Path integral approach to generating functions for multistep post-transcription and post-translation processes and arbitrary initial conditions

Jaroslav Albert

Received: 30 October 2018 / Revised: 5 August 2019 / Published online: 5 September 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract
Stochastic fluctuations in the copy number of gene products have perceivable effects on the functioning of gene regulatory networks (GRNs). The master equation (ME) provides a theoretical basis for studying such effects. However, solving the ME can be a task that ranges from simple to difficult to impossible using conventional methods. Therefore, discovering new techniques for solving the ME is an important part of research on stochastic GRNs. In this paper, we present a novel approach to obtaining the generating function (GF), which contains the same information as the ME, for a one gene system that includes multi-step post-transcription and post-translation processes. The novelty of the approach lies in solving the GF for proteins in terms of a particular path taken by the partially and fully processed mRNAs in the time-copy number plane, after which the GF is summed over all possible paths. We prove a theorem that shows the summation of all paths to be equivalent to an equation similar to the ME for the mRNAs. On a system with six gene products in total and randomly selected initial conditions, we confirm the validity of our results by comparing them with Gillespie simulations.

Keywords Stochastic gene expression · Master equation · Generating function · Path integral

Mathematics Subject Classification 92B05 · 92-08

1 Introduction
Stochastic fluctuations give rise to cell-to-cell differences in copy numbers of gene products such as mRNA and protein. Sometimes these differences are insignificant;
other times they lead to major shifts in phenotype (Blake et al. 2003). Therefore, understanding the impact of stochastic fluctuations is an important endeavor in the field of systems biology.

There exist numerous methods for modeling stochastic gene expression. Some of them are entirely numerical, such as the Gillespie algorithm (GA) (Gillespie 1977) and its derivatives (Gibson and Bruck 2000; Gillespie 2001; Cao et al. 2004, 2005a, b); others are hybrids of the GA and the master equation (ME) (Burrage et al. 2004; Jahnke and Altintan 2010; Albert 2016a, b; Duso and Zechner 2018; Alfonso et al. 2005; Kurasov et al. 2018); while the rest facilitate either exact or approximate analytic solutions to the ME (Jahnke and Huisinga 2007; Albert and Rooman 2016). However, solving the ME analytically has proven possible only for systems that are too simple and hence not very interesting. For more complex systems, the strategy is usually to find techniques that lead to approximate (but analytic) solutions of either the ME or the generating function (GF) (Shahrezaei and Swain 2008; Pendar et al. 2013; Bokes et al. 2012a, b; Popović et al. 2016; Veerman et al. 2018).

In this paper we present a novel approach to obtaining the generating function of a one-gene system comprising of mRNA and protein at different stages of post-transcription (for mRNA) and post-translation (for protein) processing. We make the observation that the master equation can be reduced to include the mRNAs only, and that a similar reduction can be achieved also for the proteins but only for a specific path in the time-copy number plane of the fully processed mRNAs. We show, by proving a theorem involving a sum of all paths of the partially and fully processed mRNA, that obtaining the generating function for arbitrary initial conditions can be reduced to solving a set of $M$ ordinary differential equations, where $M - 1$ is the number of post-transcription processes. Solving these equations numerically, and with the help of Cauchy’s contour theorem, we compute the probability distributions for the fully processed proteins. Also, we compute the first four moments for the fully processed protein starting with a randomly generated set of initial values for all the gene products. We demonstrate the validity of our approach by comparing our results with Gillespie simulations.

2 The master equation

The system we consider comprises of these reactions:

$\emptyset \xrightarrow{r} m_1$

$m_M \xrightarrow{d} \emptyset$

$m_i \xrightarrow{a_i} m_{i+1} \quad i = 1, \ldots, M - 1$

$m_i \xrightarrow{\bar{a}_i} m_{i-1} \quad i = 2, \ldots, M$

$m_M \xrightarrow{K} m_M + n_1$

$n_N \xrightarrow{q} \emptyset$
where \( m_1 \) is the copy number of freshly transcribed mRNAs, \( m_i \), for \( i = 2, 3, \ldots, M \), are the copy numbers of mRNAs that have undergone the first (\( i = 2 \)), second (\( i = 3 \)), etc. post-transcription process with \( m_M \) being the copy number of fully processed mRNA from which proteins can be translated. The same notation applies to the proteins: \( n_1 \) is the copy number of freshly translated proteins, and \( n_N \) is the copy number of fully processed proteins. For the reactions that change the variables \( m_i \), \( r \) is the transcription rate, \( d \) is the degradation rate of a fully processed mRNA, and \( a_i \) and \( \bar{a}_i \) are the forward and backward reaction rates of the post-transcription processes, respectively. For the remaining reactions, \( K \) is the translation rate, \( q \) is the degradation rate of a fully processed protein, and \( b_i \) and \( \bar{b}_i \) are the forward and backward reaction rates of the post-translation processes, respectively. The master equation for this system reads

\[
\frac{\partial}{\partial t} P(m, n, t) = r(t)[P(m_1 - 1) - P] + d((m_M + 1)P(m_M + 1) - m_M P]
\]

\[
+ \sum_{i=1}^{M-1} a_i [(m_i + 1)P(m_i + 1, m_{i+1} - 1) - m_i P]
\]

\[
+ \sum_{i=1}^{M-1} \bar{a}_{i+1} [(m_{i+1} + 1)P(m_i - 1, m_{i+1} + 1) - m_{i+1} P]
\]

\[
+ K m_M [P(n_1 - 1) - P] + q [(n_N + 1)P(n_N + 1) - n_N P]
\]

\[
+ \sum_{i=1}^{N-1} b_i [(n_i + 1)P(n_i + 1, n_{i+1} - 1) - n_i P]
\]

\[
+ \sum_{i=1}^{N-1} \bar{b}_{i+1} [(n_{i+1} + 1)P(n_i - 1, n_{i+1} + 1) - n_{i+1} P],
\]

where \( P(m, n, t) \) is the joint probability of observing the sets of copy numbers \( m = (m_1, \ldots, m_M) \) and \( n = (n_1, \ldots, n_N) \). For brevity, we only write the argument(s) of \( P \) on the right hand side of Eq. (2) explicitly if there is a change to the said argument(s); e. g. instead of writing \( P(m, n_1, \ldots, n_k + 1, \ldots, n_N, t) \), we write \( P(n_k + 1) \). If there is no change to any argument, we merely write \( P \). In principle, Eq. (2) could be solved numerically; however, even for moderate average mRNA and protein copy numbers, the dimension of the problem might be too large for such a direct approach. For example, for \( M = N = 3 \), \( \langle m_1 \rangle = \langle m_2 \rangle = \langle m_3 \rangle \sim 10 \) and \( \langle n_1 \rangle = \langle n_2 \rangle = \langle n_3 \rangle \sim 100 \), the number of equations that need to be solved are of order \( 10^3 \times 100^3 = 10^9 \).
3 The generating function

3.1 Direct approach

An alternative approach to the above problem is to obtain a generating function (GF), which contains as much information about the system as the ME. If we let $s$ be the set of all variables $(m_1, ..., m_M, n_1, ..., n_N)$, then the GF is defined as (Van Kampen 2007)

$$F(y,t) = \sum_s \prod_{k=1}^{M+N} y_{sk}^s P(s,t), \quad (3)$$

where

$$\sum_s \equiv \sum_{s_1}^{\infty} ... \sum_{s_{M+N}}^{\infty}. \quad (4)$$

The probability distribution, $P(s_k, t)$, for the variable $s_k$ is related to $F(s,t)$ through this expression:

$$P(s_k, t) = \frac{1}{s_k!} \left[ \frac{\partial^s_k F(y,t)}{\partial y_{sk}} \right]_{y=0}. \quad (5)$$

The GF can also be used to obtain statistical moments:

$$\langle s_k^l \rangle = \left[ \left( y_k \frac{\partial}{\partial y_k} \right)^l F(y,t) \right]_{y_1=1, ..., y_{M+N}=1} = \sum_s s_k^l P(s,t). \quad (6)$$

To obtain an equation for the GF, we must multiply Eq. (2) by the product $\prod_{i=1}^{M+N} y_{si}^s$ and sum over $s$. The result is a partial differential equation of the form (Walczak et al. 2012)

$$\frac{\partial F}{\partial t} = r(t)(y_1 - 1)F - d(y_M - 1) \frac{\partial F}{\partial y_M}$$

$$+ \sum_{i=1}^{M-1} a_i (y_{i+1} - y_i) \frac{\partial F}{\partial y_i} + \sum_{i=1}^{M-1} \tilde{a}_{i+1} (y_i - y_{i+1}) \frac{\partial F}{\partial y_{i+1}}$$

$$+ K(y_{M+1} - 1) y_M \frac{\partial F}{\partial y_M} - q(y_{M+N} - 1) \frac{\partial F}{\partial y_{M+N}}$$

$$+ \sum_{i=1}^{M-1} b_i (y_{M+i+1} - y_{M+i}) \frac{\partial F}{\partial y_{M+i}} + \sum_{i=1}^{M-1} \tilde{b}_{i+1} (y_{M+i} - y_{M+i+1}) \frac{\partial F}{\partial y_{M+i+1}}. \quad (7)$$
For the system at hand, even this equation is very difficult to solve, especially for arbitrary initial conditions. In the next section, we show an alternative approach for obtaining the GF.

### 3.2 Path integral approach

We begin by noticing that the stochastic evolution of the set \( m = (m_1, ..., m_M) \) is independent of the set \( n = (n_1, ..., n_N) \). To see this, we sum both sides of Eq. (2) over \( n \) to obtain

\[
\frac{\partial}{\partial t} P(m, t) = r(t)[P(m_1 - 1) - P] + d[(m_M + 1)P(m_M + 1) - m_M P] + \sum_{i=1}^{M-1} a_i[(m_i + 1)P(m_i + 1, m_{i+1} - 1) - m_i P] + \sum_{i=1}^{M-1} \bar{a}_{i+1}[(m_{i+1} + 1)P(m_i - 1, m_{i+1} + 1) - m_{i+1} P].
\] (8)

If the initial probability, \( P(m, 0) \), is Poisson, then \( P(m, t) \) is also Poisson:

\[
P(m, t) = \prod_{i=1}^{M} \frac{\gamma_i^{m_i}}{m_i!} e^{-\gamma_i},
\] (9)

where the vector \( \gamma \) satisfies

\[
\frac{d\gamma}{dt} = S\gamma + r,
\] (10)

in which

\[
S = \begin{bmatrix}
-a_1 & \bar{a}_2 & 0 & 0 & \cdots & 0 \\
\bar{a}_1 & -(a_2 + \bar{a}_2) & \bar{a}_3 & 0 & \cdots & 0 \\
0 & \bar{a}_1 & -(a_2 + \bar{a}_3) & \bar{a}_4 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & a_{M-2} - (a_{M-1} + \bar{a}_{M-1}) & \bar{a}_M & \cdots \\
0 & 0 & \cdots & a_{M-1} - (\bar{a}_{M-1} + d) & -a_M + d
\end{bmatrix}
\]
and

\[
\mathbf{r} = \begin{bmatrix} r(t) \\ 0 \\ \cdots \\ 0 \end{bmatrix}.
\]

The solution to Eq. (10) is

\[
\gamma_k = \Gamma_{k}^{(1)}(t) + \sum_{i=1}^{M} \sum_{j=1}^{M} \Gamma_{kj}^{(0)}(t)\gamma_j(0),
\]  

(11)

where

\[
\Gamma_{k}^{(1)}(t) = \int_{0}^{t} f_k(t-t')r(t')dt', \quad \Gamma_{kj}^{(0)}(t) = U_{ki}U_{ij}^{-1}e^{S_i t},
\]  

(12)

\[
f_k(t-t') = \sum_{i=1}^{M} U_{ki}U_{ij}^{-1}e^{S_i (t-t')},
\]  

(13)

\(S_i\) is the \(i\)th eigenvalue of \(S\), and \(U_{ij}\) is the unitary matrix that diagonalizes \(S\), i.e., \(U^{-1}SU = \delta_{ij}S_j\). The generating function for the variable \(m_k\) has the simple form

\[
F_k(x, t) = \prod_{j=1}^{M} \sum_{m} x^{m_k} \gamma_j^{m_j} \frac{m_j!}{m_j!} e^{-\gamma_j} = \exp \left[ (x-1) \gamma_k(t) \right].
\]  

(14)

For an initial probability distribution \(\delta_{m_1, \tilde{m}_1} \delta_{m_2, \tilde{m}_2} \cdots \delta_{m_M, \tilde{m}_M}\), where \(\tilde{m}_1, ..., \tilde{m}_M\) are the initial mRNA copy numbers, the GF takes on the following form (see “Appendix A”):

\[
F_k(\tilde{m}, x, t) = \prod_{j=1}^{M} \left[ (x-1) \Gamma_{kj}^{(0)}(t) + 1 \right]^{\tilde{m}_j} \exp \left[ (x-1) \Gamma_{k}^{(1)}(t) \right].
\]  

(15)

For an arbitrary initial probability distribution, \(P(\tilde{m}, 0)\), the GF becomes

\[
F_k(x, t) = \sum_{\tilde{m}} P(\tilde{m}, 0) \prod_{j=1}^{M} \left[ (x-1) \Gamma_{kj}^{(0)}(t) + 1 \right]^{\tilde{m}_j} \exp \left[ (x-1) \Gamma_{k}^{(1)}(t) \right],
\]  

(16)

or, in terms of the initial GF, \(F(x_1, ..., x_M, 0)\),

\(\square\) Springer
\[ F_k(x, t) = F(\phi_1(x, t), ..., \phi_M(x, t), 0) \exp \left[ (x - 1) I_k^{(1)}(t) \right], \quad (17) \]

where \( \phi_j(x, t) = (x - 1) I_{k_j}^{(0)}(t) + 1 \).

Unfortunately, variables \( n_1, ..., n_N \) cannot be decoupled from \( m \) via the same trick because the translation of proteins depends on mRNA via the term \( K_{m1} \). However, they can be decoupled in a different way.

Imagine we are able to observe the evolution, or path, of the variable \( m_M \) in real time. Then the ME for \( n \) is simply

\[
\frac{\partial}{\partial t} P(n, t) = K_{mM}(t)[P(n_1 - 1) - P] + q[(n_N + 1)P(n_N + 1) - n_N P] \\
+ \sum_{i=1}^{N-1} b_i[(n_i + 1)P(n_i + 1, n_{i+1} - 1) - n_i P] \\
+ \sum_{i=1}^{N-1} \tilde{b}_{i+1}[(n_{i+1} + 1)P(n_i - 1, n_{i+1} + 1) - n_{i+1} P], \quad (18)
\]

where \( m_M(t) \) is the observed path. Note that Eq. (18) is identical in structure to Eq. (8). Hence, starting with some initial set of protein copy numbers \((\tilde{n}_1, ..., \tilde{n}_N)\), the GF for the variable \( n_k \), given the path \( m_M(t) \), is

\[
F_k(\tilde{n}, y, t|m_M(t)) = \prod_{j=1}^{N} \left[ (y - 1) \Lambda_{k_j}^{(0)}(t) + 1 \right]^{\tilde{n}_j} \exp \left[ (y - 1) \Lambda_k^{(1)}(t) \right]. \quad (19)
\]

The functions \( \Lambda_{k_j}^{(0)}(t) \) and \( \Lambda_k^{(1)}(t) \) are defined as

\[
\Lambda_k^{(1)}(t) = \int_0^t g_k(t - t')K_{mM}(t')dt', \quad \Lambda_{k_j}^{(0)}(t) = \sum_{i=1}^{N} V_{k'i} V_{ij}^{-1} e^{T_i t}, \quad (20)
\]

where

\[
g_k(t - t') = \sum_{i=1}^{M} V_{ki} V_{lj}^{-1} e^{T_i (t-t')}, \quad (21)
\]

\( V_{ij} \) is the unitary matrix that diagonalizes the matrix.
\[ T = \begin{bmatrix}
-b_1 & \tilde{b}_2 & \cdot & \cdot & \cdot \\
 b_1 & -(b_2 + \tilde{b}_2) & \tilde{b}_3 & \cdot & \cdot \\
 & b_{N-2} & -(b_{N-1} + \tilde{b}_{N-1}) & \tilde{b}_N & \cdot \\
 & & b_{N-1} & -(b_N + q) & \cdot \\
\end{bmatrix}, \]

and \( T_i \) is the \( i \)th eigenvalue of \( T \). The GF (19) is valid only for a particular path taken by the variable \( m_M \). To obtain the true GF, we must multiply Eq. (19) by the probability of observing a particular set of paths, \( \{ m \} \), and then sum over all possible paths:

\[
F_k(\tilde{n}, y, t) = \sum_{\text{all paths}} P(\{ m(0) \}) P(\{ m \}) F_k(\tilde{n}, y, t|m_M(t)) = \prod_{j=1}^{N} \left[ (y - 1) \Lambda_{kj}^{(0)}(t) + 1 \right]^{n_j} Q(t), \tag{22}
\]

where

\[
\sum_{\text{all paths}} \equiv \prod_{j=1}^{M} \prod_{l=0}^{L} \sum_{m_j(t_l)}, \tag{23}
\]

and

\[
Q(t) = \sum_{\text{all paths}} P(\{ m(0) \}) P(\{ m \}) \exp \left[ (y - 1) \int_{0}^{t'} g_k(t - t') m_M(t') dt' \right]; \tag{24}
\]

\( \mathcal{P}(\{ m \}) \) is the joint probability to observe a particular set of paths, \( \{ m \} = \{ m(0), \ldots, m(t_L) \} \), where \( m(0) \) is the set \( m \) at \( t_0 = 0 \), \( m(t_1) \) is the set \( m \) at \( t_1 = \Delta t \), and so on until \( t_L = L \Delta t = t \); and \( P(\{ m(t_0) \}) \) is the probability of observing the set \( m \) at \( t = 0 \). In the limit \( \Delta t \to 0 \) and \( L \to \infty \), such that \( L \Delta t = t \), \( \{ m \} \) becomes a continuous path. For an example of a path, see Fig. 1.

We can work out Eq. (24) using the following theorem.

**Theorem 1** If \( P(\{ m, 0 \}) \) is the probability to observe \( m \) at \( t = 0 \) in a system governed by Eq. (8), then, for an arbitrary function \( W(\{ m \}, t, t') \),

\[
\sum_{\text{all paths}} P(\{ m(0) \}) \mathcal{P}(\{ m \}) \exp \left[ \int_{0}^{t'} W(\{ m \}, t, t') dt' \right] = \sum_{m} Q(m, t, t') \bigg|_{t'=t} \tag{25}
\]
where $Q(m, t, t')$ is the solution of

$$
\frac{dQ(m, t, t')}{dt'} = r(t')(Q(m_1 - 1) - Q) + d[(m_M + 1)Q(m_M + 1) - m_M Q] \\
+ \sum_{i=1}^{M-1} b_i[(m_i + 1)Q(m_i + 1, m_{i+1} - 1) - m_i Q] \\
+ \sum_{i=1}^{M-1} \tilde{b}_{i+1}[(m_{i+1} + 1)Q(m_i - 1, m_{i+1} + 1) - m_{i+1} Q] \\
+ W(m, t, t') Q
$$

such that $Q(m, t, 0) = P(m, 0)$. (For proof, see “Appendix B”.)

To evaluate Eq. (24), we need only to replace $W(m, t, t')$ in Eq. (26) with $(y-1)g_k(t-t')m_M$ and solve for $Q(m, t, t')$. If the initial probability distribution for mRNA is Poisson,

$$
P(m, 0) = \prod_{i=1}^{M} \frac{\eta^m_i}{m_i!} e^{-\eta_i}, \quad (27)
$$

then $Q(m, t, t')$ has the form

$$
Q(m, t, t') = \prod_{i=1}^{M} \frac{[\xi_i(t, t')]^{m_i}}{m_i!} e^{-h(t, t')}.
$$

Plugging Eq. (28) into Eq. (26) for $W(m, t, t') \to (y-1)g_k(t-t')m_M$ and collecting the factors of $m_i^1$ and $m_i^0$, we obtain the equations for $\xi$ and $h$:

$$
\frac{d\xi}{dt'} = r + S\xi + (y-1)g_k(t-t')B\xi
$$

(29)
\[
\frac{dh}{dt'} = r - d\xi_M, \tag{30}
\]

where \( B = \delta_i,M\delta_j,M \). The dependence of \( \xi \) and \( h \) on the index \( k \) was left out for the sake of simplicity. To satisfy the initial conditions \( Q(m, t, 0) = P(m, 0) \), we must have \( \xi_i(t, t' = 0) = \eta_i \) and \( h(t, t' = 0) = \sum_i \eta_i \). Summing Eq. (28) over \( m \) leads to

\[
Q(t) = \exp \left[ \sum_{i=1}^M \xi_i(t, t) - h(t, t) \right]. \tag{31}
\]

We can reduce this expression by adding up all equations in (29) and then subtracting Eq. (30):

\[
\frac{d}{dt'} \left[ \sum_{i=1}^M \xi_i(t, t') - h(t, t') \right] = (y - 1)g_k(t - t')\xi_M(t, t'). \tag{32}
\]

Integrating Eq. (32) over \( dt' \), we obtain

\[
Q(t) = \exp \left[ (y - 1) \int_0^t g_k(t - t')\xi_M(t, t')dt' \right]. \tag{33}
\]

The formal solution to Eq. (29) is

\[
\xi(t, t') = D^{(1)}(t, t') + D^{(0)}(t, t', 0)\xi(0), \tag{34}
\]

where

\[
D^{(1)}(t, t') = \int_0^{t'} D^{(0)}(t, t', t'')r(t'')dt'' \tag{35}
\]

is a vector and

\[
D^{(0)}(t, t', t'') = T \exp \left[ \int_{t''}^{t'} [S + (y - 1)g_k(t - t')B] dt_1 \right] \tag{36}
\]

is a matrix. \( T \) here is the time ordering operator. Expressions (36) and (35) are difficult to evaluate in their current form. However, by inserting Eq. (34) into Eq. (29), we obtain two separate equations, one for \( D^{(1)}(t, t') \), and another for \( D^{(0)}(t, t', 0) \):

\[
\frac{d}{dt'} D^{(1)}(t, t') = r + SD^{(1)}(t, t') + (y - 1)g_k(t - t')BD^{(1)}(t, t') \tag{37}
\]

\[
\frac{d}{dt'} D^{(0)}(t, t', 0) = SD^{(0)}(t, t', 0) + (y - 1)g_k(t - t')BD^{(0)}(t, t', 0), \tag{38}
\]
with initial conditions that follow from Eqs. (35) and (36): $D^{(1)}(t, 0) = 0$ and $D^{(0)}(t, 0, 0) = 1$. Eq. (33) can now be written as

$$Q(t) = \exp \left\{ (y - 1) \left[ \psi_k^{(1)}(t) + \sum_{j=1}^{M} \psi_{kj}^{(0)}(t) \xi_j(0) \right] \right\}, \tag{39}$$

where

$$\psi_k^{(1)}(t) = \int_0^t g_k(t - t') D_M^{(1)}(t', t') dt' \tag{40}$$

and

$$\psi_{kj}^{(0)}(t) = \int_0^t g_k(t - t') D_{Mj}^{(0)}(t', t', 0) dt'. \tag{41}$$

Note that in Eqs. (39), (40) and (41) we have put back the index $k$. Results in Eqs. (39), (40) and (41) are correct only if the initial probability distribution, $P(m, 0)$, is Poisson. For the initial probability distribution $\delta_{m_1, \tilde{m}_1} \delta_{m_2, \tilde{m}_2} \ldots \delta_{m_M, \tilde{m}_M}$, where, as before, $\tilde{m}_1, \ldots, \tilde{m}_M$ are the initial mRNA copy numbers, Eq. (39) becomes (see “Appendix C”)

$$Q(t) = \prod_{j=1}^{M} \left[ (y - 1) \psi_{kj}^{(0)}(t) + 1 \right]^{\tilde{m}_j} \exp \left[ (y - 1) \psi_k^{(1)}(t) \right]. \tag{42}$$

Hence, the GF in Eq. (22) reads

$$F_k(\tilde{n}, y, t) = \exp \left[ (y - 1) \psi_k^{(1)}(t) \right] \times \prod_{i=1}^{M} \prod_{j=1}^{N} \left[ (y - 1) \psi_{ki}^{(0)}(t) + 1 \right]^{\tilde{m}_i} \left[ (y - 1) \Lambda_{kj}^{(0)}(t) + 1 \right]^{\tilde{n}_j}. \tag{43}$$

Finally, for an arbitrary initial joint probability distribution, $P(m, n, 0)$, we obtain

$$F_k(y, t) = \exp \left[ (y - 1) \psi_k^{(1)}(t) \right] \times \sum_{m} \sum_{n} P(m, n, 0) \prod_{j=1}^{N} \left[ (y - 1) \psi_{kj}^{(0)}(t) + 1 \right]^{m_j} \left[ (y - 1) \Lambda_{kj}^{(0)}(t) + 1 \right]^{n_j} \exp \left[ (y - 1) \psi_k^{(1)}(t) \right] F(\mu_1(t), \ldots, \mu_M(t), \nu_1(t), \ldots, \nu_N(t), 0), \tag{44}$$
where

\[
F(x_1, \ldots x_M, y_1, \ldots, y_N, 0) = \sum_m \sum_n P(m, n, 0) \prod_{i=1}^M x_i^m_i \prod_{j=1}^N y_j^m_j
\]

(45)
is the initial GF, \( \mu_i(t) = (y - 1)\psi_{kj}^{(0)}(t) + 1 \) and \( \nu_i(t) = (y - 1)\Lambda_{kj}^{(0)}(t) + 1 \).

4 Numerical experiments

In this section we compute central moments and probability distribution functions (PDFs) using result (44) and compare them to Gillespie simulations.

4.1 Central moments

To compute statistical moments, we take successive derivatives of Eqs. (37) and (38) with respect to \( y \), and set \( y = 1 \). This leads to the following sets of equations:

\[
\frac{d}{dt'} D^{(1)}(t, t') = r + S D^{(1)}(t, t')
\]

\[
\frac{d}{dt'} D^{(0)}(t, t', 0) = S D^{(0)}(t, t', 0),
\]

(46)

for the zeroth derivative, and

\[
\frac{d}{dt'} \left[ \partial^p D^{(1)}(t, t') \right] = S \left[ \partial^p D^{(1)}(t, t') \right] + g_k(t - t')B \left[ \partial^{p-1} D^{(1)}(t, t') \right]
\]

(47)

\[
\frac{d}{dt'} \left[ \partial^p D^{(0)}(t, t', 0) \right] = S \left[ \partial^p D^{(0)}(t, t', 0) \right] + g_k(t - t')B \left[ \partial^{p-1} D^{(0)}(t, t', 0) \right]
\]

(48)

for the \( p \)th derivative, where \( \partial^p \equiv \partial^p / (\partial y)^p \). The solution to Eqs. (37) and (38) can now be written as a Taylor series in terms of \((y - 1)\):

\[
D^{(1)}(t, t') = \sum_{p=0}^\infty \left[ \partial^p D^{(1)}(t, t') \right] (y - 1)^p
\]

\[
D^{(0)}(t, t', 0) = \sum_{p=0}^\infty \left[ \partial^p D^{(0)}(t, t', 0) \right] (y - 1)^p,
\]

(49)

This makes the functions \( \psi_k^{(1)}(t) \) and \( \psi_{kj}^{(0)}(t) \) polynomials on \((y - 1)\):

\[
\psi_k^{(1)}(t) = \sum_{p=0}^\infty (y - 1)^{p+1} \left[ \int_0^t g_k(t - t') \left[ \partial^p D_M^{(1)}(t, t') \right] dt' \right]
\]

(50)
\[ \psi_{kj}^{(0)}(t) = \sum_{p=0}^{\infty} (y - 1)^{p+1} \left[ \int_0^t g_k(t - t') \left[ \partial^p D_{Mj}^{(0)}(t, t', 0) \right] dt' \right]. \quad (51) \]

With respect to \( y \), the GF in Eq. (44) is now analytic. Taking the first four derivatives of Eq. (44) and setting \( y = 1 \), we obtain the following statistical moments for the variable \( n_N \), which, to simplify notation, we write as \( n \):

\[
\begin{align*}
F_1 &= \langle n \rangle \\
F_2 &= \langle n(n - 1) \rangle = \langle n^2 \rangle - \langle n \rangle \\
F_3 &= \langle n(n - 1)(n - 2) \rangle = \langle n^3 \rangle - 3\langle n^2 \rangle + 2\langle n \rangle \\
F_4 &= \langle n(n - 1)(n - 2)(n - 3) \rangle = \langle n^4 \rangle - 6\langle n^3 \rangle + 11\langle n^2 \rangle - 6\langle n \rangle \\
\end{align*}
\quad (52)
\]

where \( F_l = [d^l F_N(y, t)/(dy)^l]_{y=1} \). The first four central moments, defined as \( \sigma_l = \langle (n - \langle n \rangle)^l \rangle \) for \( l = 2, 3, ... \), are

\[
\begin{align*}
\sigma_1 &= \langle n \rangle \\
\sigma_2 &= \langle n^2 \rangle - \langle n \rangle^2 \\
\sigma_3 &= \langle n^3 \rangle - 3\langle n \rangle \langle n^2 \rangle + 2\langle n \rangle^3 \\
\sigma_4 &= \langle n^4 \rangle - 4\langle n \rangle \langle n^3 \rangle + 6\langle n \rangle^2 \langle n^2 \rangle - 3\langle n \rangle^4. \\
\end{align*}
\quad (53)
\]

Solving Eq. (52) for the moments \( \langle n^l \rangle \) and inserting the solution into Eq. (53), we obtain

\[
\begin{align*}
\sigma_1 &= F_1 \\
\sigma_2 &= F_2 + F_1 - F_1^2 \\
\sigma_3 &= F_3 + 3F_2 - 3F_1 (F_1 + F_2) + F_1 + 2F_1^3 \\
\sigma_4 &= F_4 + 6F_3 + 7F_2 + F_1 - 4F_1 (F_1 + 3F_2 + F_3) + 6F_1^2 (F_1 + F_2) - 3F_1^4. \\
\end{align*}
\quad (54)
\]

Figure 2 shows a comparison between the calculations above and Gillespie simulations (see captions for details).

### 4.2 Probability distributions

The probability of observing \( n_N \) fully processed proteins is given by Eq. (5). To take high derivatives of a function \( f(x) \), it is convenient to use Cauchy’s integral formula, which states that

\[
\frac{d^p f(x)}{dx^p} = \frac{p!}{2\pi i} \oint dz \frac{f(z)}{(z - x)^{p+1}},
\quad (55)
\]

where \( f(z) \) is analytic at the complex number \( x \). The integral over the complex variable \( z \) must enclose the point \( x \) but is otherwise arbitrary. In the following calculations,
Fig. 2 Comparison of the path integral method and Gillespie simulations for the first four central moments for a system with $M = N = 3$ and the following reaction rates: $r = 5/\text{min}, d = 0.1/\text{min}, K = 1/\text{min}, a_f = a_i = 0.2/\text{min}, b_f = b_i = 0.1/\text{min}$ and $q = 0.05/\text{min}$. The initial conditions were chosen randomly: $\tilde{m}_1 = 59, \tilde{m}_2 = 88, \tilde{m}_3 = 44, \tilde{n}_1 = 982, \tilde{n}_2 = 316$ and $\tilde{n}_3 = 1977$

we choose the contour to be a unit circle in the complex plain, i.e. $z = e^{i\theta}$, simply because it is easy to work with.

To apply Eq. (55) to the GF given by Eq. (44), we must solve Eqs. (37) and (38), but with $y$ replaced by an imaginary parameter $z = e^{i\theta}$, and set $x$ in Eq. (55) to zero. Separating the real and the imaginary terms in Eqs. (37) and (38), we obtain

$$\frac{d}{dt'} \mathbf{D}_R^{(1)}(t, t') = \mathbf{r} + SD_R^{(1)}(t, t') + (\cos \theta - 1)g_k(t - t')\mathbf{B}_D^{(1)}(t, t') - \sin \theta g_k(t - t')\mathbf{B}_I^{(1)}(t, t')$$

$$\frac{d}{dt'} \mathbf{D}_I^{(1)}(t, t') = SD_I^{(1)}(t, t') + (\cos \theta - 1)g_k(t - t')\mathbf{B}_D^{(1)}(t, t') + \sin \theta g_k(t - t')\mathbf{B}_I^{(1)}(t, t')$$

$$\frac{d}{dt'} \mathbf{D}_R^{(0)}(t, t', 0) = SD_R^{(0)}(t, t', 0) + (\cos \theta - 1)g_k(t - t')\mathbf{B}_D^{(0)}(t, t', 0)$$

$$\frac{d}{dt'} \mathbf{D}_I^{(0)}(t, t', 0) = SD_I^{(0)}(t, t', 0) + (\cos \theta - 1)g_k(t - t')\mathbf{B}_D^{(0)}(t, t', 0) + \sin \theta g_k(t - t')\mathbf{B}_I^{(0)}(t, t', 0),$$

(56)
Fig. 3  a Two examples of the GF as a function of \( \theta \); b an example of a distribution constructed from 100k realizations simulated with the GA (left) and the corresponding cumulative distribution (right)

Fig. 4 Average \( n_3 \) as a function of time for \( r = 0.5, k = 0.1, K = 10, q = 0.05 \) and ten different sets of \( a_i, \bar{a}_i, b_i, \) and \( \bar{b}_i \). The circles label the times at which the average \( n_3 \) have reached half of the steady state values

where the indices \( R \) and \( I \) refer to the real and imaginary parts. Since in general Eqs. (56) cannot be solved analytically, we must solve them numerically for every value of \( \theta \) from 0 to 2\( \pi \). This can be done approximately by dividing \( \theta \) into \( J \) equal parts, i.e. \( \theta = 2\pi j / J \), and then solving Eq. (56) numerically for each \( j \). Inserting the solution into Eqs. (40) and (41), and performing the integration, we obtain the GF in Eq. (44) in terms of the discretized parameter \( \theta \). Finally, to perform the contour integral in Eq. (55), the GF must be interpolated. Figure 3a shows two examples of the interpolated GF. Notice that the real part of the GF is symmetric around \( \pi \), while the imaginary part is anti-symmetric. This means that we only need to divide \( \theta \) up to \( \pi \); for \( \pi < \theta \leq 2\pi \), we merely take the complex conjugate of the GA and change its argument to \( 2\pi - \theta \). Thus, the actual division we consider is \( \theta = \pi j / J \) between 0 and \( \pi \).

To check the validity of this discretization method, we compare it to Gillespie simulations for a system with \( M = N = 3 \) for 10 parameter sets, such that the average \( m_3 \) and \( n_3 \) at the steady state remains the same for each set. We do this by randomly selecting 10 values for each \( a_i, \bar{a}_i, b_i, \) and \( \bar{b}_i \) (\( i = 1, 2, 3 \)) from a square distribution.
Fig. 5  Probability distributions for five out of ten parameter sets (cases) for a $J = 100$, $IO = 1$; b $J = 100$, $IO = 2$; c $J = 200$, $IO = 1$; d $J = 200$, $IO = 2$

with the range $[0.01, 0.5]$ min, while keeping the other parameters fixed at the values $r = 0.1$, $d = 0.1$, $K = 10$ and $q = 0.05$. The steady state average of $m_3$ and $n_3$ effectuated by these parameter sets is 1 and 200, respectively. Because the noise in GRNs tends to be more prominent when the rising protein levels are half way to the steady state, we choose to compare our method against Gillespie simulations at these temporal points, which we label as $t_{1/2}$ (see Fig. 4).

To compute probability distributions (PDs) using Eq. (55), we start at $p = n_{av}$, where $n_{av}$ is the closest integer to the average, and then compute the PD at $n_{av} + 1$, $n_{av} + 2$, etc. until we reach a value $\leq 0.1\%$ of the PD at $n_{av}$; we do the same for the $ps$ to the left of $n_{av}$. In order to study how $J$ and the interpolation order (IO) affect the accuracy of our method, we compute the PDs for $J = 100, 200$ and the IO 1, 2, 3 and 4.

One way to compare two distributions is to look at the root mean square (RMS) of the distance between them. However, because of the finite ensemble of realizations (100k in our case), fluctuations around the true values may appear, which can unfairly add value to the RMS. A better comparison is the RMS between two cumulative PDs.
5 Discussion

We presented a novel approach to solving the master equation for a one gene system comprising of partially and fully processed mRNA and protein. The success of this (CPDs) because they tend to reduce the ensemble driven fluctuations, as Fig. 3b shows. A sample of our results is shown in Fig. 5 for $j = 100, 200$ and the IO 1 and 2. The RMS for all ten cases and for $j = 100, 200$ and the IO 1, 2, 3 and 4 is shown in Fig. 6a. Figure 6b and c show the RMS for the same ten sets of $(a_i, \bar{a}_i, b_i, \bar{b}_i)$ but with $(r = 5, K = 1)$ and $(r = 10, K = 1)$, respectively, which gives the average $n_3(t_{1/2})$ of 500 for the former and 1000 for the latter.
Fig. 7 Relative differences between the normalization of the PDs generated by our method and 1 for the ten parameter sets (cases) \((a_1, \bar{a}_1, b_1, \bar{b}_1)\) and \(a = 0.5, d = 0.1, K = 10, q = 0.05; b = 5, d = 0.1, K = 1, q = 0.05; \) and \(c = 10, d = 0.1, K = 1, q = 0.05\)

approach lies in the fact that the mRNAs influence the production of the proteins but not vice versa, and in a theorem which allows for the conversion of an integral over infinitely many paths taken by the mRNAs into an equation similar to the master equation for the mRNAs. This method is far more efficient than, e. g. the Gillespie algorithm, thanks to the analytic steps leading to Eqs. (37) and (38). The numerical efficiency of solving these equations depends on the number of divisions \(J\) of the variable \(y\) and the interpolation order (IO). Figure 6 shows the accuracy of our method against the GA for various \(J\) and IO. For \(J = 100\), the results do tend to improve with an increase in IO, in particular for b) and c), where a jump occurs between the IO 1 and 2. However, in a), there is more of a continuity between the IOs, and in fact, the first two cases show a diminishing improvement over the four IOs. Furthermore, higher IOs cannot be relied upon to fill in missing data. For \(J = 200\), the situation is
more positive: for all cases in a), b) and c), a jump occurs between IO 1 and 2, and remains more or less unchanged for IO 3 and 4.

The dependence of our method on \( J \) and the IO raises the question as to how can one know what \( J \) and IO to use in order to obtain the correct result. To answer this question, we computed the RMS distance between the normalization of each distribution (computed via our method) and unity. The results are shown in Fig. 7, where the x-axes and the legends are exactly as in Fig. 6a–c, but with the y-axes representing RMS between the normalization of the PDs and 1. We see that the graphs in Fig. 7 have a very similar trend to those in Fig. 6. The important feature that is present in both Figs. 6 and 7 is the jump between IO 1 and 2 for \( J = 200 \), after which the quality does not improve significantly for higher IOs. Based on these results, we conclude that the relative difference in normalization is a reliable measure for determining sufficient \( J \) and IO.

The success of our method relies on the separation of the variables, which reduces the dimension of the problem to one. Although we only focused on probability distributions for each variable separately, the steps detailed in the main text also allow for the calculation of joint probability distributions of two or more variables. Since Theorem 1 is true for any system in which a part of the system, \( A \), affects another part, \( B \), unidirectionaly, this approach can be applied to problems involving more than one gene, whereby the problem’s dimension would be reduced from \( \text{dim}(A) + \text{dim}(B) \) to \( \text{dim}(A) \).

Acknowledgements This work was not funded by any institution or organization. JA thanks Ekaterina Ejkova for her technical support.

Appendix A

In this section we prove that the GF for the initial probability distribution \( \delta_{m_1, \tilde{m}_1} \delta_{m_2, \tilde{m}_2} ... \delta_{m_M, \tilde{m}_M} \), where \( \tilde{m}_1, ..., \tilde{m}_M \) are the initial mRNA copy numbers, is given by Eq. (15).

Before we continue, it will prove useful to reformulate Eq. (8) in terms of a tensor product of basis vectors \( |m\rangle = |m_1\rangle |m_2\rangle ... |m_M\rangle \) and its transpose \( \langle m| = \langle m_1| \langle m_2| ... \langle m_M| \), and operators \( A_i \), \( A_i^+ \) and \( A_i^+ A_i \), whose action on \( |m\rangle \) is as follows [see review (Walczak et al. 2012) and references therein]:

\[
A_i |m\rangle = m_i |m_1\rangle |m_2\rangle ... |m_i-1\rangle ... |m_M\rangle, \\
A_i^+ |m\rangle = |m_1\rangle |m_2\rangle ... |m_i+1\rangle ... |m_M\rangle, \\
A_i^+ A_i |m\rangle = m_i |m\rangle. \tag{57}
\]

In this notation, Eq. (8) can be written as

\[
\frac{d}{dt} |\psi\rangle = H |\psi\rangle, \tag{58}
\]
where
\[
|\psi\rangle = \sum_{m} P(m, t)|m\rangle, \quad \text{(59)}
\]
\[
\sum_{m} \equiv \sum_{m_1=0}^{\infty} \cdots \sum_{m_M=0}^{\infty}. \quad \text{(60)}
\]
and
\[
\begin{align*}
H &= r(t)[A_1^+ - 1] + d[A_M - A_M^+ A_M] \\
&\quad + \sum_{i=1}^{M-1} a_i[A_i^+ A_{i+1}^+ - A_i^+ A_i] \\
&\quad + \sum_{i=1}^{M-1} \bar{a}_{i+1}[A_{i+1}^+ A_i^+ - A_{i+1}^+ A_{i+1}]. \quad \text{(61)}
\end{align*}
\]

The formal solution to Eq. (58) can be written as
\[
|\psi\rangle = e^{H(t_L-1)dt} e^{H(t_{L-2})dt} \ldots e^{H(t_0)dt} |\psi(0)\rangle, \quad \text{(62)}
\]
where
\[
|\psi(0)\rangle = \sum_{m} P(m, 0)|m\rangle \quad \text{(63)}
\]
is the initial state, and \(t_0 = 0, t_1 = dt, \ldots, Ldt = t\). Using the orthogonality relations \(\langle m_i | m_j' \rangle = \delta_{m_i m_j'} \delta_{ij}\), we obtain the probability of observing \(m\) at time \(t\) by multiplying Eq. (62) by \(|m|\):
\[
P(m, t) = \langle m | \psi \rangle = \langle m | e^{H(t_{L-1})dt} e^{H(t_{L-2})dt} \ldots e^{H(t_0)dt} |\psi(0)\rangle, \quad \text{(64)}
\]
or, in a short hand notation
\[
P(m, t) = \langle m | \psi \rangle = \langle m | G(t) |\psi(0)\rangle, \quad \text{(65)}
\]
where
\[
G(t) = T e^{\int_0^t H(t')dt'} \quad \text{(66)}
\]
and \(T\) is the time ordering operator. The generating function for the variable \(m_k\) can be written as
\[
F_k(x, t) = \sum_{m} x^{m_k} P(m, t) = \sum_{m} x^{m_k} \langle m | G(t) |\psi(0)\rangle. \quad \text{(67)}
\]
Finally, we will also need the following relation:

$$|m\rangle = \oint dz \frac{m!}{2\pi i z^{m+1}} e^{\tilde{z} z} |z\rangle,$$

(68)

where $z$ is a complex number integrated around a unit circle in the complex plain, and the state $|z\rangle$ is defined as

$$|z\rangle = \sum_{m} \frac{z^m}{m!} e^{-z^2} |m\rangle.$$

(69)

Setting $z = e^{i\theta}$, the proof becomes straightforward:

$$|m\rangle = \int_{0}^{2\pi} d\theta \frac{m!}{2\pi} e^{-i m \theta} e^{i \theta} \left[ \sum_{m'} \frac{e^{i m' \theta}}{m'} e^{-e^{i \theta} |m'|} \right] = \sum_{m'} \frac{m!}{m'}! \left[ \frac{1}{2\pi} \int_{0}^{2\pi} d\theta e^{i (m'-m) \theta} \right] |m'| = \sum_{m} \frac{m!}{m'}! \delta_{m'm} |m\rangle = |m\rangle.$$

(70)

With the initial conditions $\tilde{m} = (\tilde{m}_1, ... \tilde{m}_M)$, the generation function for the variable $m_k$ is given by

$$F_k(\tilde{m}, x, t) = \sum_{m} x^{m_k} \langle m | G(t) | \tilde{m}\rangle.$$

(71)

According to Eq. (68), the product state $|\tilde{m}\rangle$ reads

$$|\tilde{m}\rangle = \prod_{j=1}^{M} |\tilde{m}_j\rangle = \prod_{j=1}^{M} \oint dz_j \frac{\tilde{m}_j!}{2\pi i z_j^{m_j+1}} e^{z_j |z_j\rangle}.$$

(72)

Acting with $G(t)$ on $|\tilde{m}\rangle$ yields

$$G(t) |\tilde{m}\rangle = \prod_{j=1}^{M} \oint dz_j \frac{\tilde{m}_j!}{2\pi i z_j^{m_j+1}} e^{z_j} [G(t) |z_j\rangle].$$

(73)

The generating function can now be written as

$$F_k(\tilde{m}, x, t) = \sum_{m} x^{m_k} \langle m | G(t) | \tilde{m}\rangle$$

$$= \prod_{j=1}^{M} \oint dz_j \frac{\tilde{m}_j!}{2\pi i z_j^{m_j+1}} e^{z_j} \left[ \sum_{m=0}^{\infty} x^{m_k} \langle m | G(t) | z_j\rangle \right].$$

(74)
Since the state \( |z_j\rangle \) is Poisson, the inner product

\[
\prod_{j=1}^{M} \langle m|G(t)|z_j\rangle = \prod_{j=1}^{M} \frac{\gamma_j^{mj}}{m_j!} e^{-\gamma_j}
\]  

(75)

is the Poisson distribution, where, as in Eq. (10), \( \gamma_j \) satisfies

\[
\frac{d\gamma_j}{dt} = S_j \gamma + r_j,
\]  

(76)

such that \( \gamma_j(0) = z_j \). The solution of Eq. (76) can be written as

\[
\gamma_k(t) = \gamma_k^{(1)}(t) + \sum_{j=1}^{M} \gamma_{kj}^{(0)}(t) z_j,
\]  

(77)

where

\[
\gamma_k^{(1)}(t) = \int_{0}^{t} f_k(t - t') r(t') dt', \quad \text{and} \quad \gamma_{kj}^{(0)}(t) = \sum_{i=1}^{M} U_{ki} U_{ij}^{-1} e^{S_i t}.
\]  

(78)

Equation (67) can now be written as

\[
F_k(\hat{m}, x, t) = \sum_{m} x^{m_k} \langle m|G(t)|\hat{m}\rangle \\
= \prod_{j=1}^{M} \int d z_j \frac{\tilde{m}_j !}{2\pi i z_j^{m_j + 1}} e^{z_j} \left[ \sum_{m} x^{m_k} \langle m|G(t)|z_j\rangle \right] \\
= \prod_{j=1}^{M} \int d z_j \frac{\tilde{m}_j !}{2\pi i z_j^{m_j + 1}} e^{z_j} \left[ \sum_{m} x^{m_k} \frac{\gamma_j^{mj}}{m_j!} e^{-\gamma_j} \right] \\
= \prod_{j=1}^{M} \int d z_j \frac{\tilde{m}_j !}{2\pi i z_j^{m_j + 1}} e^{z_j} e^{(x-1)\gamma_k} \\
\times \exp \left[ (x - 1) \left( \gamma_k^{(1)}(t) + \sum_{l=1, l \neq k}^{M} \gamma_{kl}^{(0)}(t) z_l \right) \right] \\
= e^{-(x-1)\gamma_k^{(1)}(t)} \prod_{j=1}^{M} \left[ \int d z_j \frac{\tilde{m}_j !}{2\pi i z_j^{m_j + 1}} e^{z_j} [1 + (x-1)\gamma_{kj}^{(0)}(t)] \right].
\]  

(79)
Invoking the integral identity

\[
\int_0^{2\pi} d\theta \frac{m_i}{2\pi} e^{-im\theta} e^{ae^{i\theta}} = a^m. \tag{80}
\]

we obtain

\[
F_k(\mathbf{m}, x, t) = e^{-(x-1)\gamma_k^{(1)}(t)} \prod_{j=1}^{M} \left[ 1 + (x-1)\gamma_k^{(0)} \right] \tilde{m}_j. \tag{81}
\]

**Appendix B**

In this section we give a proof of Theorem 1.

**Proof** The probability to observe a path \( \{m\} \) is given by

\[
\langle \mathbf{m}(t_L) | e^{H(t_{L-1})dt} | \mathbf{m}(t_{L-1}) \rangle \langle \mathbf{m}(t_{L-1}) | e^{H(t_{L-2})dt} | \mathbf{m}(t_{L-2}) \rangle \ldots \langle \mathbf{m}(t_1) | e^{H(t_0)dt} | m(t_0) \rangle. \tag{82}
\]

Then,

\[
\sum_{\text{all paths}} P(\mathbf{m}(t_0), 0) \mathcal{P}(\{\mathbf{m}\}) \exp \left[ \int_0^t W(\mathbf{m}(t'), t, t')dt' \right] = \sum_{\text{all paths}} P(\mathbf{m}(t_0), 0) \prod_{i=0}^{L-1} e^{W(\mathbf{m}(t_i), t, t_i)dt}
\times \langle \mathbf{m}(t_L) | e^{H(t_{L-1})dt} | \mathbf{m}(t_{L-1}) \rangle \langle \mathbf{m}(t_{L-1}) | e^{H(t_{L-2})dt} | \mathbf{m}(t_{L-2}) \rangle \ldots \langle \mathbf{m}(t_1) | e^{H(t_0)dt} | m(t_0) \rangle. \tag{83}
\]

We can rearrange these products so as to combine the indices in \( W(\mathbf{m}(t_i), t, t_i) \) with those labeling the basis, like so

\[
\sum_{\text{all paths}} \langle \mathbf{m}(t_L) | e^{H(t_{L-1})dt} \left[ e^{W(\mathbf{m}(t_{L-1}), t, t_{L-1})dt} | \mathbf{m}(t_{L-1}) \rangle \right] \langle \mathbf{m}(t_{L-1}) | e^{H(t_{L-2})dt} \left[ e^{W(\mathbf{m}(t_{L-2}), t, t_{L-2})dt} | \mathbf{m}(t_{L-2}) \rangle \right] \ldots \langle \mathbf{m}(t_1) | e^{H(t_0)dt} \left[ e^{W(\mathbf{m}(t_0), t, t_0)dt} | \mathbf{m}(t_0) \rangle \right] \right] P(\mathbf{m}(t_0), 0). \tag{84}
\]

Since \( A_i^+ A_j | \mathbf{m}(t_k) \rangle = m_i(t_k) | \mathbf{m}(t_k) \rangle \), for \( i = 1, 2, \ldots, M \), we can replace the set \( \mathbf{m} = (m_1, m_2, \ldots, m_M) \) in \( W(\mathbf{m}(t_i), t, t_i) \) with the set \( A^+ A = (A_1^+ A_1, A_2^+ A_2, \ldots, A_M^+ A_M) \). This allows us to move the sums over individual times from the very front of Eq. (84) to the immediate left of each basis:
\[
\sum_{m(t_L)} \langle m(t_L) | e^{H(t_L-1)dt} e^{W(A^+A,t,t_L-1)dt} \left[ \sum_{m(t_{L-1})} |m(t_{L-1})\rangle \langle m(t_{L-1})| \right] \\
\times e^{H(t_{L-2})dt} e^{W(A^+A,t,t_{L-2})dt} \left[ \sum_{m(t_{L-2})} |m(t_{L-2})\rangle \langle m(t_{L-2})| \right] \ldots \\
\ldots e^{H(t_0)dt} e^{W(A^+A,t,t_0)dt} \sum_{m(t_0)} |m(t_0)\rangle P(m(t_0), 0),
\]

where we have rearranged the square brackets to call attention to the identity

\[
\mathbb{1} = \sum_m |m\rangle \langle m|.
\]

Hence, Eq. (85) becomes

\[
\sum_{m(t_L)} \langle m(t_L) | e^{H(t_L-1)dt} e^{W(A^+A,t,t_L-1)dt} \ldots e^{H(t_0)dt} e^{W(A^+A,t,t_0)dt} \\
\times \sum_{m(t_0)} |m(t_0)\rangle P((m(t_0), 0).
\]

The entire operator in the square brackets now acts on the initial state \(|\psi(0)\rangle\) defined in Eq. (63). Invoking the relation \(e^{sV}e^{sU} = e^{s(V+U)} + \mathcal{O}(s^2)\) for arbitrary matrices \(V\) and \(W\) in the limit \(s \to 0\), expression (87) can be written as

\[
\sum_{m'} \left[ \langle m' | e^{\tilde{H}(t,t_L-1)dt} e^{\tilde{H}(t,t_{L-2})dt} \ldots e^{\tilde{H}(t,t_0)dt} |\psi(0)\rangle \right],
\]

where \(m\) and \(m'\) refer to the set \((m_1, m_2, \ldots, m_M)\) at \(t' = 0\) and \(t' = t\), respectively, and \(\tilde{H}(t,t') = H(t') + W(A^+A, t, t')\). The expression in Eq. (88), satisfies the equation

\[
\frac{d}{dt'} |\varphi\rangle = \tilde{H}(t, t') |\varphi\rangle,
\]

or, if we define the state \(|\varphi\rangle\) as

\[
|\varphi\rangle = \sum_m Q(m, t, t') |m\rangle,
\]

\(Q(m, t, t')\) must satisfy Eq. (26). For \(t' = t_0 = 0\), the only term in the brackets of Eq. (88) is \(e^{\tilde{H}(t,0)dt} = 1 + \tilde{H}(t, 0)dt + \mathcal{O}(dt^2)\); hence, as \(dt \to 0\), Eq. (88) reduces to

\[
\sum_{m'} \langle m' | \varphi\rangle = \sum_{m'} \langle m' | \left[ 1 + \tilde{H}(t, 0)dt + \mathcal{O}(dt^2) \right] |\psi(0)\rangle
\]

and \(|\varphi\rangle \to |\psi(0)\rangle\), or \(Q(m, t, 0) = P(m, 0)\). \(\square\)
Appendix C

To prove Eq. (42), we follow the same logic as in “Appendix A”, except that now the operator $G(t)$ must be replaced with

$$G(t) = T \exp \left[ \int_0^t [H(t') + (y - 1)g_k(t - t')mM] \, dt' \right].$$

(92)

The action of $G(t)$ on a Poisson state $|z\rangle$, where $z$ is any complex number, produces the state

$$G(t)|z\rangle = |z(t)\rangle = \sum_m \prod_{i=1}^M [\xi_i(t, t)]^{m_i} \frac{m_i!}{m_i!} e^{-h(t, t)} |m\rangle,$$

(93)

where $\xi_i(t, t')$ and $h(t, t')$ are the solutions of Eqs. (29) and (30), respectively, with the initial conditions $\xi_i(t, 0) = z_i$ and $h(t, 0) = \sum_i z_i$. Invoking Eq. (68), we can write

$$\sum_m \langle m | z(t) \rangle = \sum_m \prod_{i=1}^M [\xi_i(t, t)]^{m_i} \frac{m_i!}{m_i!} e^{-h(t, t)} = \exp \left[ \sum_{i=1}^M \xi_i(t, t) - h(t, t) \right] = \exp \left[ (y - 1) \int_0^t g_k(t - t') \xi_M(t, t') \, dt \right] = \exp \left[ (y - 1) \left( \psi_k^{(1)}(t) + \sum_{l=1}^M \psi_{kl}^{(0)}(t)z_l \right) \right].$$

(94)

Finally for the initial conditions $Q(m', t, 0) = 0_{m_1, \tilde{m}_1} ... 0_{M, \tilde{m}_M}$, we obtain

$$Q(t) = \sum_m \langle m | G(t)|\tilde{m}\rangle$$

$$= \prod_{j=1}^M \left[ \int_0^{2\pi} d\theta_j \frac{\tilde{m}_j!}{2\pi} e^{-i\theta_j\tilde{m}_j} e^{i\theta_j} \right] \sum_m \langle m | z(t) \rangle$$

$$= \prod_{j=1}^M \left[ \int_0^{2\pi} d\theta_j \frac{\tilde{m}_j!}{2\pi} e^{-i\theta_j\tilde{m}_j} e^{i\theta_j} \right] \exp \left[ (y - 1) \left( \psi_k^{(1)}(t) + \sum_{l=1}^M \psi_{kl}^{(0)}(t)e^{i\theta_l} \right) \right]$$

$$= \prod_{j=1}^M \left[ \int_0^{2\pi} d\theta_j \frac{\tilde{m}_j!}{2\pi} e^{-i\theta_j\tilde{m}_j} e^{i\theta_j} \right] \exp \left[ (y - 1) \psi_k^{(1)}(t) \right]$$

$$= \prod_{j=1}^M (y - 1) \psi_{kj}^{(0)}(t) + 1 \right]^{m_j} e^{i\theta_j} \left[ \right]$$

(95)
References

Albert J (2016a) A hybrid of the chemical master equation and the Gillespie algorithm for efficient stochastic simulations of sub-networks. PLoS ONE 11(3):e0149909

Albert J (2016b) Stochastic simulation of reaction subnetworks: exploiting synergy between the chemical master equation and the Gillespie algorithm. In: AIP conference proceedings, vol 1790, no 1, p 150026

Albert J, Rooman M (2016) Probability distributions for multimeric systems. J Math Biol 72(1–2):157–169

Alfonsi A, Cancès E, Turinici G, Ventura BD, Huisinga W (2005) Adaptive simulation of hybrid stochastic and deterministic models for biochemical systems. In: ESAIM: proceedings, vol 14, pp 1–13

Blake WJ, Kærn M, Cantor CR, Collins JJ (2003) Noise in eukaryotic gene expression. Nature 422:633–637

Bokes P, King JR, Wood ATA, Loose M (2012a) Exact and approximate distributions of protein and mRNA levels in the low-copy regime of gene expression. J Math Biol 64(5):829–854

Bokes P, King JR, Wood ATA, Loose M (2012b) Multiscale stochastic modelling of gene expression. J Math Biol 65(3):493–520

Burrage K, Tian T, Burrage P (2004) A multi-scaled approach for simulating chemical reaction systems. Prog Biophys Mol Biol 85:217–234

Cao Y, Li H, Petzold L (2004) Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. J Chem Phys 121:4059

Cao Y, Gillespie DT, Petzold LR (2005a) Avoiding negative populations in explicit Poisson tau-leaping. J Chem Phys 123(5):054104

Cao Y, Gillespie DT, Petzold LR (2005b) Efficient step size selection for the tau-leaping simulation method. J Chem Phys 124(4):044109

Duso L, Zechnier C (2018) Selected-node stochastic simulation algorithm. J Chem Phys 148:164108

Gibson MA, Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. J Phys Chem 104(9):1876–1889

Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81(25):2340–2361

Gillespie DT (2001) Approximate accelerated stochastic simulation of chemically reacting systems. J Chem Phys 115(4):1716

Jahnke T, Altıntan D (2010) Efficient simulation of discrete stochastic reaction systems with a splitting method. BIT Numer Math 50(4):797–822

Jahnke T, Huisinga W (2007) Solving the chemical master equation for monomolecular reaction systems analytically. J Math Biol 54(1):1–26

Kurasov P, Lück A, Mugnolo D, Wolf V (2018) Stochastic hybrid models of gene regulatory networks. arXiv:1803.10958

Pendar H, Platini T, Kulkarni RV (2013) Exact protein distributions for stochastic models of gene expression using partitioning of Poisson processes. Phys Rev E 87:042720

Popović N, Marr C, Swain PS (2016) A geometric analysis of fast–slow models for stochastic gene expression. J Math Biol 72(1–2):87–122

Shahrezaei V, Swain PS (2008) Analytical distributions for stochastic gene expression. PNAS 105(45):17256–17261

Van Kampen NG (2007) Stochastic processes in physics and chemistry 3rd print. North Holland, Amsterdam

Veerman F, Marr C, Popović N (2018) Time-dependent propagators for stochastic models of gene expression: an analytical method. J Math Biol 77(2):261–312

Walczak AM, Mugler A, Wiggins CH (2012) Analytic methods for modeling stochastic regulatory networks computational modeling of signaling. Networks 880:273–322

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.