Stochastic gene expression with delay

BY MARTIN JANSEN¹ AND PETER PFAFFELHUBER²

May 30, 2013

Abstract

The expression of genes usually follows a two-step procedure. First, a gene (encoded in the genome) is transcribed resulting in a strand of (messenger) RNA. Afterwards, the RNA is translated into protein. Classically, this gene expression is modeled using a Markov jump process including activation and deactivation of the gene, transcription and translation rates together with degradation of RNA and protein. We extend this model by adding delays (with arbitrary distributions) to transcription and translation. Such delays can e.g. mean that RNA has to be transported to a different part of a cell before translation can be initiated.

Already in the classical model, production of RNA and protein come in bursts by activation and deactivation of the gene, resulting in a large variance of the number of RNA and proteins in equilibrium. We derive precise formulas for this second-order structure with the model including delay in equilibrium. As a general fact, the delay decreases the variance of the number of RNA and proteins.

1 Introduction

The central dogma of molecular biology is that a gene (encoded within the genome) is transcribed into (messenger) RNA (also abbreviated mRNA), which in turn is translated into protein, the whole process also being called gene expression. Mathematical models for this process have by now been studied for a long time; see e.g. Rigney and Schieve (1977), Berg (1978), McAdams and Arkin (1997), Swain et al. (2002), Paulsson (2005), Cottrell et al. (2012).

Within a single cell, there are either one or two copies of the genome, and only a few genes code for the same protein. For this reason, there is only a small copy number of RNA and protein molecules, making them a noisy (i.e. stochastic) quantity. While this stochasticity has been assumed to be detrimental to the cellular function, it can also help a cell to adapt to fluctuating environments, or help to explain genetically homogeneous but phenotypically heterogeneous cellular populations (Kaern et al., 2005).

In order to consider stochasticity in gene expression, Swain et al. (2002) distinguish between intrinsic and extrinsic noise. The latter accounts for changing environments of the cell, while the former accounts for the stochastic process of transcription and translation. Let us look at the possible sources of intrinsic noise in more detail; see e.g. Zhu et al. (2007), Roussel and Zhu (2006).

(i) Various mechanisms for gene expression require random events to occur. In order to understand this let us have a closer look at the mechanisms of gene expression. Transcription starts when RNA polymerase (which are enzymes helping in the synthetization of RNA) binds to the promoter region of the gene, forming an elongation complex. This elongation complex is then ready to start
walking along the DNA, reading off DNA and making RNA. Before the transcript is released, a
ribosome binding site (which is needed for translation) is being produced on the transcript. Then
follows translation which starts when a free ribosome binds to the ribosome binding site of the
transcript and again is a complex process involving many chemical reactions, which lead to fluc-
tuations.

(ii) Another source of noise comes from turning genes on and off. This means that transcription
factors can bind to promoter regions of the gene and only bound (or unbound) promoters can ini-
tiate transcription. This process has been found to be the most important source of randomness
for gene expression \cite{Zhu2008, Iyer-Biswas2009}. The effect of this activation and inactivation of
genes is a burst-like behavior of protein production, already apparent in
McAdams and Arkin \cite{1997}. When considering the amount of RNA within the cell during the
production of a specific protein, it is hence not surprising that production of RNA comes in bursts,
which are related to times when the gene is turned on. This burst-like behavior is inherited to
protein formation, which also comes in bursts during transcription.

In our paper, we are studying yet another effect leading to intrinsic noise in gene expression,
which is (random) delay. To understand this, assume that the RNA transcript is available to
the system only after some time lag after initiation of transcription, and the same applies to the
translated protein. In real-life applications, such gene expression delays have been found e.g. by
Lewis \cite{2003}, Monk \cite{2003}, Barrio et al. \cite{2006}, Bratsun et al. \cite{2005}.

While our modeling approach only takes a single gene/RNA/protein triple into account, the
field of systems biology aims at unraveling interactions between genes in so-called pathways. It
seems clear that randomness as well as delays can accumulate in such networks of interacting genes
and proteins. As a simple example, the transcription factor regulating the expression of gene A
is coded by a gene B which in turn may be regulated by gene A (or by itself), which can lead to
a bi-modal distribution of the number of proteins encoded by gene A or B; see e.g. Kaern et al.
\cite{2005}. Although such feedback systems are highly interesting, we are not touching on this level
of complexity.

Simulation of chemical systems, or in silico modeling, today paves the way to understanding
complex cellular processes. While the Gillespie algorithm is a classical approach for stochastic
simulations \cite{Gillespie1977} – see also the review Gillespie et al. \cite{2013} – chemical delay models
have as well been algorithmically studied. Various explicit simulation schemes for delay models –
in particular in the field of stochasticity in gene expression – have been given; see Bratsun et al.
\cite{2005}, Roussel and Zhu \cite{2006}, Barrio et al. \cite{2006}, Cai \cite{2007}, Tian et al. \cite{2007}, Anderson
\cite{2007}, Ribeiro \cite{2010}, Tian \cite{2013}. In addition, Barrio et al. \cite{2013} provide methods in order
to obtain the correct form of delay distribution which can be fed into such simulation studies, at
least for first-order reaction networks.

Outline: In Section 2 we introduce our delay model using a classical approach of stochastic
time-change equations as well as a description of the system in equilibrium. Then, we present
our main results in Section 3. Basically, Theorems 3.2 and 3.4 give the second-order structure
of the number of RNA and protein in equilibrium under the delay model, respectively. Finally, we
discuss an application in Section 4 which comes with almost deterministic delay.
2 The model

In order to be able to model gene expression in a sophisticated way, we now give our delay model. Using the terminology from Roussel (1996), we may write

\[ \begin{align*}
\text{inactive gene} & \xrightarrow{\lambda_1^+} \text{active gene} \\
\text{active gene} & \xrightarrow{\lambda_1^-} \text{active gene(t) + RNA transcript(t + G_2)} \\
\text{RNA transcript} & \xrightarrow{\lambda_3} \text{RNA transcript(t) + protein(t + G_3)} \\
\text{RNA transcript} & \xrightarrow{1/\tau_2} \emptyset \\
\text{protein} & \xrightarrow{1/\tau_3} \emptyset.
\end{align*} \]

Essentially, (1) is an extension of the well-studied model of gene expression, as e.g. given in Paulsson (2005). Gene expression of \( n_{\text{max}} \) similar genes is studied. Each gene is activated and deactivated at rates \( \lambda_1^+ \) and \( \lambda_1^- \), respectively. (Additionally, we will set \( \tau_1 = 1/(\lambda_1^+ + \lambda_1^-) \)). Every active gene creates the RNA transcript at rate \( \lambda_2 \), which is degraded at rate \( 1/\tau_2 \). However, an RNA molecule is available for the system (i.e. can be translated) only some random delay time \( G_2 \) after its creation, where \( G_2 \) is an independent random variable with distribution \( \mu \). Then, each RNA transcript available for the system initiates translation of protein at rate \( \lambda_3 \) which in turn degrades at rate \( 1/\tau_3 \). Again, it takes a delay of a random time \( G_3 \), distributed according to \( \nu \) and independent of everything else, that the protein molecule is available for the system (i.e. for other downstream processes).

We note that (1) is a special case of a model studied in Zhu et al. (2007). Since they consider the ribosome binding site as an own chemical species, their model requires more delay random variables. Moreover, they distinguish gene expression in prokaryotes (bacteria) and eukaryotes (higher organisms), the main difference being that only eukaryotes have a cellular core. As a consequence, in prokaryotes translation can already be initiated when transcription is not complete yet. (The ribosome can bind to the ribosome binding site while the RNA transcript is still being produced.) The simplification (6) and (7) in Zhu et al. (2007) for gene expression in prokaryotes, are in line with (1) for \( G_2 = 0 \). In addition, for the same simplification in eukaryotes (see their equation (5)), we exactly recover (1).

The question we ask is about the equilibrium behavior of the number of available RNA molecules and proteins. We restrict our study to \( n_{\text{max}} = 1 \), because all genes are independent. We define, with \( t \in \mathbb{R} \),

\[ \begin{align*}
\mathcal{N}_1(t) &= \text{number of active genes (either 0 or 1) at time } t, \\
\mathcal{N}_2(t) &= \text{number of available RNA at time } t, \\
\mathcal{N}_3(t) &= \text{number of available proteins at time } t.
\end{align*} \]

Putting (1) into well-established time-change equations (Anderson and Kurtz, 2011), in equilib-
rium we get

\[ N_1(t) = \mathcal{P}_1 \left( \int_{-\infty}^{t} \lambda_1^2 ds \right) - \mathcal{P}_2 \left( \int_{-\infty}^{t} \lambda_1^2 ds \right), \]
\[ N_2(t) = \mathcal{P}_3 \left( \int_{-\infty}^{t} \int_{0}^{\infty} \lambda_2 N_1(s - \delta) \mu(\delta) ds \right) - \mathcal{P}_4 \left( \int_{-\infty}^{t} \int_{0}^{\infty} \frac{1}{\tau_2} N_2(s) ds \right), \]
\[ N_3(t) = \mathcal{P}_5 \left( \int_{-\infty}^{t} \int_{0}^{\infty} \lambda_3 N_2(s - \delta) \nu(\delta) ds \right) - \mathcal{P}_6 \left( \int_{-\infty}^{t} \int_{0}^{\infty} \frac{1}{\tau_3} N_3(s) ds \right) \]

for independent, unit rate Poisson processes \( \mathcal{P}_1, \ldots, \mathcal{P}_6 \). (Formally, we need to write integrals \( \int_{-\infty}^{t} ds \) and then obtain the equilibrium by letting \( T \to \infty \). Also note that [Anderson and Kurtz (2011)] indicate that such delay time-change equations do have a unique solution.) Another description gives the distribution of RNA and protein in equilibrium. If the gene is active, only some random delay time \( G_2 \geq 0 \) with distribution \( \mu \) later, the RNA is available. Therefore,

\[ N_2(t) \sim \text{Poi} \left( \lambda_2 \int_{0}^{\infty} e^{-\frac{r}{\tau_2}} \int_{0}^{\infty} N_1(t - r - \delta) \mu(\delta) dr \right) \]
\[ = \text{Poi} \left( \lambda_2 \tau_2 \int_{0}^{\infty} N_1(t - s) \left( \exp \left( \frac{1}{\tau_2} \right) * \mu \right)(ds) \right), \]

where \( \exp(1/\tau) \) is the exponential distribution with expectation \( \tau \) and \( * \) denotes the convolution of measures. Indeed, every RNA available at time \( t \) was produced at some time \( t - r \), which only works if at time \( t - r - G_2 \) (where \( G_2 \sim \mu \), the gene was active. In addition, the RNA must not be degraded during time \( r \), which happens with probability \( e^{-r/\tau_2} \). Using the same kind of arguments, we set for another delay \( G_3 \geq 0 \) with distribution \( \nu \)

\[ N_3(t) \sim \text{Poi} \left( \lambda_3 \int_{0}^{\infty} e^{-\frac{r}{\tau_3}} \int_{0}^{\infty} N_2(t - r - \delta) \nu(\delta) dr \right) \]
\[ = \text{Poi} \left( \lambda_3 \tau_3 \int_{0}^{\infty} N_2(t - s) \left( \exp \left( \frac{1}{\tau_3} \right) * \nu \right)(ds) \right). \]

Frequently (see e.g. [Paulsson (2005)]) stochasticity in gene expression is studied through the Master equation. We stress that the usual Master equation is unsuitable to be used for an arbitrary delay of RNA and protein production. The reason is simply that by a non-exponentially distributed equation. We stress that the usual Master equation is unsuitable to be used for an arbitrary delay

\[ \text{Proposition 3.1 (First and second order structure of } N_1 \text{.) The expectation, variance and covari-} \]

3 Results

We are now ready to formulate and prove our results on the number of active genes (Proposition 3.1), the amount of RNA molecules (Theorem 3.2) and the number of available protein (Theorem 3.4). We will start with first and second moments of the active genes.

**Proposition 3.1** (First and second order structure of \( N_1 \)). The expectation, variance and covari-
The expectation, variance and covariance of RNA in equilibrium are given by

\[
E[N(t)] = \lambda_1^+ \tau_1, \quad \tau \geq 0
\]  

\[
Cov[N(0), N(t)] = e^{-\mu t} \lambda_1^+ \lambda_1^- \tau_1^2, \quad \tau \geq 0
\]  

\[
Var[N(t)] = \lambda_1^+ \lambda_1^- \tau_1^2.
\]  

Proof. Let \( P \) be a Poisson process with rate \( \lambda_1^+ + \lambda_1^- \). For activation and inactivation of the gene, we can say that upon an event in \( P \), the gene is activated with probability \( \lambda_1^+ \tau_1 \), and deactivated with probability \( \lambda_1^- \tau_1 \). Hence,

\[
E[N(t)] = P(N(t) = 1) = \frac{\lambda_1^+}{\lambda_1^+ + \lambda_1^-} = \lambda_1^+ \tau_1
\]

because the gene is active iff the last activation event was after the last de-activation event. Moreover, for \( t \geq 0 \)

\[
Cov[N(0), N(t)] = E[Cov[N(0), N(t)] | P[0, t]] = P[P[0, t] = \theta] \cdot Var[N(0)]
\]

\[
e^{-\mu t} \lambda_1^+ \lambda_1^- \tau_1^2.
\]

Finally, (7) follows from (6) by taking \( t = 0 \).

Next, we derive our results for the amount of RNA.

**Theorem 3.2** (Expectation, variance and covariance of RNA in equilibrium). The expectation, variance and covariance of \( N_2 \) in equilibrium are given by

\[
E[N_2(t)] = \lambda_1^+ \lambda_2 \tau_1 \tau_2, \quad \tau \geq 0
\]

\[
Cov[N_2(0), N_2(t)] = \lambda_1^+ \lambda_2 \tau_1 \tau_2 e^{-\mu t}
\]

\[
+ \lambda_1^+ \lambda_1^- \lambda_2^2 \tau_1^2 \frac{\tau_2}{\tau_1 + \tau_2} \frac{\tau_1}{\tau_1 - \tau_2} - \tau_2 E[e^{-\mu \frac{G_2 - G_2'}{\tau_2}}],
\]

\[
Var[N_2(t)] = \lambda_1^+ \lambda_2 \tau_1 \tau_2
\]

\[
+ \lambda_1^+ \lambda_1^- \lambda_2^2 \tau_1^2 \tau_2^2 \frac{\tau_1}{\tau_1 + \tau_2} \frac{\tau_2}{\tau_1 - \tau_2} - \tau_2 E[e^{-\mu \frac{G_2 - G_2'}{\tau_2}}],
\]

where \( G_2, G_2' \) are independent and \( G_2, G_2' \sim \mu \).

**Remark 3.3** (Convexity and deterministic delay). We note that the last terms in (10) and (11) stay bounded, even for \( \tau_1 \to \tau_2 \). Moreover, for \( X \geq 0 \) the map \( t \mapsto E[e^{-\mu t}] \) is convex. Hence, assuming \( \tau_1 > \tau_2 \) without loss of generality, we obtain, using the convex combination \( 1 = \frac{\tau_1}{\tau_1 - \tau_2} - \frac{\tau_2}{\tau_1 - \tau_2} \) (and noting that \( \frac{\tau_1}{\tau_1 - \tau_2} > 1 \)) that for a random variable \( S \)

\[
f_S(\tau_1, \tau_2) := \frac{\tau_1 E[e^{-\mu \frac{S}{\tau_1}}] - \tau_2 E[e^{-\mu \frac{S}{\tau_2}}]}{\tau_1 - \tau_2}
\]

\[
\leq E\left[\exp\left( - \left( \frac{1}{\tau_1} \frac{\tau_1}{\tau_1 - \tau_2} - \frac{1}{\tau_2} \frac{\tau_2}{\tau_1 - \tau_2} \right) S \right) \right] = 1
\]

with equality if and only if \( V[S] = 0 \). In particular we see that

\[
Var[N_2(t)] \leq \lambda_1^+ \lambda_2 \tau_1 \tau_2 + \lambda_1^+ \lambda_1^- \lambda_2^2 \tau_1^2 \frac{\tau_1}{\tau_1 + \tau_2}
\]
with equality if and only if \( \mu \) is a delta-measure. In particular, we see that stochastic delay leads to a decrease in the variance for \( N_2 \) in equilibrium. For a deterministic delay, we obtain

\[
\frac{\text{Var}[N_2(t)]}{\mathbb{E}[N_2(t)]^2} = \frac{1}{\mathbb{E}[N_2(t)]} + \frac{\lambda_1^+ \lambda_1^- \lambda_2^+ \tau_1^3 \tau_2^2}{(\tau_1 + \tau_2) \lambda_2^+(\lambda_1^+)^2 \tau_1^2 \tau_2^2} = \frac{1}{\mathbb{E}[N_2]} + \frac{\lambda_1^- \tau_1}{\lambda_1^+ \tau_1 + \tau_2},
\]

which equals the numerical value in the absence of delay, \( \mu = \delta_0 \); see equation (5) in Paulsson (2005).

Proof of Theorem 3.2. Using (3) and Proposition 3.1, we write

\[
\mathbb{E}[N_2(t)] = \mathbb{E} \left[ \lambda_2 \tau_2 \int_0^\infty N_1(t-s) \left( \exp \left( \frac{1}{\tau_2} \right) \ast \mu \right)(ds) \right]
= \lambda_2 \tau_2 \int_0^\infty \mathbb{E}[N_1(t-s)] \left( \exp \left( \frac{1}{\tau_2} \right) \ast \mu \right)(ds)
= \lambda_2 \tau_2 \lambda_1^+ \tau_1.
\]

For (10) and (11), it clearly suffices to prove (10), since the variance formula just requires to take \( t = 0 \). From our model, we can write in the case \( t \geq 0 \) (compare with the explanation given below (3))

\[
N_2(0) = X + Y,
N_2(t) = X + Z,
\]

where \( X, Y \) and \( Z \) are conditionally independent given \( N_1 \) such that

\[
X \sim \text{Poi} \left\{ \lambda_2 \int_0^\infty e^{-\frac{r + s}{\tau_2}} \int_0^\infty N_1(-r - \delta) \mu(d\delta) dr \right\}
= \text{Poi} \left\{ \lambda_2 \tau_2 e^{-\frac{t}{\tau_2}} \int_0^\infty N_1(-s) \left( \exp \left( \frac{1}{\tau_2} \right) \ast \mu \right)(ds) \right\}
\]

is the number of RNAs available by time 0 which will not be degraded by time \( t \),

\[
Y \sim \text{Poi} \left\{ \lambda_2 \int_0^\infty e^{-\frac{r}{\tau_2}} \left( 1 - e^{-\frac{t}{\tau_2}} \right) \int_0^\infty N_1(-r - \delta) \mu(d\delta) dr \right\}
= \text{Poi} \left\{ \lambda_2 \tau_2 \left( 1 - e^{-\frac{t}{\tau_2}} \right) \int_0^\infty N_1(-s) \left( \exp \left( \frac{1}{\tau_2} \right) \ast \mu \right)(ds) \right\}
\]

is the number of RNAs available by time 0 which will be degraded by time \( t \),

\[
Z \sim \text{Poi} \left\{ \lambda_2 \int_0^t e^{-\frac{r}{\tau_2}} \int_0^\infty N_1(t-r-\delta) \mu(d\delta) dr \right\}
= \text{Poi} \left\{ \lambda_2 \int_0^\infty e^{-\frac{r}{\tau_2}} \chi_{t \leq r} \int_0^\infty N_1(t-r-\delta) \mu(d\delta) dr \right\}
\]

is the number of RNAs available only after time 0 and present by time \( t \).
Hence, we get
\[
\text{Cov}[N_2(0), N_2(t)] = \text{Cov}[X + Y, X + Z]
\]
\[
= E[\text{Var}[X|N_1]] + \text{Var}[E[X|N_1]] + \text{Cov}[E[X|N_1], E[Z|N_1]]
\]
\[
+ \text{Cov}[E[X|N_1], E[Y|N_1]] + \text{Cov}[E[Y|N_1], E[Z|N_1]]
\]
\[
= E\left[\lambda_2 \tau_2 e^{-\frac{1}{\tau_2}} \int_0^\infty N_1(-s) \left( \exp \left( \frac{1}{\tau_2} \right) * \mu \right) (ds) \right]
\]
\[
+ \text{Var}\left[\lambda_2 \tau_2 e^{-\frac{1}{\tau_2}} \int_0^\infty N_1(-s) \left( \exp \left( \frac{1}{\tau_2} \right) * \mu \right) (ds) \right]
\]
\[
+ \text{Cov}\left[\lambda_2 \tau_2 e^{-\frac{1}{\tau_2}} \int_0^\infty N_1(-s) \left( \exp \left( \frac{1}{\tau_2} \right) * \mu \right) (ds),
\lambda_2 \int_0^\infty e^{-\frac{1}{\tau_2}} 1_{r \leq t} \int_0^\infty N_1(t - r - \delta) \mu(d\delta) dr \right]
\]
\[
+ \text{Cov}\left[\lambda_2 \tau_2 e^{-\frac{1}{\tau_2}} \int_0^\infty N_1(-s) \left( \exp \left( \frac{1}{\tau_2} \right) * \mu \right) (ds),
\lambda_2 \int_0^\infty e^{-\frac{1}{\tau_2}} 1_{r \leq t} \int_0^\infty N_1(t - r - \delta) \mu(d\delta) dr \right]
\]
\[
=: A_1 + A_2 + A_3 + A_4 + A_5.
\]
We treat the terms separately and write, using Proposition 3.1 for independent, \( \exp(1/\tau_2) \)-distributed random variables \( T_2, T'_2 \),

\[
A_1 = \lambda_1^+ \lambda_2 \tau_1 \tau_2 e^{-\frac{1}{\tau_2}},
\]

\[
A_2 + A_4 = \lambda_2^2 \tau_2^2 \left( e^{-\frac{1}{\tau_2}} + e^{-\frac{1}{\tau_2}} (1 - e^{-\frac{1}{\tau_2}}) \right) \int_0^\infty \int_0^\infty \text{Cov}[N_1(-s), N_1(-r)]
\]
\[
\left( \exp \left( \frac{1}{\tau_2} \right) * \mu \right) (ds) \left( \exp \left( \frac{1}{\tau_2} \right) * \mu \right) (dr)
\]
\[
= \lambda_1^+ \lambda_1 \lambda_2^2 \tau_1^2 \tau_2^2 e^{-\frac{1}{\tau_2}} \int_0^\infty \int_0^\infty e^{-\frac{1}{\tau_2}} \left( \exp \left( \frac{1}{\tau_2} \right) * \mu \right) (ds) \left( \exp \left( \frac{1}{\tau_2} \right) * \mu \right) (dr)
\]
\[
= \lambda_1^+ \lambda_1 \lambda_2^2 \tau_1^2 \tau_2^2 e^{-\frac{1}{\tau_2}} E \left[ e^{-\frac{(T_2 + G - T'_2 - G')}{\tau_1}} \right],
\]

\[
A_3 + A_5 = \lambda_2^2 \tau_2^2 \int_0^t \int_0^\infty 1_{r \leq t} \text{Cov}[N_1(-s), N_1(t - r - \delta)] \mu(d\delta) \exp \left( \frac{1}{\tau_2} \right) (dr)
\]
\[
\left( \exp \left( \frac{1}{\tau_2} \right) * \mu \right) (ds)
\]
\[
= \lambda_1^+ \lambda_1 \lambda_2^2 \tau_1^2 \tau_2^2 \int_0^t \int_0^\infty 1_{r \leq t} e^{-\frac{|r + s - r - \delta|}{\tau_1}} \mu(d\delta) \exp \left( \frac{1}{\tau_2} \right) (dr)
\]
\[
\left( \exp \left( \frac{1}{\tau_2} \right) * \mu \right) (ds)
\]
\[
= \lambda_1^+ \lambda_1 \lambda_2^2 \tau_1^2 \tau_2^2 E \left[ e^{-\frac{(T_2 + G - T'_2 - G')}{\tau_1}}, T_2 \leq t \right]
\]
\[
= \lambda_1^+ \lambda_1 \lambda_2^2 \tau_1^2 \tau_2^2 \left( E \left[ e^{-\frac{(T_2 + G - T'_2 - G')}{\tau_1}} \right] - E \left[ e^{-\frac{(T_2 + G - T'_2 - G')}{\tau_1}} \right] 1_{T_2 > t} e^{-\frac{1}{\tau_2}} \right)
\]
\[
= \lambda_1^+ \lambda_1 \lambda_2^2 \tau_1^2 \tau_2^2 \left( E \left[ e^{-\frac{(T_2 + G - T'_2 - G')}{\tau_1}} \right] - E \left[ e^{-\frac{(T_2 + G - T'_2 - G')}{\tau_1}} \right] e^{-\frac{1}{\tau_2}} \right)
\]
because, given $T'_2 > t$, the random variable $T'_2 - t$ is $\exp\left(\frac{1}{T'_2}\right)$-distributed. Altogether,

$$\text{Cov}[N_2(0), N_2(t)] = \lambda_t^+ \lambda_2 \tau_1 \tau_2 e^{-\frac{|t|}{T'_2}} + \lambda_t^+ \lambda_1 \lambda_2^2 \tau_1^2 \tau_2^2 \mathbb{E}\left[ e^{-\frac{|t|+T'_2+G-T'_2-G^t|}{\tau_1}} \right].$$

(13)

Now, (10) follows since by Lemma A.1,

$$\mathbb{E}\left[ e^{\frac{|G_2-G_2+G^t|}{\tau_1}} \right] = \mathbb{E}\left[ e^{\frac{|G_1-G_1+G^t|}{\tau_2}} \right] - \mathbb{E}\left[ e^{\frac{|G_1-G_1+G^t|}{\tau_3}} \right].$$

Therefore, using

$$\mathbb{E}\left[ e^{\frac{|G_2-G_2+G^t|}{\tau_1}} \right] = \frac{\tau_1}{(\tau_1 + \tau_2)(\tau_1 - \tau_2)} \left( \tau_1 \mathbb{E}\left[ e^{\frac{|G_2-G_2+G^t|}{\tau_1}} \right] - \tau_2 \mathbb{E}\left[ e^{\frac{|G_2-G_2+G^t|}{\tau_2}} \right] \right).$$

Theorem 3.4 (Expectation and variance of protein in equilibrium). The expectation and variance of $N_3$ in equilibrium are given by

$$\mathbb{E}[N_3(t)] = \lambda_t^+ \lambda_2 \lambda_3 \tau_1 \tau_2 \tau_3,$$

(14)

$$\text{Var}[N_3(t)] = \lambda_t^+ \lambda_2 \lambda_3 \tau_1 \tau_2 \tau_3 + \lambda_t^+ \lambda_2 \lambda_3^2 \tau_1 \tau_2 \tau_3^2 \cdot A + \lambda_t^+ \lambda_1 \lambda_2^2 \lambda_3^2 \tau_1^2 \tau_2^2 \tau_3^2 \cdot B,$$

(15)

$$A = \frac{\tau_2}{\tau_2 + \tau_3} \frac{\tau_2 - \tau_3}{\tau_2 - \tau_3},$$

(16)

$$B = \frac{\tau_1}{\tau_1 + \tau_3} \left( \frac{\tau_1^2}{\tau_1 + \tau_2} \frac{\tau_1}{\tau_1 - \tau_3} - \frac{\tau_2}{\tau_2 + \tau_3} \frac{\tau_2}{\tau_2 - \tau_3} \right) \frac{(\tau_1 - \tau_3)(\tau_1 - \tau_2)}{(\tau_1 + \tau_3)(\tau_1 - \tau_3)} - \frac{\tau_2^2}{\tau_2 + \tau_3} \frac{\tau_2}{\tau_2 - \tau_3} \frac{\tau_2}{\tau_2 - \tau_3} \frac{(\tau_1 - \tau_3)(\tau_1 - \tau_2)}{(\tau_1 - \tau_3)(\tau_1 - \tau_3)} + \tau_3 \mathbb{E}\left[ e^{\frac{|G_1-G_1+G^t|}{\tau_3}} \right].$$

(17)

where $G_2, G'_2, G_3, G'_3$ and independent and $G_2, G'_2 \sim \mu$ and $G_3, G'_3 \sim \nu$.

Remark 3.5 (Convexity and deterministic delay). We note that the last terms in (16) and (17) stay bounded, even for $\tau_1 \to \tau_2$, $\tau_2 \to \tau_3$ and $\tau_1 \to \tau_3$. In addition, $A, B \geq 0$ can easily be shown, which means that $\text{Var}[N_3(t)] \geq \mathbb{E}[N_3(t)]$ in all cases. Moreover, we will – as for the second moments of the amount of RNA – argue that $\text{Var}[N_3(t)]$ is largest in the absence of delay. First, recall from (12) that $A \leq 1$ with equality if and only if $\mu$ is a delta-measure (i.e. $\text{Var}[G_2] = 0$). Moreover, for a similar bound on $B$, assume without loss of generality that $\tau_1 > \tau_3$ and write, using again (12) and $S := |G_3 - G'_3 + G_2 - G'_2|

$$\frac{\tau_1 + \tau_3}{\tau_1} B = \frac{\tau_2^2 f_S(\tau_1, \tau_2)}{(\tau_1 + \tau_2)(\tau_1 - \tau_3)} - \frac{\tau_3^2 f_S(\tau_2, \tau_3)}{(\tau_2 + \tau_3)(\tau_1 - \tau_3)} \frac{\tau_2 f_S(\tau_1, \tau_2) + \tau_3 f_S(\tau_1, \tau_2) - \tau_1^2 f_S(\tau_2, \tau_3) - \tau_2^2 f_S(\tau_2, \tau_3)}{(\tau_1 + \tau_2)(\tau_2 + \tau_3)(\tau_1 - \tau_3)} + \frac{\tau_1 f_S(\tau_1, \tau_2) - \tau_3 f_S(\tau_3, \tau_2)}{\tau_1 + \tau_3}.$$
the convex combinations $1 = \frac{\tau_2}{\tau_1 - \tau_2} - \frac{\tau_1}{\tau_1 - \tau_3}$ and $1 = \frac{\tau_1}{\tau_1 - \tau_3} - \frac{\tau_2}{\tau_1 - \tau_2}$ (recall $\tau_1 > \tau_3$), we see that

\[
\frac{\tau_2 f_S(\tau_1, \tau_2) - \tau_3 f_S(\tau_3, \tau_2)}{\tau_1 - \tau_3} 
\leq f_S\left(\frac{\tau_1}{\tau_1 - \tau_3} - \frac{\tau_3}{\tau_1 - \tau_3}, \tau_2\right) = f_S\left(\frac{\tau_1^2 + \tau_1 \tau_3 + \tau_3^2}{\tau_1 + \tau_3}, \tau_2\right) \leq 1,
\]

with equality if and only if $\text{Var}[S] = 0$. In total, we see that

\[
\begin{align*}
\text{Var}[\mathcal{N}_3(t)] &\leq \lambda_1^+ \lambda_2 \lambda_3 \tau_1 \tau_2 \tau_3 + \lambda_1^+ \lambda_2 \lambda_3^2 \tau_1 \tau_2 \tau_3^2 \frac{\tau_2}{\tau_2 + \tau_3} \\
&\quad + \lambda_1^+ \lambda_2^2 \lambda_3 \tau_1^2 \tau_2^2 \tau_3 \frac{\tau_1}{\tau_1 + \tau_3} (\tau_2 + \tau_3)(\tau_1 + \tau_2)
\end{align*}
\]

with equality if $\mu$ and $\nu$ are delta-measures. In this case,

\[
\frac{\text{Var}[\mathcal{N}_3(t)]}{\text{E}[\mathcal{N}_3(t)]^2} = \frac{1}{\text{E}[\mathcal{N}_3(t)]} + \frac{1}{\lambda_1^+ \lambda_2 \tau_1 \tau_2 + \tau_3} + \frac{\lambda_1^+ \lambda_1^- \lambda_2 \lambda_3 \tau_1^2 \tau_2^2 \tau_3 \frac{\tau_1}{\tau_1 + \tau_3} (\tau_2 + \tau_3)(\tau_1 + \tau_2)}{\text{E}[\mathcal{N}_3(t)]},
\]

which is a well-known result; see equation (4) in Paulsson (2005).

**Proof.** Using (4) and Theorem 3.2,

\[
\text{E}[\mathcal{N}_3(t)] = \text{E}\left[\lambda_3 \tau_3 \int_0^\infty \mathcal{N}_2(t-s)\left(\exp\left(-\frac{1}{\tau_3}\right) * \nu\right)(ds)\right]
= \lambda_3 \tau_3 \lambda_2 \lambda_3 \tau_1 \tau_2.
\]

For the variance, we again use (4) and two independent exp($1/\tau_3$)-distributed random variables $T_3, T_\nu$ for

\[
\text{Var}[\mathcal{N}_3(t)] = \text{E}[\text{Var}[\mathcal{N}_3(t)|\mathcal{N}_1, \mathcal{N}_2]] + \text{Var}[\text{E}[\mathcal{N}_3(t)|\mathcal{N}_1, \mathcal{N}_2]],
\]

with

\[
\text{E}[\text{Var}[\mathcal{N}_3(t)|\mathcal{N}_1, \mathcal{N}_2]] = \text{E}\left[\lambda_3 \tau_3 \int_0^\infty \mathcal{N}_2(t-r)\left(\exp\left(-\frac{1}{\tau_3}\right) * \nu\right)(dr)\right]
= \lambda_1^+ \lambda_2 \lambda_3 \tau_1 \tau_2 \tau_3
\]
and with \( i \) in the third equality

\[
\text{Var} [\mathbb{E} [N_3(t)|N_1, N_2]] = \text{Var} \left[ \lambda_3 \tau_3 \int_0^\infty \mathcal{N}_2(t-r) \left( \exp \left( \frac{1}{\tau_3} \right) * \nu \right) (dr) \right]
\]

\[
= \lambda_3^2 \tau_3^2 \int_0^\infty \int_0^\infty \text{Cov} [\mathcal{N}_2(t-s), \mathcal{N}_2(t-r)]
\]

\[
\left( \exp \left( \frac{1}{\tau_3} \right) * \nu \right) (dr) \left( \exp \left( \frac{1}{\tau_3} \right) * \nu \right) (ds)
\]

\[
= \lambda^2 \lambda_2 \lambda_3 \tau_1 \tau_2 \tau_3^2 \int_0^\infty \int_0^\infty \exp \left( \frac{1}{\tau_3} \right) \left( \exp \left( \frac{1}{\tau_3} \right) * \nu \right) (dr) \left( \exp \left( \frac{1}{\tau_3} \right) * \nu \right) (ds)
\]

\[
+ \lambda^2 \lambda_2 \lambda_3 \tau_1 \tau_2 \tau_3^2 \left( \int_0^\infty \int_0^\infty \mathbb{E} \left[ e^{-\frac{|T_3 + G_3 - G_3'|}{\tau_3}} \right] \right) \left( \exp \left( \frac{1}{\tau_3} \right) * \nu \right) (ds)
\]

\[
= \lambda^2 \lambda_2 \lambda_3 \tau_1 \tau_2 \tau_3^2 \mathbb{E} \left[ e^{-\frac{|T_3 + G_3 - G_3'|}{\tau_3}} \right] + \lambda^2 \lambda_2 \lambda_3 \tau_1 \tau_2 \tau_3^2 \mathbb{E} \left[ e^{-\frac{|T_3 + G_3 - G_3'| + \tau_2 + G_2 - G_2'}{\tau_3}} \right].
\]

(18)

Now, by Lemma A.1 [21],

\[
\mathbb{E} \left[ e^{-\frac{T_3 + G_3 - G_3'}{\tau_3}} \right] = \frac{\tau_2}{\tau_2 + \tau_3} \frac{\tau_2}{\tau_2 - \tau_3} \mathbb{E} \left[ e^{-\frac{T_3 + G_3 - G_3'}{\tau_3}} \right] - \tau_3 \mathbb{E} \left[ e^{-\frac{|G_3 - G_3'|}{\tau_3}} \right]
\]

(19)

and by applying (22) twice (first for \( S = T_2 + G_2' - T_2 - G_2 \) and then for \( S = G_2 - G_2' + G_3 - G_3' \))

\[
\mathbb{E} \left[ e^{-\frac{|T_3 + G_3 - G_3'| + \tau_2 + G_2 - G_2'}{\tau_3}} \right] = \frac{\tau_1}{\tau_1 + \tau_3} \frac{\tau_1}{\tau_1 - \tau_3} \mathbb{E} \left[ e^{-\frac{|T_3 + G_3 - G_3'| + \tau_2 + G_2 - G_2'}{\tau_3}} \right] - \tau_3 \mathbb{E} \left[ e^{-\frac{|G_3 - G_3'|}{\tau_3}} \right]
\]

(20)

Hence, plugging (19) and (20) in (18), we get the result.

\[\boxed{}\]

4 Application: Small variance

As [Anderson and Kurtz (2011)] argue, a delay is the result of many mechanisms. Hence, by the central limit theorem, it is reasonable to assume that the delay distribution has a small variance. For this case, we obtain the following result.

**Corollary 4.1.** Let \( G_2, G_2' \sim \mu \) and \( G_3, G_3' \sim \nu \) are independent, and such that \( \text{Var}[G_2], \text{Var}[G_3] \) is small. Then, if \( \delta_2 := \mathbb{E} [\left| G_2 - \mathbb{E} [G_2] \right|^3], \quad \delta_3 := \mathbb{E} [\left| G_3 - \mathbb{E} [G_3] \right|^3], \)

\[\boxed{}\]
it holds for $\delta_2, \delta_3 \to 0$ that

\[
\begin{align*}
\Var[\mathcal{N}_2(t)] &= \lambda_1^+ \lambda_2 \tau_1 \tau_2 + \lambda_1^+ \lambda_2 \lambda_2 \tilde{\tau}_1 \tilde{\tau}_2 \frac{1}{\tau_1 + \tau_2} \left( 1 - \frac{2\Var[G_2]}{\tau_1 \tau_2} \right) + O(\delta_2), \\
\Var[\mathcal{N}_3(t)] &= \lambda_1^+ \lambda_2 \tau_1 \tau_2 \tau_3 + \lambda_1^+ \lambda_2 \lambda_2 \frac{\tau_1 \tau_2 \tau_3}{\tau_2 + \tau_3} \left( 1 - \frac{2\Var[G_3]}{\tau_2 \tau_3} \right) \\
&\quad + \lambda_1^+ \lambda_1^+ \lambda_2 \lambda_2^2 \frac{\tau_1 \tau_2 \tau_3}{(\tau_1 + \tau_2)(\tau_1 + \tau_3)(\tau_2 + \tau_3)} - \Var[G_2 + G_3] + O(\delta_2, \delta_3).
\end{align*}
\]

Proof. We write

\[
E\left[ e^{-\frac{G_2'-G_2}{\tau}} \right] = E\left[ 1 - \frac{|G_2' - G_2|}{\tau} + \frac{(G_2' - G_2)^2}{\tau^2} + O\left( |G_2' - G_2|^3 \right) \right]
\]

\[
= E\left[ 1 - \frac{|G_2' - G_2|}{\tau} + \frac{2(\Var[G_2] + \Var[G_3])}{\tau^2} \right] + O(\delta_2, \delta_3).
\]

Hence, the first result concerning $\Var[\mathcal{N}_2(t)]$ follows directly from Theorem 3.2 For the variance of proteins,

\[
E\left[ e^{-\frac{G_3'-G_3+G_2-G_3'}{\tau}} \right] = E\left[ 1 - \frac{|G_3 - G_3' + G_2 - G_3'|}{\tau} + \frac{(G_3 - G_3' + G_2 - G_3')^2}{\tau^2} \right]
\]

\[
= E\left[ 1 - \frac{|G_3 - G_3' + G_2 - G_3'|}{\tau} + \frac{2(\Var[G_2] + \Var[G_3])}{\tau} \right] + O(\delta_2, \delta_3).
\]

So, denoting the values of $A$ and $B$ from (17) with deterministic $G_2$ and $G_3$ by $A_0$ and $B_0$, respectively (compare with Remark 3.3, (17) gives

\[
A_0 - A = \frac{2\Var[G_3]}{\tau_2 + \tau_3} \tau_3,
\]

\[
B_0 - B = \frac{\tau_1}{(\tau_1 + \tau_3)(\tau_1 + \tau_3)} \left( \frac{2\Var[G_2 + G_3] \tau_1}{(\tau_1 + \tau_2) \tau_2} - \frac{2\Var[G_2 + G_3] \tau_3}{(\tau_2 + \tau_3) \tau_2} \right) + O(\delta_2, \delta_3)
\]

\[
= \frac{2\Var[G_2 + G_3] \tau_1}{(\tau_1 + \tau_2)(\tau_1 + \tau_3)(\tau_2 + \tau_3)} + O(\delta_2, \delta_3).
\]

Hence, if $\Var$ is the variance of $\mathcal{N}_3(t)$ for deterministic $G_2, G_3$,

\[
\Var[\mathcal{N}_3(t)] = \Var - 2 \left( \lambda_1^+ \lambda_2 \lambda_2 \frac{\Var[G_3] \tau_1}{\tau_2 + \tau_3} \right)
\]

\[
+ \lambda_1^+ \lambda_1^+ \lambda_2 \lambda_2^2 \frac{\Var[G_2 + G_3] \tau_1}{(\tau_1 + \tau_2)(\tau_1 + \tau_3)(\tau_2 + \tau_3)} + O(\delta_2, \delta_3)
\]

and the result follows. \( \square \)

A A key lemma

Within this section, we summarize some frequently used computations in the following lemma.

**Lemma A.1.** Let $T, T'$ be independent exponentially distributed with expectation $\sigma$, and $G, G' \geq 0$ be independent and identically distributed. Then, for $t \in \mathbb{R}, \tau > 0$

\[
E\left[ \frac{e^{-|t+T+G-t'-G'|}}{\tau} \right] = \frac{\tau}{\tau + \sigma} \frac{\tau E\left[ e^{-\frac{|G'-G-\tau|}{\tau}} \right]}{\tau - \sigma} - \sigma E\left[ e^{-\frac{|G'-G-\tau|}{\tau}} \right].
\]
Moreover, if $S$ has a symmetric distribution, i.e. $S \overset{d}{=} -S$,
\[
E\left[e^{-\frac{|T + G - T' - G'|}{\sigma}}\right] = E\left[e^{-\frac{|T + G - T' - G'|}{\sigma}}\right] = \frac{\tau}{\tau + \sigma} E\left[e^{-\frac{|G' - G|}{\sigma}}\right] - \sigma E\left[e^{-\frac{|G' - G|}{\sigma}}\right].
\]

**Remark A.2.** Note that in (21) and (22) all left hand sides include the exponentially distributed random variables $T, T'$, while the right hand sides only depend on the distribution of $G, G'$ (and $S$).

**Proof.** We start with the proof of (21). First,
\[
E\left[e^{-\frac{t+T + G - T' - G'}{\sigma}}\right] = E\left[E\left[e^{-\frac{t+T + G - T' - G'}{\sigma}}1_{T > T' + G' - G - t}\right|G - G']\right] + E\left[E\left[e^{-\frac{t+T + G - T' - G'}{\sigma}}1_{T > T' + G' - G - t}\right|G - G']\right]
\]

Then,
\[
E\left[e^{-\frac{t+T + G - T' - G'}{\sigma}}1_{T > T' + G' - G - t}\right|G - G'] = 1_{G' - G - t \geq 0} E\left[e^{-\frac{t+T + G - T' - G'}{\sigma}}1_{T > T' + G' - G - t}\right|G - G']
+ 1_{G' - G - t < 0} E\left[e^{-\frac{t+T + G - T' - G'}{\sigma}}\left(1_{T + G'} < G' + t + 1_{T > T' + G' - G - t}\right)\right|G - G']
\]

Now, on the set \{G' - G - t \geq 0\},
\[
E\left[e^{-\frac{t+T + G - T' - G'}{\sigma}}1_{T > T' + G' - G - t}\right|G - G'] = \frac{1}{2} e^{-\frac{G' - G}{\sigma}} E\left[e^{-\frac{t}{\tau}}\right] = \frac{1}{2 \tau} e^{-\frac{G' - G}{\sigma} t}
\]

since $P[T > T' + G' - G - t|G - G'] = \frac{1}{2} e^{-\frac{G' - G}{\sigma}}$ and given $T > T' + G' - G - t$ and conditioned on $G - G'$, the random variable $T + G + T' - G'$ is again exponentially distributed with expectation $\sigma$. Next, on $\{G' - G - t < 0\}$
\[
E\left[e^{-\frac{t+T + G - T' - G'}{\sigma}}1_{T' < G - G + t}\right|G - G'] = \int_0^\infty \int_0^{G' + t} \frac{1}{\sigma^2} e^{-\frac{t+T + G - T' - G'}{\sigma}} ds' ds
\]
\[
= \frac{1}{\sigma^2} e^{-\frac{G' - G + t}{\sigma}} \int_0^\infty e^{-s(s + \frac{t}{\tau})} ds \int_0^{G' + t} e^{-s'(s' + \frac{t}{\tau})} ds' = \frac{1}{\tau^2} \left(e^{-\frac{G' - G + t}{\sigma}} - e^{-\frac{G' - G + t}{\sigma}}\right)
\]

as well as
\[
E\left[e^{-\frac{t+T + G - T' - G'}{\sigma}}1_{T' > G - G + t}\right|G - G'] = e^{-\frac{G' - G'}{\sigma}} E\left[e^{-\frac{t}{\tau}}1_{T > T'}\right|G - G'] = \frac{1}{2} e^{-\frac{G' - G'}{\sigma}} \frac{\tau}{\tau + \sigma}
\]
since $P[T' > G - G' + t | G - G'] = e^{-\frac{t + G - G'}{\sigma}}$ and given $T' > G - G' + t$, the random variable $T' + G' - G - t$ is again $\exp(\sigma)$ distributed. Plugging the last three computations into (21),

$$e^{\frac{t + G - G'}{\sigma}} E[T > T' + G' - G - t] G - G'] = \frac{1}{2} \tau e^{-\frac{|G' - G|}{\sigma}} + G - G' \leq 0 \frac{\tau^2}{(\tau + \sigma)(\tau - \sigma)} \left( e^{-\frac{G - G' + t}{\tau}} - e^{-\frac{G - G' - t}{\sigma}} \right).$$

Then, by symmetry of $G - G'$, from (21),

$$E[e^{\frac{t + G - G'}{\sigma} T'}] = \frac{1}{2} \tau e^{-\frac{|G' - G|}{\sigma}} + G - G' \leq 0 \frac{\tau^2}{(\tau + \sigma)(\tau - \sigma)} \left( e^{-\frac{G - G' + t}{\tau}} - e^{-\frac{G - G' - t}{\sigma}} \right) + E\left[ \frac{1}{2} \tau e^{-\frac{|G' - G'|}{\sigma}} + G - G' \leq 0 \frac{\tau^2}{(\tau + \sigma)(\tau - \sigma)} \left( e^{-\frac{G - G' + t}{\tau}} - e^{-\frac{G - G' - t}{\sigma}} \right) \right]

= \frac{\tau}{\tau + \sigma} E[e^{-\frac{|G' - G|}{\sigma}}] + \frac{\tau^2}{(\tau + \sigma)(\tau - \sigma)} E[e^{-\frac{|G' - G|}{\tau}} - e^{-\frac{|G' - G|}{\sigma}}]

= \frac{\tau}{\tau + \sigma} \sigma E[e^{-\frac{|G' - G|}{\tau}}] - \sigma E[e^{-\frac{|G' - G|}{\sigma}}].$$

For (22), we only need to show the first equality since the second follows from (21) by conditioning on $S$. Using the symmetry of $S$ we can write

$$E[e^{-\frac{|T + G - T' + G'|}{\tau}}] = E\left[ e^{-\frac{|T + G - T' + G'|}{\tau}} | T + G - T' - G' \geq 0 \right] + E\left[ e^{-\frac{|T + G - T' + G'|}{\tau}} | T + G - T' - G' < 0 \right]

= E[e^{-\frac{|T + G - T' + G'|}{\tau}}]$$

which shows the first equality in (22).

\[\square\]

**Acknowledgments**

We thank Bence Melykuti for helpful comments on the manuscript.

**References**

D. Anderson. A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *J. Chem. Phys.* 127 (21), 214107–214107 (2007).

D. Anderson and T. G. Kurtz. Continuous time Markov chain models for chemical reaction networks. In *Design and Analysis of Biomolecular Circuits: Engineering Approaches to Systems and Synthetic Biology*. Springer (2011).

M. Barrio, K. Burrage, A. Leier and T T. Tian. Oscillatory regulation of hes1: Discrete stochastic delay modelling and simulation. *PLoS Comput. Biol.* 2 (9) (2006).

M. Barrio, A. Leier and T. T. Marquez-Lago. Reduction of chemical reaction networks through delay distributions. *J. Chem. Phys.* 138 (10), 104114–104114 (2013).

O. G. Berg. A model for statistical fluctuations of protein numbers in a microbial-population. *J. Theor. Biol.* 71, 587–603 (1978).

D. Bratsun, D. Volfson, L. S. Tsimring and J. Hasty. Delay-induced stochastic oscillations in gene regulation. *Proc. Natl. Acad. Sci. USA* 102 (41), 14593–14598 (2005).
X. Cai. Exact stochastic simulation of coupled chemical reactions with delays. *J. Chem. Phys.* **126** (12), 124108–124108 (2007).

D. Cottrell, P. S. Swain and P. F. Tupper. Stochastic branching-diffusion models for gene expression. *Proc. Natl. Acad. Sci. USA* **109** (25), 9699–9704 (2012).

D. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361 (1977).

D. Gillespie, A. Hellander and L. Petzold. Perspective: Stochastic algorithms for chemical kinetics. *J. Chem. Phys.* **138**, 170901 (2013).

S. Iyer-Biswas, F. Hayot and C. Jayaprakash. Stochasticity of gene products from transcriptional pulsing. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **79** (3 Pt 1), 031911–031911 (2009).

M. Kaern, T. C. Elston, W. J. Blake and J. J. Collins. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.* **6** (6), 451–464 (2005).

J. Lewis. Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. *Curr. Biol.* **13** (16), 1398–1408 (2003).

H. H. McAdams and A. Arkin. Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA* **94** (3), 814–819 (1997).

N. A. Monk. Oscillatory expression of Hes1, p53, and NF-kappaB driven by transcriptional time delays. *Curr. Biol.* **13** (16), 1409–1413 (2003).

J. Paulsson. Models of stochastic gene expression. *Physics of Life Review* **2**, 157–175 (2005).

A. S. Ribeiro. Stochastic and delayed stochastic models of gene expression and regulation. *Math. Biosci.* **223** (1), 1–11 (2010).

D. R. Rigney and W. C. Schieve. Stochastic model of linear, continuous protein synthesis in bacterial populations. *J. Theor. Biol.* **69** (4), 761–766 (1977).

M R Roussel. The use of delay differential equations in chemical kinetics. *J. Phys. Chem.* **100**, 8323–8330 (1996).

M. R. Roussel and R. Zhu. Stochastic kinetics description of a simple transcription model. *Bull. Math. Biol.* **68** (7), 1681–1713 (2006a).

M. R. Roussel and R. Zhu. Validation of an algorithm for delay stochastic simulation of transcription and translation in prokaryotic gene expression. *Phys. Biol.* **3** (4), 274–284 (2006b).

P. S. Swain, M. B. Elowitz and E. D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. USA* **99** (20), 12795–12800 (2002).

T. Tian. Chemical memory reactions induced bursting dynamics in gene expression. *PLoS One* **8** (1) (2013).

T. Tian, K. Burrage, P. M. Burrage and M. Carletti. Stochastic delay differential equations for genetic regulatory networks. *Journal of Computational and Applied Mathematics* **205**, 696–707 (2007).

R. Zhu, A. S. Ribeiro, D. Salahub and S. A. Kauffman. Studying genetic regulatory networks at the molecular level: delayed reaction stochastic models. *J. Theor. Biol.* **246** (4), 725–745 (2007).

R. Zhu and D. Salahub. Delay stochastic simulation of single-gene expression reveals a detailed relationship between protein noise and mean abundance. *FEBS Lett.* **582** (19), 2905–2910 (2008).