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.

[*]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.

1 Motivation

A system of chemical reactions at thermal equilibrium is traditionally modelled as a collection of ordinary differential equations (ODEs), using the law of mass action. In this so-called reaction rate equation (RRE), each component represents the real-valued concentration of one species. By contrast, the Chemical Master Equation (CME) models the system as a discrete, integer-valued Markov process and consists of a (typically massive) ODE where each component represents the probability of one particular state. The pioneering work of Gillespie [1, 2] showed how the CME can be derived from first principle arguments and how the RRE arises from the CME in the thermodynamic limit. For the large number of molecules involved in typical chemical reactions, the RRE is a perfectly adequate model. However, there are some application areas, most notably in molecular biology [3, 4, 5, 6], where the numbers of molecules present is so small that the RRE model is inappropriate. In these cases, stochastic models are relevant and noise is typically summarized through means, correlations and variances [7, 8, 9, 10, 11].

From a computational viewpoint, it is possible to simulate from the CME in a pathwise sense using the Stochastic Simulation Algorithm (SSA), also known as Gillespie's algorithm [1, 2]. However, since this algorithm records every reaction, it is usually prohibitively expensive for even modest systems and time intervals. The Chemical Langevin Equation (CLE), introduced by Gillespie [12], can be regarded as a compromise between the CME and RRE. The CLE is a stochastic differential equation (SDE) where each component represents the level of one species as a time-dependent, real-valued, random variable. The CLE can be derived from the CME by making certain modelling assumptions and, in turn, the RRE can be derived from the CLE by further simplification. Simulating with the CLE is much more computationally attractive than with the CME, and hence it is important to understand what circumstances are needed for the CLE to be a good model. In this work, we contribute to this topic by analysing the first and second moments, and correlations, of the CLE for a generic class of first-order reaction networks that encompasses many models in cell biology. Using Ito's Lemma we derive ODEs for these quantities. Gadgil *et al.* [7] recently presented ODEs for the corresponding first and second moments and correlations of the CME. Our main finding is that these quantities match in the two regimes, despite the fact that one is inherently discrete and the other inherently continuous. This work was motivated by the analysis in [13], which showed the first and second moment matching result for a specific gene regulation network. We also note that Gillespie [14, Chapter 6] proved this equivalence for scalar problems and showed that third and higher moments do not match. Our contribution here is to extend this work to general first-order networks and to a hybrid model. In

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

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

We emphasize that in this setting $\mathbf{X}(t)$ is a non-negative integer-valued random variable, and the CME is a (typically massive) ODE showing how the probabilities evolve for each possible state.

In the CLE framework, we represent the amount of species i present at time t by a real-valued random variable $Y_i(t)$. The state vector $\mathbf{Y}(t) \in \mathbb{R}^N$ is then a real-valued stochastic process. The CLE, which describes the evolution of $\mathbf{Y}(t)$, takes the form of an Ito Stochastic differential equation [16]

$$d\mathbf{Y}(t) = \sum_{j=1}^{M} \boldsymbol{\nu}_j a_j(\mathbf{Y}(t))dt + \sum_{j=1}^{M} \boldsymbol{\nu}_j \sqrt{a_j(\mathbf{Y}(t))} \, dW_j(t). \tag{2}$$

Here, the $W_j(t)$ are independent scalar Brownian motions.

Throughout this work we assume that the SDE (2) has a well-defined, unique solution for which all first and second moments and correlation coefficients exist. Establishing conditions under which this property holds appears an open problem.

3 First-Order Networks

We now consider the four categories of reaction that Gadgil *et al.* [7] used to define a general class of first-order reaction networks.

Production From a Source, $\emptyset \xrightarrow{k_i^s} S_i$: Here, species i is created according to a rate constant k_i^s. This process does not depend on the current state of the system. The stoichiometric vector takes the form $\boldsymbol{\nu} = \mathbf{e}_i$, where $\mathbf{e}_i \in \mathbb{R}^N$ is the vector of zeros, except for a 1 in position i. The corresponding propensity function is k_i^s.

Degradation, $S_i \xrightarrow{k_i^d X_i} \emptyset$: Here, species i is degraded at a rate proportional to its current state, with an associated proportionality constant k_i^d. The stoichiometric vector takes the form $\boldsymbol{\nu} = -\mathbf{e}_i$, with propensity function $k_i^d X_i(t)$.

Conversion, $S_j \xrightarrow{k_{ij}^{con} X_j} S_i$: Here, species j is converted to species i at a rate proportional to the current state of S_j, with an associated proportionality constant k_{ij}^{con}. The stoichiometric vector takes the form $\boldsymbol{\nu} = -\mathbf{e}_j + \mathbf{e}_i$, with propensity function $k_{ij}^{con} X_j(t)$.

Catalytic Production From a Source, $\emptyset \xrightarrow{k_{ij}^{cat} X_j} S_i$: Here, species i is created at a rate proportional to the current state of S_j, with an associated propor-

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

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 k^{con}_{ii}, 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)]

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

and [7, equation (29)]

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

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

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

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

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

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

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

where $\mathbf{Y}(t) \in \mathbb{R}^N$, $\mathbf{b} : \mathbb{R}^N \to \mathbb{R}^N$, $\boldsymbol{\sigma} : \mathbb{R}^N \to \mathbb{R}^{N \times M}$ and $\mathbf{W}(t) \in \mathbb{R}^M$. There is one Brownian motion for each of the M reactions, and we order them in a manner that follows naturally from the description in section 3, so that

$$
\mathbf{W}(t) = \begin{bmatrix}
\mathbf{W}^{\mathrm{s}}(t) \\
\mathbf{W}^{\mathrm{d}}(t) \\
\mathbf{W}^{\mathrm{con}[1]}(t) \\
\mathbf{W}^{\mathrm{con}[2]}(t) \\
\vdots \\
\vdots \\
\mathbf{W}^{\mathrm{con}[N]}(t) \\
\mathbf{W}^{\mathrm{cat}[1]}(t) \\
\mathbf{W}^{\mathrm{cat}[2]}(t) \\
\vdots \\
\vdots \\
\mathbf{W}^{\mathrm{cat}[N]}(t)
\end{bmatrix}.
\tag{6}
$$

Here,

$\mathbf{W}^{\mathrm{s}}(t) \in \mathbb{R}^N$ has ith element $W_i^{\mathrm{s}}(t)$ corresponding to the source reaction $\emptyset \xrightarrow{\mathrm{k}_i^{\mathrm{s}}} S_i$,

$\mathbf{W}^{\mathrm{d}}(t) \in \mathbb{R}^N$ has ith element $W_i^{\mathrm{d}}(t)$ corresponding to the degradation reaction $S_i \xrightarrow{\mathrm{k}_i^{\mathrm{d}} X_i} \emptyset$,

$\mathbf{W}^{\mathrm{con}[j]}(t) \in \mathbb{R}^N$ has ith element $W_i^{\mathrm{con}[j]}(t)$ corresponding to the conversion reaction $S_j \xrightarrow{\mathrm{k}_{ij}^{\mathrm{con}} X_j} S_i$, and

$\mathbf{W}^{\mathrm{cat}[j]}(t) \in \mathbb{R}^N$ has ith element $W_i^{\mathrm{cat}[j]}(t)$ corresponding to the catalytic reaction $\emptyset \xrightarrow{\mathrm{k}_{ij}^{\mathrm{cat}} X_j} S_i$.

Taking account of the precise form of the propensity functions, we find that the CLE (2) has kth component

$$
\begin{aligned}
dY_k(t) &= \left\{ \mathrm{k}_k^{\mathrm{s}} - \mathrm{k}_k^{\mathrm{d}} Y_k(t) + \sum_{r=1}^N \mathrm{k}_{kr}^{\mathrm{con}} Y_r(t) - \sum_{r=1}^N \mathrm{k}_{rk}^{\mathrm{con}} Y_k(t) + \sum_{r=1}^N \mathrm{k}_{kr}^{\mathrm{cat}} Y_r(t) \right\} dt \\
&\quad + \sqrt{\mathrm{k}_k^{\mathrm{s}}}\, dW_k^{\mathrm{s}}(t) - \sqrt{\mathrm{k}_k^{\mathrm{d}} Y_k(t)}\, dW_k^{\mathrm{d}}(t) + \sum_{r=1}^N \sqrt{\mathrm{k}_{kr}^{\mathrm{con}} Y_r(t)}\, dW_k^{\mathrm{con}[r]}(t) \\
&\quad - \sum_{r=1}^N \sqrt{\mathrm{k}_{rk}^{\mathrm{con}} Y_k(t)}\, dW_r^{\mathrm{con}[k]}(t) + \sum_{r=1}^N \sqrt{\mathrm{k}_{kr}^{\mathrm{cat}} Y_r(t)}\, dW_k^{\mathrm{cat}[r]}(t).
\end{aligned}
\tag{7}
$$

It follows that the drift coefficient $\mathbf{b}\left(\mathbf{Y}(t)\right)$ in (5) has kth component given by
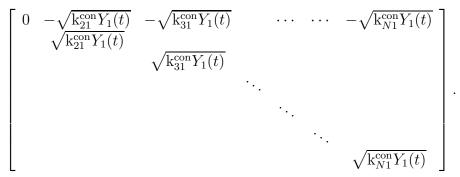
$$b_k\left(\mathbf{Y}(t)\right) = \mathrm{k}_k^{\mathrm{s}} - \mathrm{k}_k^{\mathrm{d}}Y_k(t) + \sum_{r=1}^{N}\mathrm{k}_{kr}^{\mathrm{con}}Y_r(t) - \sum_{r=1}^{N}\mathrm{k}_{rk}^{\mathrm{con}}Y_k(t) + \sum_{r=1}^{N}\mathrm{k}_{kr}^{\mathrm{cat}}Y_r(t). \quad (8)$$

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

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

and hence

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

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, $T^{\mathrm{s}}\left(T^{\mathrm{s}}\right)^T = \mathrm{diag}\left(\mathrm{k}_k^{\mathrm{s}}\right)$ 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

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

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

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

In words, $T^{\text{con}[r]}$ has all zeros in its rth column. All other columns have exactly two possible nonzeros; column j has $\sqrt{\text{k}_{jr}^{\text{con}} Y_r(t)}$ in position (j, j) and $-\sqrt{\text{k}_{jr}^{\text{con}} Y_r(t)}$ in position (j, r). It follows that $T^{\text{con}[r]} \left(T^{\text{con}[r]}\right)^T$ is given by

$$Y_r(t) \begin{bmatrix} \text{k}_{1r}^{\text{con}} & & & & -\text{k}_{1r}^{\text{con}} & & & \\ & \text{k}_{2r}^{\text{con}} & & & -\text{k}_{2r}^{\text{con}} & & & \\ & & \ddots & & \vdots & & & \\ & & & \text{k}_{r-1,r}^{\text{con}} & -\text{k}_{r-1,r}^{\text{con}} & & & \\ -\text{k}_{1r}^{\text{con}} & -\text{k}_{2r}^{\text{con}} & \cdots & -\text{k}_{r-1,r}^{\text{con}} & \sum_{s=1}^{N} \text{k}_{sr}^{\text{con}} & -\text{k}_{r+1,r}^{\text{con}} & \cdots & -\text{k}_{Nr}^{\text{con}} \\ & & & & -\text{k}_{r+1,r}^{\text{con}} & \text{k}_{r+1,r}^{\text{con}} & & \\ & & & & \vdots & & \ddots & \\ & & & & -\text{k}_{Nr}^{\text{con}} & & & \text{k}_{Nr}^{\text{con}} \end{bmatrix}.$$

For catalytic production, the general block $T^{\text{cat}[r]}$ arises from reactions where species S_r acts as a catalyst, and is diagonal, of the form $\text{diag}\left(\sqrt{\text{k}_{kr}^{\text{cat}} Y_r(t)}\right)$. Hence $T^{\text{cat}[r]} \left(T^{\text{cat}[r]}\right)^T = \text{diag}\left(\text{k}_{kr}^{\text{cat}} Y_r(t)\right)$.

Using these expressions in (9), it follows that $\mathbf{a}\left(\mathbf{Y}(t)\right) := \boldsymbol{\sigma}\left(\mathbf{Y}(t)\right) \boldsymbol{\sigma}^T\left(\mathbf{Y}(t)\right) \in \mathbb{R}^{N \times N}$ has diagonal element

$$a_{ii}\left(\mathbf{Y}(t)\right) = \text{k}_i^{\text{s}} + \text{k}_i^{\text{d}} Y_i(t) + \sum_{j=1}^{N} \left(\text{k}_{ij}^{\text{con}} Y_j(t) + \text{k}_{ji}^{\text{con}} Y_i(t)\right) + \sum_{j=1}^{N} \text{k}_{ij}^{\text{cat}} Y_j(t) \qquad (10)$$

and off-diagonal element

$$a_{ij}\left(\mathbf{Y}(t)\right) = -\text{k}_{ij}^{\text{con}} Y_j(t) - \text{k}_{ji}^{\text{con}} Y_i(t), \qquad \text{for } i \neq j. \qquad (11)$$

We are now in a position to examine the moments of $\mathbf{Y}(t)$.

Theorem 4.1. *Consider any first-order reaction network for which a unique solution to (2) exists along with the quantities $\mathbb{E}\left[Y_k(t)\right]$ and $\mathbb{E}\left[Y_k(t)Y_l(t)\right]$ for all $1 \leq k, l \leq N$. Then all first and second moments and correlations for the CLE match those for the CME, that is, $\mathbb{E}\left[X_k(t)\right] = \mathbb{E}\left[Y_k(t)\right]$ and $\mathbb{E}\left[X_k(t)X_l(t)\right] = \mathbb{E}\left[Y_k(t)Y_l(t)\right]$, for all $1 \leq k, l \leq N$.*

Proof
We see from (8) that the drift is linear, and hence taking expectations in (5) we obtain the ODE
$$\frac{d\mathbb{E}\left[\mathbf{Y}(t)\right]}{dt} = \mathbf{b}\left(\mathbb{E}\left[\mathbf{Y}(t)\right]\right).$$

Comparing (3) and (8) we see that $\mathbb{E}\left[\mathbf{X}(t)\right]$ and $\mathbb{E}\left[\mathbf{Y}(t)\right]$ 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 $U_{kl} = \mathbb{E}\left[X_k(t)X_l(t)\right]$ for $k \neq l$ and $U_{kk} = \mathbb{E}\left[X_k^2(t)\right]$, so that $\mathbf{U}(t) = \mathbf{V}(t) + \mathrm{diag}\left(\mathbf{M}(t)\right)$, it follows from (3) and (4) that

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

and

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

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

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

where "mart." denotes a martingale, whose precise form is not relevant to this analysis. Taking expected values, we conclude that

$$
\begin{aligned}
\frac{d\mathbb{E}\left[Y_k(t)Y_l(t)\right]}{dt} \;=\;& \mathrm{k}_k^{\mathrm{s}}\mathbb{E}\left[Y_l(t)\right] - \mathrm{k}_k^{\mathrm{d}}\mathbb{E}\left[Y_l(t)Y_k(t)\right] + \sum_{r=1}^{N}\mathrm{k}_{kr}^{\mathrm{con}}\mathbb{E}\left[Y_l(t)Y_r(t)\right] \\
& - \mathbb{E}\left[Y_l(t)Y_k(t)\right]\sum_{r=1}^{N}\mathrm{k}_{rk}^{\mathrm{con}} - \mathrm{k}_{kl}^{\mathrm{con}}Y_l(t) + \sum_{r=1}^{N}\mathrm{k}_{kr}^{\mathrm{cat}}\mathbb{E}\left[Y_l(t)Y_r(t)\right] \\
& + \mathrm{k}_l^{\mathrm{s}}\mathbb{E}\left[Y_k(t)\right] - \mathrm{k}_l^{\mathrm{d}}\mathbb{E}\left[Y_k(t)Y_l(t)\right] + \sum_{r=1}^{N}\mathrm{k}_{lr}^{\mathrm{con}}\mathbb{E}\left[Y_k(t)Y_r(t)\right] \\
& - \mathbb{E}\left[Y_k(t)Y_l(t)\right]\sum_{r=1}^{N}\mathrm{k}_{rl}^{\mathrm{con}} - \mathrm{k}_{lk}^{\mathrm{con}}Y_k(t) + \sum_{r=1}^{N}\mathrm{k}_{lr}^{\mathrm{cat}}\mathbb{E}\left[Y_k(t)Y_r(t)\right], \quad \text{for } k \neq l.
\end{aligned}
\tag{14}
$$

Similarly, applying Ito's lemma to $Y_k^2(t)$, using (8) and (10), we find that

$$
\begin{aligned}
d\left(Y_k^2(t)\right) \;=\;& \left\{2Y_k(t)b_k\left(\mathbf{Y}(t)\right) + a_{kk}\left(\mathbf{Y}(t)\right)\right\}dt + \mathrm{mart.}, \\
=\;& \left\{2\mathrm{k}_k^{\mathrm{s}}Y_k(t) - 2\mathrm{k}_k^{\mathrm{d}}Y_k(t)^2 + 2\sum_{r=1}^{N}\mathrm{k}_{kr}^{\mathrm{con}}Y_k(t)Y_r(t) - 2Y_k^2(t)\sum_{r=1}^{N}\mathrm{k}_{rk}^{\mathrm{con}} \right. \\
& + 2\sum_{r=1}^{N}\mathrm{k}_{kr}^{\mathrm{cat}}Y_k(t)Y_r(t) + \mathrm{k}_k^{\mathrm{s}} + \mathrm{k}_k^{\mathrm{d}}Y_k(t) \\
& \left. + \sum_{r=1}^{N}\left(\mathrm{k}_{kr}^{\mathrm{con}}Y_r(t) + \mathrm{k}_{rk}^{\mathrm{con}}Y_k(t)\right) + \sum_{r=1}^{N}\mathrm{k}_{kr}^{\mathrm{cat}}Y_r(t)\right\}dt + \mathrm{mart.}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\frac{d\mathbb{E}\left[Y_k^2(t)\right]}{dt} \;=\;& 2\mathrm{k}_k^{\mathrm{s}}\mathbb{E}\left[Y_k(t)\right] - 2\mathrm{k}_k^{\mathrm{d}}\mathbb{E}\left[Y_k^2(t)\right] + 2\sum_{r=1}^{N}\mathrm{k}_{kr}^{\mathrm{con}}\mathbb{E}\left[Y_k(t)Y_r(t)\right] - 2\mathbb{E}\left[Y_k^2(t)\right]\sum_{r=1}^{N}\mathrm{k}_{rk}^{\mathrm{con}} \\
& + 2\sum_{r=1}^{N}\mathrm{k}_{kr}^{\mathrm{cat}}\mathbb{E}\left[Y_k(t)Y_r(t)\right] + \mathrm{k}_k^{\mathrm{s}} + \mathrm{k}_k^{\mathrm{d}}\mathbb{E}\left[Y_k(t)\right] \\
& + \sum_{r=1}^{N}\mathrm{k}_{kr}^{\mathrm{con}}\mathbb{E}\left[Y_r(t)\right] + \mathbb{E}\left[Y_k(t)\right]\sum_{r=1}^{N}\mathrm{k}_{rk}^{\mathrm{con}} + \sum_{r=1}^{N}\mathrm{k}_{kr}^{\mathrm{cat}}\mathbb{E}\left[Y_r(t)\right].
\end{aligned}
\tag{15}
$$

Comparing (14) and (15) with (12) and (13), we see that the second moments and correlations for the CME and CLE statisfy the same closed system of ODEs, which proves the result.

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

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

and

$$
M \quad \xrightarrow{k_p} \quad M + P \tag{17}
$$

$$
M \quad \xrightarrow{\gamma_r} \quad \emptyset \tag{18}
$$

$$
P \quad \xrightarrow{\gamma_p} \quad \emptyset. \tag{19}
$$

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.

The motivation is that the mRNA and protein (case (a)) or protein (case (b)) may be much more abundant than the other quantities, so the continuous-valued ODE framework may be appropriate for that part of the system. This, of course, allows for a cheaper simulation than the full CME. Paszek looked at the steady-state means and variances and showed that moving partially to the RRE in this way introduces an error in the variance that can be quantified in terms of the reaction rate parameters. Our aim in this section is to consider partially using the CLE instead of the RRE—this also gives a much cheaper model to simulate but, using our insights from section 4, should give a more accurate variance. In fact, we will see that this version reproduces the exact CME variance for all time and for any value of m.

We also emphasize that "noise strength" in stochastic regulatory network models is typically measured as the ratio of mean to variance (often at steady state) [8, 9, 10, 11], so our focus on first and second moments is highly relevant.

We begin by stating the ODEs arising for the first and second moments and correlations in the CME version of (16)–(19). With a slight abuse of notation, we will let $D_i(t)$, $D_i^\star(t)$, $M(t)$ and $P(t)$ denote the random variables describing the state of the corresponding species at time t in the CME, and for convenience we will not explicitly display the time dependence. Using (3) and (4) (with some extra details given in the appendix) we then obtain

$$\frac{d\mathbb{E}[M]}{dt} = \sum_{s=1}^{m} k_r \mathbb{E}[D_s^\star] - \gamma_r \mathbb{E}[M] \tag{20}$$

$$\frac{d\mathbb{E}[P]}{dt} = k_p \mathbb{E}[M] - \gamma_p \mathbb{E}[P] \tag{21}$$

and

$$\frac{d\mathbb{E}[M^2]}{dt} = 2k_r \sum_{s=1}^{m} \mathbb{E}[MD_s^\star] - 2\gamma_r \mathbb{E}[M^2] + \gamma_r \mathbb{E}[M] + k_r \sum_{s=1}^{m} \mathbb{E}[D_s^\star] \tag{22}$$

$$\frac{d\mathbb{E}[P^2]}{dt} = 2k_p \mathbb{E}[MP] - 2\gamma_p \mathbb{E}[P^2] + \gamma_p \mathbb{E}[P] + k_p \mathbb{E}[M] \tag{23}$$

and, after summing the ODEs for $\mathbb{E}[D_s^\star M]$ and those for $\mathbb{E}[D_s^\star P]$,

$$\frac{d\sum_{s=1}^{m} \mathbb{E}[D_s^\star M]}{dt} = -(\gamma_r + k_d) \sum_{s=1}^{m} \mathbb{E}[D_s^\star M] + k_a \sum_{s=1}^{m} \mathbb{E}[D_s M] + k_r \sum_{k=1}^{m} \sum_{s=1}^{m} \mathbb{E}[D_k^\star D_s^\star] \tag{24}$$

$$\frac{d\sum_{s=1}^{m} \mathbb{E}[D_s^\star P]}{dt} = -(k_d + \gamma_p) \sum_{s=1}^{m} \mathbb{E}[D_s^\star P] + k_p \sum_{s=1}^{m} \mathbb{E}[D_s^\star M] + k_a \sum_{s=1}^{m} \mathbb{E}[D_s P] \tag{25}$$

and finally

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

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

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

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

probability of moving to state $i-1$ is proportional to i and the deactivation rate, giving

$$\gamma_{i,i-1} = ik_d. \tag{28}$$

Alternatively, the probability of moving to state $i+1$ is proportional to the number of inactive genes, $m-i$, and the activation rate, k_a, giving

$$\gamma_{i,i+1} = (m-i)k_a. \tag{29}$$

Since all other off-diagonal entries on the ith row are zero, we have

$$\gamma_{i,i} = -ik_d - (m-i)k_a. \tag{30}$$

In applying the generalized Ito's lemma we also need to know that

$$
\begin{aligned}
\sum_{j=1}^{m} j\gamma_{r(t),j} &= (r(t)-1)\,\gamma_{r(t),r(t)-1} + r(t)\gamma_{r(t),r(t)} + (r(t)+1)\,\gamma_{r(t),r(t)+1} \\
&= -k_d r(t) - k_a r(t) + k_a m. \tag{31}
\end{aligned}
$$

For the first moments, we simply take expectation in (27) to obtain

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

and

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

agreeing with (20) and (21).

Applying the generalized Ito lemma to $\widehat{M}r$, we find that

$$d\left(\widehat{M}r\right) = \left(r\left(k_r r - \gamma_r \widehat{M}\right) + \sum_{j=1}^{m} j\widehat{M}\gamma_{r(t),j}\right) dt + \text{mart.}$$

Using (31) and taking expectations,

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

which agrees with (24), when we note that $D_s + D_s^\star = 1$.

Similarly apply generalized Ito to $\widehat{P}r$,

$$d\left(\widehat{P}r\right) = \left(r\left(k_p \widehat{M} - \gamma_p \widehat{P}\right) + \sum_{j=1}^{m} j\widehat{P}\gamma_{r(t),j}\right) dt + \text{mart.},$$

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.

6   Dimerization Model

We saw in section 5 that the Raser and O'Shea gene transcription model (16)–(19) is a first-order reaction network in the sense of [7]. This first-order character is lost when we need to model the case where proteins combine to form complexes. Typically, these complexes degrade less rapidly than the underlying proteins, producing so-called "cooperative stability"[18]. This effect can widen the accessible range of protein levels and hence enhance biological function. Cooperative stability through nonlinear degradation was examined in [18] in a simple genetic circuit with feedback, although stochastic effects were not considered. An SSA study of the role of dimerization in noise reduction of a simple genetic circuit with negative feedback was conducted in [19], under the assumption that the dimers do not degrade but dissociate back to monomers.

The CME and CLE formulations of a simple dimerization model were compared in [13]. Here, we consider a more realistic case, which is therefore more computationally demanding. We take the gene transcription model (16)–(19) in the case of a single gene copy, $m = 1$, adding a dimerization pathway (the protein monomers P forming dimers P_2), with the dimers degrading less rapidly than the monomers ($\gamma_{p_2} < \gamma_p$):

$$P + P \quad \overset{k_1}{\to} \quad P_2 \tag{32}$$

$$P_2 \quad \overset{k_{-1}}{\to} \quad P + P \tag{33}$$

$$P_2 \quad \overset{\gamma_{p_2}}{\to} \quad \emptyset \tag{34}$$

Here, reaction (32) represents the dimerization process—two proteins combine to form a new species. In reaction (33) a dimer molecule dissociates back into two protein moelcules. Reaction (34) represents decay of a dimer molecule. Theorem 4.1 and the results in [7] do not apply, because $P + P \to P_2$ is not first-order, so our aim is to test via numerical simulation how closely the first and second moments for the CME match those for the CLE.

Writing the state vector as

$$
\begin{bmatrix}
X_1(t) \\
X_2(t) \\
X_3(t) \\
X_4(t) \\
X_5(t)
\end{bmatrix}
=:
\begin{bmatrix}
D \\
D^* \\
M \\
P \\
P_2
\end{bmatrix},
$$

an appropriate stoichiometric matrix takes the form

$$\begin{bmatrix} \boldsymbol{\nu}_1 & \boldsymbol{\nu}_2 & \ldots & \boldsymbol{\nu}_9 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & -2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 \end{bmatrix}$$

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 $\pm 1.96 \, std/\sqrt{K}$, 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.

	$\mathbb{E}[P]$	$\mathbb{E}[P^2]$	$\mathbb{E}[P_2]$	$\mathbb{E}[P_2^2]$
CME	$[64.96, 65.06]$	$[4276.9, 4289.1]$	$[105.06, 105.11]$	$[11054, 11064]$
CLEa	$[64.99, 65.08]$	$[4280.2, 4292.5]$	$[105.05, 105.10]$	$[11051, 11061]$
CLEb	$[64.98, 65.07]$	$[4279.4, 4291.7]$	$[105.06, 105.11]$	$[11053, 11063]$

Table 1: 95% confidence intervals for Monte Carlo sample mean approximations to $\mathbb{E}[P]$, $\mathbb{E}[P^2]$, $\mathbb{E}[P_2]$ and $\mathbb{E}[P_2^2]$ at time $t = 20$ in the CME and CLE formulations of the model (16)–(19), with $m = 1$, and (32)–(34). CLEa uses Euler–Maruyama with stepsize 0.04 and CLEb uses Euler–Maruyama with stepsize 0.004.

7   Summary and Conclusions

Means, variances and correlations are the standard measures that biologists use to summarize the intrinsic noise in regulatory networks. In the commonly arising case where reactions are first-order, we have shown that a stochastic differential equation model known as the Chemical Langevin Equation gives exactly the same results, for all time, as the more fundamental Chemical Master Equation. We also showed that a particular multi-scale model that combines elements of the Langevin and Master equations has the same property. From the viewpoint of computational biology, this result supports the use of the numerically discretized Chemical Langevin Equation as the basis of a simulation tool that is typically orders of magnitude more efficient than the Stochastic Simulation Algorithm approach, which uses the Chemical Master Equation. Further, just as deterministic ordinary differential equations are generally more amenable to theoretical analysis than discrete maps, the Langevin regime, which can be analyzed conveniently with the tools of stochastic calculus, will typically succumb more readily to analysis than the discrete-valued master equation. There are, of course, many interesting open questions regarding the extent to which the Langevin and Master Equation approaches match for higher order networks.

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

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

Using these expressions in (3) and (4) we obtain the ODEs (21), (22) and (23).

References

[1] Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comp Phys. 1976;22:403—434.

[2] Gillespie DT. Exact stochastic simulation of coupled chemical reactions. J Phys Chem. 1977;81:2340–2361.

[3] de Jong H. Modeling and simulation of genetic regulatory systems: A literature review. Journal of Computational Biology. 2002;9:69–105.

[4] Samad HE, Khammash M, Petzold L, Gillespie DT. Stochastic modeling of gene regulatory networks. Int J Robust and Nonlinear Control. 2005;15:691–711.

[5] Turner TE, Schnell S, Burrage K. Stochastic approaches for modelling in vivo reactions. Computational Biology and Chemistry. 2004;28:165–178.

[6] Wilkinson DJ. Stochastic Modelling for Systems Biology. Chapman & Hall/CRC; 2006.

[7] Gadgil C, Lee CH, Othmer HG. A stochastic analysis of first-order reaction networks. Bulletin of Mathematical Biology. 2005;67:901–946.

[8] Paszek P. Modeling stochasticity in gene regulation: characterization in the terms of the underlying distribution function. Bulletin of Mathematical Biology. 2007;69:1567–1601.

[9] Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. Stochastic mRNA synthesis in mammalian Cells. PLoS Biology. 2006;4:e309.

[10] Raser JM, O'Shea EK. Control of stochasticity in eukaryotic gene expression. Science. 2004;304(5678):1811–4.

[11] Thattai M, van Oudenaarden A. Intrinsic noise in gene regulatory networks. Proc Natl Acad Sci USA. 2001;98(15):8614–19.

[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.