}[\mathbf{X}(t)]$ and $\mathbb{E}[\mathbf{Y}(t)]$ satisfy the same ODE. We note that in this first-order case, the ODE for the means is precisely the RRE system, and it is well known that the RRE then matches the mean of the CME; see [7] for a comprehensive historical overview.

Next, we write the correlation result from [7] in a more convenient form. Letting $\mathbf{U}(t) \in \mathbb{R}^{N\times N}$ denote the symmetric matrix with $U_{kl} = \mathbb{E}[X_k(t)X_l(t)]$ for $k \neq l$ and $U_{kk} = \mathbb{E}[X_k^2(t)]$, so that $\mathbf{U}(t) = \mathbf{V}(t) + \mathrm{diag}(\mathbf{M}(t))$, it follows from (3) and (4) that

$$\begin{aligned}\frac{dU_{kl}(t)}{dt} &= \mathrm{k}_k^{\mathrm{s}} M_l(t) - \mathrm{k}_k^{\mathrm{d}} U_{lk}(t) + \sum_{r=1}^{N} \mathrm{k}_{kr}^{\mathrm{con}} U_{lr}(t) \\ &\quad - U_{lk}(t)\sum_{r=1}^{N} \mathrm{k}_{rk}^{\mathrm{con}} - \mathrm{k}_{kl}^{\mathrm{con}} M_l(t) + \sum_{r=1}^{N} \mathrm{k}_{kr}^{\mathrm{cat}} U_{lr}(t) \\ &\quad + \mathrm{k}_l^{\mathrm{s}} M_k(t) - \mathrm{k}_l^{\mathrm{d}} U_{kl}(t) + \sum_{r=1}^{N} \mathrm{k}_{lr}^{\mathrm{con}} U_{kr}(t) \\ &\quad - U_{kl}(t)\sum_{r=1}^{N} \mathrm{k}_{rl}^{\mathrm{con}} - \mathrm{k}_{lk}^{\mathrm{con}} M_k(t) + \sum_{r=1}^{N} \mathrm{k}_{lr}^{\mathrm{cat}} U_{kr}(t), \quad \text{for } k \neq l, \end{aligned} \tag{12}$$

and

$$\begin{aligned}\frac{dU_{kk}(t)}{dt} &= 2\mathrm{k}_k^{\mathrm{s}} M_k(t) - 2\mathrm{k}_k^{\mathrm{d}} U_{kk}(t) + 2\sum_{r=1}^{N} \mathrm{k}_{kr}^{\mathrm{con}} U_{kr}(t) - 2U_{kk}(t)\sum_{r=1}^{N} \mathrm{k}_{rk}^{\mathrm{con}} \\ &\quad + 2\sum_{r=1}^{N} \mathrm{k}_{kr}^{\mathrm{cat}} U_{kr}(t) + \mathrm{k}_k^{\mathrm{s}} + \mathrm{k}_k^{\mathrm{d}} M_k(t) \\ &\quad + \sum_{r=1}^{N} \mathrm{k}_{kr}^{\mathrm{con}} M_r(t) + M_k(t)\sum_{r=1}^{N} \mathrm{k}_{rk}^{\mathrm{con}} + \sum_{r=1}^{N} \mathrm{k}_{kr}^{\mathrm{cat}} M_r(t). \end{aligned} \tag{13}$$

Now applying Ito's lemma [16] to $Y_k(t)Y_l(t)$, for $k \neq l$, using (8) and (11), we find that

$$\begin{aligned} d\left(Y_k(t)Y_l(t)\right) &= \left\{Y_l(t) b_k\left(\mathbf{Y}(t)\right) + Y_k(t) b_l\left(\mathbf{Y}(t)\right) + \tfrac{1}{2}\left(a_{lk}\left(\mathbf{Y}(t)\right) + a_{kl}\left(\mathbf{Y}(t)\right)\right)\right\} dt + \text{mart.}, \\ &= \Bigg\{Y_l(t)\left(\mathrm{k}_k^{\mathrm{s}} - \mathrm{k}_k^{\mathrm{d}} Y_k(t) + \sum_{r=1}^{N} \mathrm{k}_{kr}^{\mathrm{con}} Y_r(t) - Y_k(t)\sum_{r=1}^{N} \mathrm{k}_{rk}^{\mathrm{con}} + \sum_{r=1}^{N} \mathrm{k}_{kr}^{\mathrm{cat}} Y_r(t)\right) \\ &\quad + Y_k(t)\left(\mathrm{k}_l^{\mathrm{s}} - \mathrm{k}_l^{\mathrm{d}} Y_l(t) + \sum_{r=1}^{N} \mathrm{k}_{lr}^{\mathrm{con}} Y_r(t) - Y_l(t)\sum_{r=1}^{N} \mathrm{k}_{rl}^{\mathrm{con}} + \sum_{r=1}^{N} \mathrm{k}_{lr}^{\mathrm{cat}} Y_r(t)\right) \\ &\quad + \tfrac{1}{2}\left(-\mathrm{k}_{lk}^{\mathrm{con}} Y_k(t) - \mathrm{k}_{kl}^{\mathrm{con}} Y_l(t) - \mathrm{k}_{kl}^{\mathrm{con}} Y_l(t) - \mathrm{k}_{lk}^{\mathrm{con}} Y_k(t)\right)\Bigg\} dt + \text{mart.}, \end{aligned}$$

5 Multi-Scale Gene Transcription Model

Raser and O'Shea [10] were able to explain noise in eukaryotic gene expression with a CME model where DNA exists in two distinct states; active and inactive. If we assume that m gene copies are present, each of which can switch independently between the two states, then the model may expressed as

$$\left.\begin{array}{l} D_i \stackrel{k_a}{\rightarrow} D_i^\star \\ D_i \stackrel{k_d}{\leftarrow} D_i^\star \\ D_i^\star \stackrel{k_r}{\rightarrow} D_i^\star + M \end{array}\right\} \quad 1 \le i \le m. \tag{16}$$

and

$$M \stackrel{k_p}{\rightarrow} M + P \tag{17}$$

$$M \stackrel{\gamma_r}{\rightarrow} \emptyset \tag{18}$$

$$P \stackrel{\gamma_p}{\rightarrow} \emptyset. \tag{19}$$

Here D_i and $D_i^\star$ represent the ith gene copy in its inactive and active states, respectively. The first two reactions in (16) are therefore of conversion type. The third reaction in (16) models the active ith gene producing mRNA and has the form of catalytic production from a source. The reaction (17) also involves catalytic production from a source; in this case mRNA causes protein to be produced. Reactions (18) and (19) model the degradation of mRNA and protein, respectively. A simpler version of this model, where there is no inactive state, was proposed in [11], and that reference gives an interesting biologically-motivated discussion of the merits and limitations of the use of first-order CME kinetics for transcription/translation processes. ODEs for means and variances in the full CME model are derived in [9, Supplementary Information] (these could also be found from the general derivation of [7]) along with an expression for the full steady-state distribution of mRNA.

CME and CLE formulations of the system (16)–(19) were analyzed in [13], where it was shown that first and second moments of the CLE match those for the CME. That result was the motivation for the general theory of section 4.

Paszek [8] looked at the case of a single gene, $m = 1$, in (16)–(19) and considered

(a) treating gene switching with the CME, and mRNA and protein production/decay with the RRE, or, alternatively,

(b) treating gene switching and mRNA production/decay with the CME, and protein production/decay with the RRE.

and finally

$$\frac{d\mathbb{E}[MP]}{dt} = -(\gamma_r + \gamma_p)\mathbb{E}[MP] + k_p\mathbb{E}[M^2] + k_r \sum_{s=1}^{m} \mathbb{E}[D_s^\star P]. \tag{26}$$

Now we consider the case where the reversible reactions in (16) governing the state of the genes are modeled with the CME, but the transcription/translation/degradation reactions involving M and P are modeled with the CLE. We will use $\widehat{M}$ and $\widehat{P}$ to denote the time-dependent random variables that describe the amount of mRNA and protein from this hybrid model, respectively.

We will also let $r = r(t) := \sum_{i=1}^{m} D_i^\star$ record the number of active genes at time t. Note that r takes values in $\{0, 1, 2, 3, \ldots, m\}$, driven by a continuous time Markov chain. Using the CLE for mRNA and protein then produces a pair of SDEs driven by an independent Markovian switch. This type of stochastic evolution equation has been used as a model in several application areas, including mathematical finance, where, for example the market may switch from 'confident' to 'nervous' causing a change in the volatility of an asset. The recent monograph [17] discusses analytical and computational issues surrounding SDEs with switches. Of particular relevance to our work is the fact that there is a generalized version of Ito's lemma that is valid even for functions depending on $r(t)$; [17, Theorem 1.45 and Lemma 1.9].

The Markovian switching SDE for $\widehat{M}$ and $\widehat{P}$ describing the CLE takes the form

$$d\begin{bmatrix} \widehat{M} \\ \widehat{P} \end{bmatrix} = \begin{bmatrix} k_r r - \gamma_r \widehat{M} \\ k_p \widehat{M} - \gamma_p \widehat{P} \end{bmatrix} dt + \begin{bmatrix} \sqrt{k_r r} & -\sqrt{\gamma_r \widehat{M}} & 0 & 0 \\ 0 & 0 & \sqrt{k_p \widehat{M}} & -\sqrt{\gamma_p \widehat{P}} \end{bmatrix} \begin{bmatrix} dW_1 \\ dW_2 \\ dW_3 \\ dW_4 \end{bmatrix}. \tag{27}$$

In order to analyse this system, we first compute the transition rates between states of the Markovian switch. We define γ_{ij} to be the transition rate from state i to state j so that, for $j \neq i$,

$$\mathbb{P}\left(r(t+\delta) = j, \text{ given } r(t) = i\right) = \gamma_{ij}\delta + o(\delta).$$

Then $\gamma_{ii} := -\sum_{j \neq i} \gamma_{ij}$ is such that

$$\mathbb{P}\left(r(t+\delta) = i, \text{ given } r(t) = i\right) = 1 + \gamma_{ii}\delta + o(\delta).$$

For convenience, we also define $\gamma_{0,-1} = 0 = \gamma_{m,m+1}$.

The matrix of transition rates is clearly tridiagonal. For a general state i we have three non-zero transition rates. If there are currently i active genes, the

and using (31) and taking expectations,

$$\frac{d\mathbb{E}[\widehat{P}r]}{dt} = k_p\mathbb{E}[r\widehat{M}] - \gamma_p\mathbb{E}[r\widehat{P}] - k_d\mathbb{E}[r\widehat{P}] - k_a\mathbb{E}[r\widehat{P}] + k_a m\,\mathbb{E}[\widehat{P}],$$

which agrees with (25).

Ito on $\widehat{M^2}$ gives

$$d\left(\widehat{M^2}\right) = \left(2\widehat{M}\left(k_r r - \gamma_r\widehat{M}\right) + \frac{1}{2}\times 2\left(k_2 r + \gamma_r\widehat{M}\right)\right)dt + \text{mart.},$$

and hence

$$\frac{d\mathbb{E}[\widehat{M^2}]}{dt} = 2k_r\mathbb{E}[r\widehat{M}] - 2\gamma_r\mathbb{E}[\widehat{M^2}] + k_r\mathbb{E}[r] + \gamma_r\mathbb{E}[\widehat{M}],$$

agreeing with (22).

Also, Ito on $\widehat{M}\widehat{P}$ gives

$$d\left(\widehat{M}\widehat{P}\right) = \left(\widehat{P}\left(k_r r - \gamma_r\widehat{M}\right) + \widehat{M}\left(k_p\widehat{M} - \gamma_p\widehat{P}\right)\right)dt + \text{mart.},$$

and hence

$$\frac{d\mathbb{E}[\widehat{M}\widehat{P}]}{dt} = k_r\mathbb{E}[r\widehat{P}] - \gamma_r\mathbb{E}[\widehat{M}\widehat{P}] + k_p\mathbb{E}[\widehat{M^2}] - \gamma_p\mathbb{E}[\widehat{M}\widehat{P}],$$

agreeing with (26).

Finally, Ito on P^2 gives

$$d\left(\widehat{P^2}\right) = \left(2\widehat{P}\left(k_p\widehat{M} - \gamma_p\widehat{P}\right) + \tfrac{1}{2}\times 2\left(k_p\widehat{M} + \gamma_p\widehat{P}\right)\right)dt + \text{mart.},$$

and hence

$$\frac{d\mathbb{E}[\widehat{P^2}]}{dt} = 2k_p\mathbb{E}[\widehat{M}\widehat{P}] - 2\gamma_p\mathbb{E}[\widehat{P^2}] + k_p\mathbb{E}[\widehat{M}] + \gamma_p\mathbb{E}[\widehat{P}],$$

agreeing with (23)

Overall, we have shown that using the CME model for the genes and the CLE model for the mRNA and protein gives the same means, variances and correlations, for all time, as for the full CME model. This provides further support for the use of the CLE in regimes where a fully discrete simulation is not computationally feasible.

an appropriate stoichiometric matrix takes the form

$$\left[\begin{array}{cccc} \boldsymbol{\nu}_1 & \boldsymbol{\nu}_2 & \ldots & \boldsymbol{\nu}_9 \end{array}\right] = \left[\begin{array}{rrrrrrrrr} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & -2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 \end{array}\right]$$

and the propensity functions are $a_1 = k_a D$, $a_2 = k_d D^*$, $a_3 = k_r D^*$, $a_4 = \gamma_r M$, $a_5 = k_p M$, $a_6 = \gamma_p P$, $a_7 = k_1 P(P-1)/2$, $a_8 = k_{-1} P_2$, $a_9 = \gamma_{P_2} P_2$.

Our main source for rate constants is [19], where the authors used a specific implementation of the genetic circuit in *E.coli* to choose biologically reasonable values. The rate constant of transcription $k_r = 0.0078 s^{-1}$ is similar to the one used in [11]. The typical mRNA half-life is 3min ($\gamma_r = 0.0039 s^{-1}$). The average number of proteins per mRNA is fixed at $b = 11$ ($k_p = b\gamma_r$). The protein monomer decay rate is $\gamma_p = 0.0007 s^{-1}$, which is chosen to match the degradation rate due to cell-growth induced dilution [19]. According to [18], many dimers are diluted by cell division in the rapid exponential growth phase (≈ 50 min half-life). For the dimerization reaction, [19] made a reasonable choice $k_1 = 0.025 s^{-1}(nM)^{-1}$ and $k_{-1} = 0.5 s^{-1}$, noticing that the actual values of these rate constants do not matter as long as they are larger than those for the other reactions in the system, with the ratio between k_1 and k_{-1} playing a key role. For prokaryotes, the processes of DNA activation and de-activation are much slower than the transcription ($k_a \ll k_r$ and $k_d \ll k_r$): we chose the values $k_a = k_d = 0.001$ [10]. Our conclusions remain the same for the case wherein $k_a \gg k_r, k_d \gg k_r$ (not shown).

Using Monte Carlo simulation over $K = 10^5$ paths we computed sample mean approximations to $\mathbb{E}[P]$, $\mathbb{E}[P^2]$, $\mathbb{E}[P_2]$ and $\mathbb{E}[P_2^2]$. Initial conditions were determined from the steady-state of the ODE-based model: $D(0) = k_d D_T/(k_a + k_d)$, $D^*(0) = k_d D_T/(k_a + k_d)$, $M(0) = k_r D^*/\gamma_r$, $P(0) = (-b + \sqrt{d})/(2a)$, and $P_2(0) = k_1 P(0)^2/(k_{-1} + \gamma_{p_2})$. Here, $d = \gamma_p^2 + 4ak_p M(0)$ and $a = 2k_1(1 - k_{-1}/(k_{-1} + \gamma_{p_2}))$, and we consider the time interval $0 \le t \le 20$. In Table 1 we give approximate 95% confidence intervals for each sample mean, found by adding and subtracting $\pm 1.96\,\mathrm{std}/\sqrt{K}$, where std denotes the standard deviation. CLEa and CLEb denote the results for Euler–Maruyama [16] using a fixed stepsize of 0.04 and 0.004, respectively, in order to confirm that numerical discretization errors are not significant.

Table 1 shows that to typical Monte Carlo accuracy, the CLE does a good job of reproducing the means and variances of the CME. Along with the theoretical results in sections 4 and 6, this adds further support for the use of a stochastic differential equation model as a comprise between the extremes of the expensive discrete/stochastic CME and the crude continuous/deterministic RRE.

A Appendix: CME Moments for Gene Transcription System

For the system described by (16)–(19), we may write the state vector for the CME in the form

$$\begin{bmatrix} D_1 \\ D_1^\star \\ D_2 \\ D_2^\star \\ \vdots \\ \vdots \\ D_m \\ D_m^\star \\ M \\ P \end{bmatrix} \in \mathbb{R}^{2m+2}.$$

It then follows that

$$\mathcal{K} = \begin{bmatrix} -k_a & k_d & & & & & & & & \\ k_a & -k_d & & & & & & & & \\ & & -k_a & k_d & & & & & & \\ & & k_a & -k_d & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & -k_a & k_d & & \\ & & & & & & k_a & -k_d & & \\ 0 & k_r & 0 & k_r & \dots & & 0 & k_r & -\gamma_r & 0 \\ & & & & & & & & k_p & -\gamma_p \end{bmatrix} \in \mathbb{R}^{(2m+2)\times(2m+2)},$$

$K^{\mathrm{s}} = 0 \in \mathbb{R}^{(2m+2)\times(2m+2)}$, and $\Gamma \in \mathbb{R}^{(2m+2)\times(2m+2)}$ is such that

$$\Gamma+\Gamma^T = \begin{bmatrix} 0 & 0 & \dots & & \dots & 0 & 0 & 0 \\ 0 & \ddots & & & & \vdots & k_r\mathbb{E}[D_1^\star] & \vdots \\ \vdots & & & & & & 0 & \vdots \\ & & & & & & k_r\mathbb{E}[D_2^\star] & \\ & & & & & & \vdots & \\ & & & & & & \vdots & \\ \vdots & & & & \ddots & \vdots & 0 & \vdots \\ 0 & \ddots & & & & 0 & k_r\mathbb{E}[D_m^\star] & 0 \\ 0 & k_r\mathbb{E}[D_1^\star] & 0 & k_r\mathbb{E}[D_2^\star] & \dots \quad \dots & 0 & k_r\mathbb{E}[D_m^\star] & 0 & k_p\mathbb{E}[M] \\ 0 & 0 & \dots & \dots & & & 0 & k_p\mathbb{E}[M] & 0 \end{bmatrix}.$$

[12] Gillespie DT. The chemical Langevin equation. J Chem Phys. 2000;113:297–306.

[13] Khanin R, Higham DJ. Chemical Master Equation and Langevin regimes for a gene transcription model. In: Calder M, Gilmour S, editors. Computational Mathematics and Systems Biology. Lecture Notes in Bioinformatics 4695, Springer-Verlag; 2007. p. 1–14.

[14] Gillespie DT. Markov Processes: An Introduction for Physical Scientists. San Diego: Academic Press; 1991.

[15] Higham DJ. Modeling and simulating chemical reactions. SIAM Review. 2007;50:347–368.

[16] Mao X. Stochastic Differential Equations and Applications. Chichester: Horwood; 1997.

[17] Mao X, Yuan C. Stochastic Differential Equations with Markovian Switching. London: Imperial College Press; 2006.

[18] Buchler NE, Gerland U, Hwa T. Nonlinear protein degradation and the function of genetic circuits. Proc Nat Acad Sci USA. 2005;102:9559–64.

[19] Bundschuh R, Hayot F, Jayaprakash C. The role of dimerization in noise reduction of simple genetic networks. J Theor Biol. 2003;220:261–269.