

# Translational Cross Talk in Gene Networks

William H. Mather,<sup>†</sup> Jeff Hasty,<sup>‡§¶||</sup> Lev S. Tsimring,<sup>‡||</sup> and Ruth J. Williams<sup>‡\*\*\*</sup>

<sup>†</sup>Department of Physics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia; <sup>‡</sup>Biocircuits Institute, University of California, San Diego, La Jolla, California; <sup>§</sup>Department of Bioengineering, University of California, San Diego, La Jolla, California; <sup>¶</sup>Molecular Biology Section, Division of Biological Sciences, University of California, San Diego, La Jolla, California; <sup>||</sup>San Diego Center for Systems Biology, La Jolla, California; and <sup>\*\*\*</sup>Department of Mathematics, University of California, San Diego, La Jolla, California

**ABSTRACT** It has been shown experimentally that competition for limited translational resources by upstream mRNAs can lead to an anticorrelation between protein counts. Here, we investigate a stochastic model for this phenomenon, in which gene transcripts of different types compete for a finite pool of ribosomes. Throughout, we utilize concepts from the theory of multiclass queues to describe a qualitative shift in protein count statistics as the system transitions from being underloaded (ribosomes exceed transcripts in number) to being overloaded (transcripts exceed ribosomes in number). The exact analytical solution of a simplified stochastic model, in which the numbers of competing mRNAs and ribosomes are fixed, exhibits weak positive correlations between steady-state protein counts when total transcript count slightly exceeds ribosome count, whereas the solution can exhibit strong negative correlations when total transcript count significantly exceeds ribosome count. Extending this analysis, we find approximate but reasonably accurate solutions for a more realistic model, in which abundances of mRNAs and ribosomes are allowed to fluctuate randomly. Here, ribosomal fluctuations contribute positively and mRNA fluctuations contribute negatively to correlations, and when mRNA fluctuations dominate ribosomal fluctuations, a strong anticorrelation extremum reliably occurs near the transition from the underloaded to the overloaded regime.

## INTRODUCTION

Biological cells are often forced to cope with limited resources. Their survival hinges on robust strategies that regulate the allocation of these resources to the cell's myriad processes (1). On the other hand, cells sometimes use resource distribution itself as a global regulatory mechanism that controls cooperation among various metabolic or signaling pathways. Understanding the specific mechanisms of this regulation is far from complete (2,3). This is also an important issue for the emerging field of synthetic biology, where the forward engineering of complex systems from more basic modules depends critically on mitigating cross talk between modules (4–10).

A recurrent source of potential cross talk is the competition for the processing resources that control the production, degradation, and modification of proteins. It is known, for example, that  $\sigma$ -factor competition for a finite pool of RNA polymerases leads to significant changes in RNA polymerase partitioning between transcription of housekeeping genes and that of stress response genes under stress conditions (2). Molecular competition can lead to unexpected consequences, including retroactivity that influences upstream processes (11–15), in addition to the more-often-discussed downstream coupling (16–18). We will focus here on another fundamental source of cross talk: competition between mRNA molecules for translational processing resources, here referred to simply as ribosomes. There are several lines of evidence suggesting that the copy number of ribosomes is limiting to protein synthesis at the whole-

cell level (19–25), in which case competition for a common pool of ribosomes by mRNAs can lead to cross talk when large systematic changes in transcript abundance occur. A more surprising finding is that pronounced and functionally important cross talk has been shown to arise from competition between a small number of different transcripts in the galactose utilization network of *Saccharomyces cerevisiae* (26,27), apparently caused by transcripts competing for a localized pool of ribosomes.

Due to the widespread and fundamental nature of translational competition, we seek to develop a theory that provides an understanding of cross talk due to a bottleneck in ribosomal processing. Recently, queueing theory (28), which originally was developed for engineering applications in optimizing telephone and manufacturing systems, as well as human and computer networks, has emerged as a useful tool for the description of a variety of cellular processes (13,29–35). Queueing theory deals with the processing of jobs by limited resources, and it typically classifies system behavior according to whether the system is underloaded (processing resources are more than sufficient to process all arriving jobs) or overloaded (processing resources are not sufficient to process all arriving jobs), with balance (processing resources are critically loaded) being the boundary between the underloaded and overloaded regimes. Furthermore, the theory is equipped to handle the stochastic fluctuations that are naturally present within living cells. In this article, we use concepts from queueing theory to describe qualitatively different behavioral regimes for our model of translational cross talk. We will specifically focus on what are termed multiclass queues (36), where multiple different types of units are processed by shared resources.

Submitted December 26, 2012, and accepted for publication April 11, 2013.

\*Correspondence: [williams@stochastic.ucsd.edu](mailto:williams@stochastic.ucsd.edu)

Editor: H. Wiley.

© 2013 by the Biophysical Society  
0006-3495/13/06/2564/9 \$2.00

<http://dx.doi.org/10.1016/j.bpj.2013.04.049>



In this article, we present and analyze a stochastic molecular model for translational cross talk. A particularly intriguing result is found when mRNA fluctuations dominate ribosomal fluctuations: a negative correlation extremum occurs as the system transitions from being underloaded (ribosomes exceed transcripts in number) to being overloaded (transcripts exceed ribosomes in number). We refer to this as a (negative) correlation resonance. A similar correlation resonance has been observed in certain other biologically relevant multiclass queueing networks (33), though the correlation extremum in that case was positive. As in that work, the correlation resonance observed in this study appears robust, and we expect that it is a feature that could be observed experimentally. Our results are applied to explore indirect corepression between gene activities, where induction of one gene leads to an effective repression of other genes.

Although we focus in this article on a model for translational cross talk, our analysis is based on a Markov chain model for protein production. This is a quite general formulation that has potential applications in other contexts, such as transcriptional cross talk as a result of competition for transcriptional machinery (e.g., RNA polymerase), competition between  $\sigma$ -factors, and certain types of enzymatic protein processing.

An outline of the article is as follows. The second section presents the biochemical reactions that define our basic stochastic model, in which the mRNA and ribosome counts are fixed. The third section contains analytic results for the first- and second-order statistics of steady-state protein counts of this model, including the Pearson correlation coefficient (henceforth referred to simply as correlation). Leveraging these results and using a quasi-steady-state approximation, the fourth section presents results for steady-state protein statistics in the generalized model where mRNAs and ribosomes are allowed to fluctuate slowly. A negative correlation resonance reliably occurs near the point where the system transitions from the underloaded to the overloaded regime where mRNA fluctuations dominate ribosomal fluctuations. These results are used to explore indirect repression of gene activity. Concluding remarks are given in the Discussion section.

### BASIC STOCHASTIC MODEL FOR TRANSLATIONAL CROSS TALK WITH mRNA AND RIBOSOME NUMBERS FIXED

We consider a stochastic model for translational cross talk in which there are two different types of mRNA and a limited set of identical ribosomes (see Fig. 1). In this section and the next, we introduce and analyze our model with fixed numbers of copies of each of the mRNA transcripts and the ribosomes. In later sections, we expand on this, allowing the mRNA and ribosome copy numbers to fluctuate.

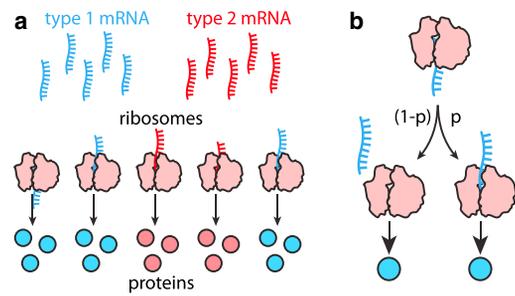


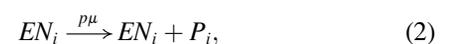
FIGURE 1 (a) Schematic of our stochastic model for translational cross talk. A finite number of copies of two types of mRNA can either be bound or unbound to a finite pool of identical ribosomes, with at most one ribosome binding to a given mRNA molecule. Once bound to an mRNA molecule, a ribosome begins translating the transcript. A single protein of the same type as the mRNA is produced upon completion of the translation process. (b) When a ribosome completes the translation process it either rebinds to the same transcript (with probability  $p$ ), or (with probability  $1 - p$ ) releases the transcript into the pool of unbound transcripts and selects an mRNA transcript at random from this pool and binds to it.

We suppose that there are two types of (free) mRNA, denoted by  $N_1$  and  $N_2$ , as well as a finite number of copies of available ribosomes,  $E$ . A molecule of the mRNA  $N_i$  can bind with rate  $\nu$  to a free ribosome to form the complex  $EN_i$  according to the reaction:

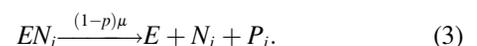


For our analysis, we assume fast binding; we assume that the binding rate,  $\nu$ , is effectively infinite, so that when a ribosome becomes free, it selects a transcript at random from the pool of unbound mRNA transcripts and then binds to it instantly. (In simulations that we do for comparison with our analytical results, we use exponentially distributed binding reaction times with a large but finite rate  $\nu$ .) The form of the reaction above implicitly assumes that at most one ribosome can bind to an mRNA at a time; this simplification is consistent with our primary focus on regimes where the number of ribosomes is limited relative to the supply of mRNA molecules.

Once the complex  $EN_i$  is formed, it produces a protein,  $P_i$ , in an exponentially distributed amount of time with mean  $1/\mu$  for fixed  $\mu > 0$ . We allow for a finite probability,  $p \in [0, 1)$ , that upon protein production the same ribosome will rebind to the same transcript. Otherwise, the ribosome dissociates (with probability  $1 - p$ ), releasing the mRNA into the pool of unbound mRNAs, and the ribosome becomes free to bind a transcript from this pool. The accompanying reactions are given by



and



Once produced, each copy of protein  $P_i$  has an exponentially distributed lifetime with a mean of  $1/\gamma$  for fixed  $\gamma > 0$ . This is represented by the degradation reaction



Here, in the main text, we assume a common degradation rate,  $\gamma$ , for both types of protein; however, it is straightforward to allow the proteins,  $P_i$ , to have their own individual degradation rates,  $\gamma_i$  (see the [Supporting Material](#)). All of the (exponential) reaction times in our stochastic model are assumed to be mutually independent.

It should be noted that the case with rebinding of ribosomes to mRNA ( $p > 0$ ) may be more or less relevant depending on the biochemical details of translation for a specific organism. Translation in eukaryotes, in particular, may exhibit rebinding, since the use of circularized RNA may increase the likelihood for the reinitiation of ribosomes (37).

In this section and the next, we assume that the total number of molecules of each type of mRNA (bound plus unbound), and the total number of copies of the ribosome (bound plus unbound), are fixed constants. We denote the total number of mRNAs of type  $i$  by  $T_i$  and the total number of ribosomes by  $R$ , so that

$$\text{count}(N_i + EN_i) = T_i, \quad (5)$$

and

$$\text{count}\left(E + \sum_{i=1}^2 EN_i\right) = R. \quad (6)$$

We define  $Y_1$  and  $Y_2$  as stochastic processes that track the numbers of molecules of the two types of protein,  $P_1$  and  $P_2$ , respectively. The steady-state moments for this pair are the focus of the next section.

### STEADY-STATE MEANS, VARIANCES, AND CORRELATION FOR PROTEIN COUNTS WITH FIXED mRNA AND RIBOSOME NUMBERS

In this section, we present exact formulas for the steady-state means, variances, covariance, and correlation of the protein counts associated with the stochastic model introduced in the previous section. The derivation of some of these formulas is explained here, and the rest of the formulas are derived in the [Supporting Material](#). In fact, the methodology we use in the [Supporting Material](#) also allows for the computation of time-dependent moments.

From our analysis, two qualitatively different regimes are apparent, depending on whether or not the total number of mRNA transcripts exceeds the number of ribosomes. The more traditional case studied in the literature is where

$T_1 + T_2 \leq R$  so that transcripts do not compete with one another for ribosomes. We call this the underloaded regime. (Usually,  $T_1 + T_2 < R$  defines the underloaded regime; however, for convenience of terminology, here we include the borderline case of balance, or critical loading, where  $T_1 + T_2 = R$ , when we reference the underloaded regime.) In this regime, under our instantaneous binding assumption ( $\nu = \infty$ ), at any time, each mRNA molecule is bound to a ribosome. Thus, in terms of protein production, the mRNAs can be treated independently of one another, where for a given mRNA transcript, a new protein is produced from it at each of the jump times of a Poisson process with rate  $\mu$ . Each of the proteins produced from a given mRNA transcript in turn independently degrades after an exponentially distributed amount of time with mean  $1/\gamma$ . Then, as in a Markovian infinite server queue (28), the steady-state number of proteins that originated from a given mRNA transcript is Poisson distributed with mean  $\mu/\gamma$ . Consequently, in the steady state,  $Y_i$  is the sum of  $T_i$  independent Poisson random variables, each of which has mean  $\mu/\gamma$ . Thus,  $Y_1$  and  $Y_2$  are independent Poisson random variables at steady state, and the steady-state mean  $\langle Y_i \rangle$ , variance  $\sigma_{Y_i}^2$ , and covariance  $\sigma_{Y_1, Y_2}^2$  for the protein counts are given for  $i = 1, 2$ , by

$$\langle Y_i \rangle = \frac{\mu T_i}{\gamma}, \quad (7)$$

$$\sigma_{Y_i}^2 = \langle Y_i \rangle, \quad (8)$$

and

$$\sigma_{Y_1, Y_2}^2 = 0. \quad (9)$$

The rebinding probability,  $p$ , does not enter here, since after being translated by the ribosome, the transcript instantly rebinds either to the same ribosome or to another ribosome (if there is one) and so is effectively always bound to a ribosome.

The qualitatively different and more interesting regime is what we term the overloaded regime, where  $T_1 + T_2 > R$  and competition for translational processing resources can be pronounced. In this case, under our instantaneous binding assumption ( $\nu = \infty$ ), we see that the one-dimensional stochastic process  $X$  that keeps track of the number of mRNAs of type 1 that are currently bound to ribosomes (the number of molecules of  $EN_1$ ) is a finite-state continuous-time Markov chain. This Markov chain drives the production of the two types of protein, and the combined process  $(X, Y_1, Y_2)$  is a continuous-time Markov chain with countably many states. Although the protein count processes  $Y_1, Y_2$  can take arbitrarily large values, the proteins collectively degrade at a rate that is proportional to their number. Consequently, it is straightforward to verify that the three-dimensional Markov chain  $(X, Y_1, Y_2)$  is irreducible and positive

recurrent, and so it has a unique steady-state distribution. Formulas for the steady-state means, variances, and covariance are derived in the [Supporting Material](#). The mean protein count in this case is a function of both mRNA counts

$$\langle Y_i \rangle = \frac{\mu R T_i}{\gamma(T_1 + T_2)}, \quad (10)$$

so that, for example, increasing the level of type 2 mRNA can decrease the expression level of type 1 proteins. This indirect repression allows for cross talk between protein species, similar to what has been seen experimentally (27). We have derived expressions for  $\sigma_{Y_i}^2$  and  $\sigma_{Y_1, Y_2}^2$  for general  $T_1$  and  $T_2$  in the overloaded case (see [Supporting Material](#)), but many features of these statistics are represented by the symmetric case  $T_1 = T_2 = T_s$ , which leads to the expressions

$$\langle Y_1 \rangle = \langle Y_2 \rangle = \frac{\mu R}{2\gamma}, \quad (11)$$

$$\sigma_{Y_1}^2 = \sigma_{Y_2}^2 = \frac{\mu R}{4\gamma(2T_s - 1)T_s[2(1 - p)\mu T_s + (2T_s - R + 1)\gamma]} \times [(R(R - 1) - (2R + 2)T_s + 8T_s^2 - 2T_s(2T_s - 2 + R)p)\mu T_s + 2(2T_s - 1)T_s(2T_s - R + 1)\gamma], \quad (12)$$

and

$$\sigma_{Y_1, Y_2}^2 = \frac{\mu^2 R(2T_s - R)(R - 2pT_s - 1)}{4\gamma(2T_s - 1)[2(1 - p)\mu T_s + (2T_s - R + 1)\gamma]}, \quad (13)$$

where  $2T_s > R$  by the overloaded condition. The correlation  $\mathcal{C}(Y_1, Y_2)$  is given by

$$\mathcal{C}(Y_1, Y_2) = \frac{\sigma_{Y_1, Y_2}^2}{\sqrt{\sigma_{Y_1}^2 \sigma_{Y_2}^2}}. \quad (14)$$

The covariance in [Eq. 13](#) is especially interesting, since its sign provides the sign of the correlation. Since  $2T_s > R$ , the correlation has the same sign as  $R - p(T_1 + T_2) - 1$ , which happens to hold true also when  $T_1 \neq T_2$  (see the [Supporting Material](#)). In particular, this correlation is positive if  $p = 0$  (no rebinding) and is negative in the limit  $p \rightarrow 1$  with other parameters fixed (the always rebind limit). Sample results for the correlation, produced from our analytic results, are illustrated in [Fig. 2](#), where it is also seen that asymmetry in  $T_1$  and  $T_2$  for fixed  $2T_s = T_1 + T_2$  tends to reduce the magnitude of the correlation.

A special limit for our system is that of infinite mRNA count ( $T_s \rightarrow \infty$ ) with fixed total count of ribosomes,  $R$ , and proportions of each type of mRNA. In this case, copies of the two types of mRNA are so abundant that the resulting

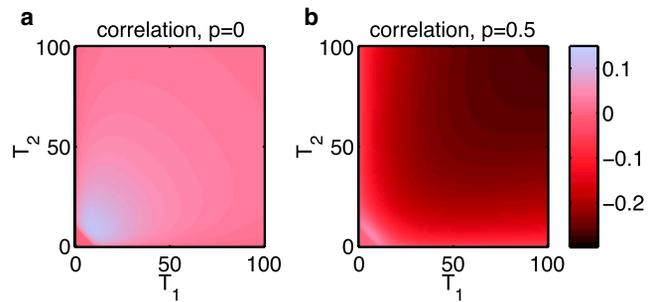


FIGURE 2 Plot of the steady-state correlation between protein counts for fixed numbers of mRNA (count for total mRNA of type  $i$ ,  $T_i$ ) and ribosomes (total count,  $R$ ). (a) The correlation is nonnegative when rebinding is absent ( $p = 0$ ). (b) However, a finite probability of rebinding ( $p = 0.5$ ) can lead to negative correlations for sufficiently large values of mRNA counts. The correlation is zero in the underloaded case ( $T_1 + T_2 \leq R$ , bottom left corners of plots). In both *a* and *b*, the correlation has a maximum along the line  $T_1 = T_2 = T_s$ , and this peak occurs slightly above the balance point (roughly in the range  $R < 2T_s < 2R$ ). Beyond the peak, along this line, the correlation decreases monotonically toward an asymptotic value which is given in [Eq. 15](#); this value is zero for  $p = 0$  and strictly negative for  $p > 0$ . Other parameters are  $R = 10$ ,  $\mu = 1$ , and  $\gamma = 0.1$ .

steady-state statistics for the protein counts are the same as those for an ensemble of independently operating ribosomes drawing from a pool with fixed proportions of mRNA of each type. This limiting model can then in principle be analyzed using known methods for first-order systems (38). The asymptotic protein correlation can be readily computed by passing to the limit in our formulas; for example, using [Eqs. 12–14](#) above, we have in the symmetric case

$$\lim_{T_s \rightarrow \infty} \mathcal{C}(Y_1, Y_2) = -\frac{p}{2 - p + 2(\gamma/\mu)}. \quad (15)$$

This expression is precisely the same as the protein correlation due to a single ribosome (see [Eqs. 12–14](#) when  $R = 1$ ). This is to be expected, since in the limit  $T_s \rightarrow \infty$ , each ribosome is effectively drawing from an infinite pool of mRNAs with fixed mRNA proportions, and its production of the two types of proteins will be independent of the production of the other ribosomes. Thus, each ribosome will produce proteins in a manner that is independently but identically distributed relative to the other ribosomes, and consequently, the associated asymptotic correlation will be the same for finitely many ribosomes as for a single ribosome. We further note that the asymptotic correlation in [Eq. 15](#) is zero when rebinding is absent ( $p = 0$ ), as can be understood by noticing that the protein production processes for a single ribosome without rebinding are equivalent in distribution to two independent Poisson processes with rates  $\mu T_i / 2T_s$ ,  $i = 1, 2$ .

It is a curious fact that competition for processing resources produces consistently positive steady-state correlations in the absence of rebinding ( $p = 0$ ) for  $R > 1$ , although these positive correlations are relatively small,

i.e., less than or equal to  $1/7$ . (For an intuitive explanation of the positive correlations, see the [Supporting Material](#)). The maximum for the correlation expression given by Eqs. 12–14 (the symmetric case) can be established by sequentially differentiating  $\mathcal{C}(Y_1, Y_2)$  with respect to certain single parameters and progressively finding maxima with respect to these parameters. This process reveals that the correlation depends monotonically on  $p$  and  $\gamma$ , and that maximum correlation requires  $p = 0$  and  $\gamma \rightarrow 0$ . Assuming the limits  $p = 0$  and  $\gamma \rightarrow 0$ , and allowing  $T_s$  to be real-valued, there exists a particular value of  $T_s$  given by

$$T_s^* = \frac{1}{2} \left( R + \sqrt{R(R-1)} \right), \quad (16)$$

where the correlation is maximal. This is approximately  $T_s^* \approx R$  for large  $R$ , so maximum correlation in this case occurs near the point where the total mRNA count is twice that of ribosomes, which can be compared to the balance point  $T_s = R/2$ . The maximum correlation at  $T_s^*$  can be shown to approach  $1/7$  from below as  $R \rightarrow \infty$ . More generally, when  $p$  and  $\gamma$  are nonzero and fixed, the correlation tends to achieve a small positive maximum value for some value of  $T_s$  and then decreases monotonically toward its negative asymptotic value as  $T_s \rightarrow \infty$  (see Eq. 15 for the asymptotic value).

## INCLUSION OF mRNA AND RIBOSOME FLUCTUATIONS

In this section, we expand the basic stochastic model introduced in the second section, Steady-state Means, Variances, and Correlation for Protein Counts with Fixed mRNA and Ribosome Numbers, to allow the total mRNA counts,  $T_1$  and  $T_2$ , and the number of ribosomes,  $R$ , to fluctuate randomly.

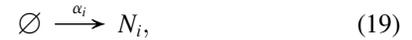
When these fluctuations occur on a slow timescale relative to the time it takes for the protein counts to reach steady-state, we use a formal quasi-steady-state type of approximation for the steady-state statistics of the protein counts in the expanded model. We expect this approximation to become increasingly precise as the limit of arbitrarily slow mRNA and ribosome fluctuations is approached. For this quasi-steady-state approximation, averages (denoted with  $\{\cdot\}$ ) of protein statistics conditional on fixed mRNA counts and ribosome counts are taken over distributions of mRNA and ribosome counts. The resulting variance,  $\Sigma_{Y_i}^2$ , and covariance,  $\Sigma_{Y_1, Y_2}^2$ , for steady-state protein counts can be written as a sum of the (co)variance of the conditional mean(s) and the mean of the conditional (co)variance,

$$\Sigma_{Y_i}^2 = \{ \langle Y_i \rangle^2 \} - \{ \langle Y_i \rangle \}^2 + \{ \sigma_{Y_i}^2 \} \quad (17)$$

and

$$\Sigma_{Y_1, Y_2}^2 = \{ \langle Y_1 \rangle \langle Y_2 \rangle \} - \{ \langle Y_1 \rangle \} \{ \langle Y_2 \rangle \} + \{ \sigma_{Y_1, Y_2}^2 \}. \quad (18)$$

In applying these formulas, we first consider the situation where only the mRNA counts,  $T_1$  and  $T_2$ , are allowed to fluctuate. We assume that production and degradation of the mRNAs are governed by the reactions



with production rates  $\alpha_i$  for  $i = 1, 2$  and common degradation rate  $\delta$ . In the associated stochastic model, for  $i = 1, 2$ , a new copy of the (free) mRNA,  $N_i$ , is produced at each of the jump times of a Poisson process with rate  $\alpha_i$ , and once produced, each copy of mRNA, whether bound or unbound, degrades after an exponentially distributed lifetime with a mean of  $1/\delta$ . Under this model, the mRNA abundances ( $T_1, T_2$ ) constitute a continuous time Markov chain with independent components, where the steady-state distribution for each component is the same as that for a Markovian infinite server queue (28), i.e., Poisson-distributed with mean  $\tau_i = \alpha_i/\delta$  for  $T_i$ . With  $\mu, \gamma, p, \tau_i, i = 1, 2$ , all fixed, slow mRNA dynamics corresponds to the mRNA degradation rate,  $\delta$ , being small relative to the protein degradation rate,  $\gamma$ , and we expect these formulas to be exact in the limit  $\delta \rightarrow 0$  (note that this implies that  $\alpha_i \rightarrow 0$  as  $\tau_i = \alpha_i/\delta$  is fixed for  $i = 1, 2$ ).

Using these steady-state distributions for computing the averages,  $\{\cdot\}$ , in Eqs. 17–18, we obtained exact expressions for the resulting formulas for protein means, variances, and correlation in terms of generalized hypergeometric functions, which are standard functions in software packages such as Maple, Mathematica, and Matlab (see [Supporting Material](#), Section C.2). Sample results produced by numerical evaluation of our formulas for the symmetric case ( $\tau_1 = \tau_2 = \tau$ ) are presented in Fig. 3. It is observed in these results that the correlation is always nonpositive, which we conjecture generally to be true in this restricted setting, though additional effects, e.g., ribosome copy number fluctuations, can lead to positive correlations, as we shall see below. The correlation exhibits a pronounced minimum slightly above the balance point,  $\tau_1 + \tau_2 = R$ . Because this negative extremum in the correlation occurs reliably near the balance point, we term this feature a correlation resonance by analogy with Mather et al. (33). The robustness of this resonance suggests that it is a feature that should be observable experimentally.

The aforementioned formulas can be used to explore the effective interactions that may occur between gene activities

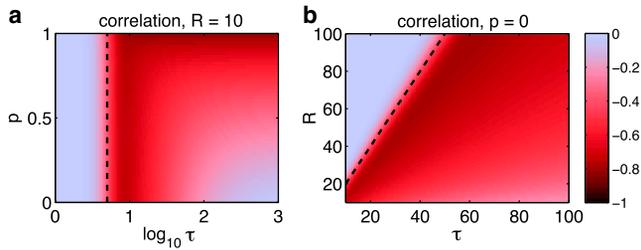


FIGURE 3 Plots of the steady-state correlation between protein counts when abundances  $T_i$  of the mRNAs are independently Poisson distributed with identical means,  $\tau$ , and number of ribosomes,  $R$ , is constant (see the fourth section). Two different scans of parameters are shown with either (a) number of ribosomes fixed,  $R = 10$ , and  $p$  and  $\tau$  variable, or (b)  $p = 0$  fixed and  $R$  and  $\tau$  variable. The balance point,  $2\tau = R$ , is plotted as a black dashed line. The correlation is very near zero for most of the underloaded regime ( $2\tau < R$ ) but exhibits a negative correlation resonance just after  $\tau$  crosses the balance point. Furthermore, the correlation is observed to be nonpositive up to numerical error. Other parameters are  $\mu = 1$  and  $\gamma = 0.1$ .

when translational resources are limited. Translational competition is seen in Fig. 4 to form an effective mutually repressive interaction between two genes, where one gene's activity can reduce another gene's protein output. In Fig. 4 b, repression of the mean level of the second protein type is seen to occur as  $\tau_1$ , the mean mRNA count for the first protein type, is increased, while  $\tau_2$  is held fixed. Corresponding correlation plots in Fig. 4 c again feature a negative correlation resonance; these plots complement the

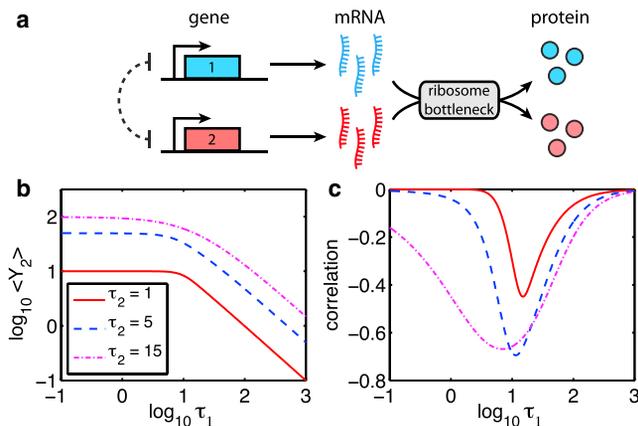


FIGURE 4 (a) Indirect repression of one gene's output by another gene's transcript. Competition for translational resources leads to an effective bottleneck for protein production (see the fourth section). This competition forms an effective repressive interaction between genes (dashed line). For illustration of our model, steady-state means and correlations are computed assuming that the abundances of mRNA copy numbers,  $T_i$ , are independently Poisson distributed with respective means,  $\tau_i$ , and ribosome copy number,  $R$ , is fixed. (b) Repression of the mean protein copy number,  $\langle Y_2 \rangle$ , occurs due to increasing the mean level  $\tau_1$  for various levels of  $\tau_2$ . (c) Correlations between the transcripts produce a substantial negative correlation resonance. There is a wider trough at the minimum when the system is unconditionally overloaded ( $\tau_2 > R$ ), e.g., see results for  $\tau_2 = 15$ . Other parameters are  $R = 10$ ,  $p = 0$ ,  $\mu = 1$ , and  $\gamma = 0.1$ .

symmetric simultaneous induction results for  $\tau_1 = \tau_2 = \tau$  shown in Fig. 3.

We have developed similar formulas based on the quasi-steady-state approximation, for the case when ribosomes are also allowed to fluctuate; these expressions again involve generalized hypergeometric functions. Although computer algebra packages such as Maple can be used to evaluate them, here we instead pursue simpler but still accurate approximations that have a more transparent structure and are quicker to evaluate. Such simplified expressions can be obtained formally by assuming an additional small noise ansatz, where the fluctuations in mRNA and ribosome counts are sufficiently small relative to their mean values (see Supporting Material, Section C.3). If ribosome counts and mRNA counts are all independently distributed, using the small noise ansatz, the lowest-order approximations for the steady-state variance and covariance are given by

$$\Sigma_{Y_i}^2 \approx \sum_{j=1,2} \bar{K}_{ij}^2 \sigma_{T_j}^2 + \bar{K}_{i3}^2 \sigma_R^2 + \bar{\sigma}_{Y_i}^2 \quad (22)$$

and

$$\Sigma_{Y_1, Y_2}^2 \approx \sum_{j=1,2} \bar{K}_{1j} \bar{K}_{2j} \sigma_{T_j}^2 + \bar{K}_{13} \bar{K}_{23} \sigma_R^2 + \bar{\sigma}_{Y_1, Y_2}^2, \quad (23)$$

where  $\bar{\sigma}_{Y_i}^2$  and  $\bar{\sigma}_{Y_1, Y_2}^2$  are protein-count variances;  $\sigma_{T_j}^2$  is the variance of the mRNA count,  $T_j$ ;  $\sigma_R^2$  is the variance of the ribosome count,  $R$ ;  $\bar{K}_{ij}$  for  $j = 1, 2$  is the derivative of the formula for the conditional mean  $\langle Y_i \rangle$  with respect to  $T_j$ ,  $\bar{K}_{i3}$  is the derivative of the formula for  $\langle Y_i \rangle$  with respect to  $R$ , and all expressions with a bar ( $\bar{\cdot}$ ) are evaluated at the mean mRNA and ribosome counts. The contribution from  $\sum_{j=1,2} \bar{K}_{1j} \bar{K}_{2j} \sigma_{T_j}^2$  to the covariance is zero in the underloaded regime, but this contribution is negative in the overloaded regime and thus can lead to negative protein correlation in the overloaded regime. The contribution  $\bar{K}_{13} \bar{K}_{23} \sigma_R^2$  to the covariance is also zero in the underloaded regime but is positive in the overloaded regime. Thus, the ribosome fluctuation term can offset or even dominate the negative contribution from the mRNA fluctuation term in the overloaded regime if the ribosome variance is sufficiently large. This observation also carries over to the case where mRNA distributions are identical:

$$\Sigma_{Y_1}^2 = \Sigma_{Y_2}^2 \approx (\bar{K}_{11}^2 + \bar{K}_{12}^2) \sigma_T^2 + \bar{K}_3^2 \sigma_R^2 + \bar{\sigma}_{Y_1}^2 \quad (24)$$

$$\Sigma_{Y_1, Y_2}^2 \approx 2\bar{K}_{11} \bar{K}_{12} \sigma_T^2 + \bar{K}_3^2 \sigma_R^2 + \bar{\sigma}_{Y_1, Y_2}^2, \quad (25)$$

where we have used symmetry to simplify:  $\sigma_T^2 = \sigma_{T_j}^2$ ,  $\bar{K}_{12} = \bar{K}_{21}$ ,  $\bar{K}_{11} = \bar{K}_{22}$ , and  $\bar{K}_3 = \bar{K}_{13} = \bar{K}_{23}$ . In this symmetric case, the relative magnitude of  $2\bar{K}_{11} \bar{K}_{12} \sigma_T^2$  versus

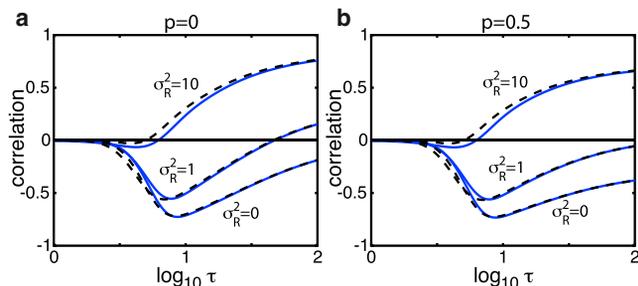
$\overline{K}_3 \sigma_R^2$  plays a significant role when determining the sign of the protein correlation. Note that though this small noise approximation breaks down at the balance point, where derivatives of protein moments suffer a discontinuity, a slight modification of the small-noise expressions can lead to uniform accuracy (see [Supporting Material](#), Section C.3).

To illustrate these results, we suppose that mRNA counts are sampled from independent Poisson distributions with identical means, and that ribosome counts are sampled from a discrete Gaussian steady-state distribution

$$\mathcal{P}(R = r) = \mathcal{N}^{-1} e^{-(r-R_0)^2/2\sigma_R^2}, \quad r = 0, 1, 2, \dots, \quad (26)$$

with  $\mathcal{N} = \sum_{r=0}^{\infty} e^{-(r-R_0)^2/2\sigma_R^2}$  a normalization factor,  $\sigma_R^2$  the approximate variance for the distribution, and  $R_0$  the approximate mean for the distribution. This Gaussian distribution is chosen such that ribosomes may then be over- or underdispersed relative to a Poisson distribution, which mimics the experimental observation that some intracellular molecules such as ribosomes are not Poisson-distributed (39–41). Note that the Gaussian distribution can also approximate a variety of unimodal distributions, including a Poisson distribution. [Fig. 5](#) demonstrates that protein correlations then tend to be negative for small mean mRNA count but positive for large mean mRNA count. If mRNA count fluctuations were larger than the naive Poisson estimate (not displayed in [Fig. 5](#)), then stronger negative correlations would be exhibited.

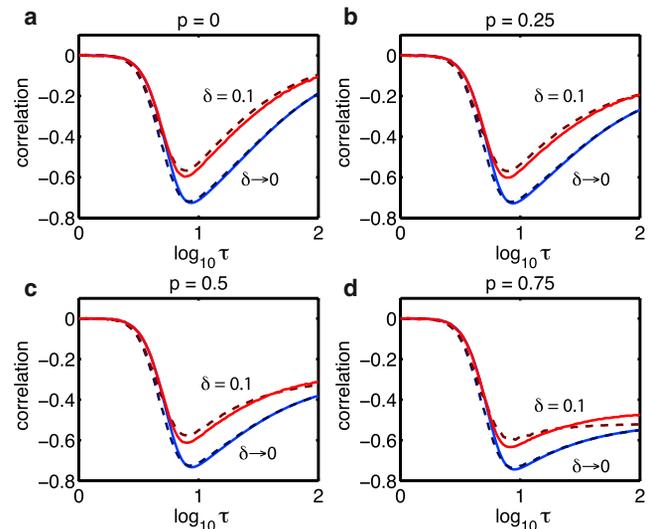
We can slightly relax the quasi-steady-state approximation to allow mRNA fluctuations to occur on a finite time-



**FIGURE 5** Steady-state correlations between protein counts obtained using a quasi-steady-state approximation (*solid lines*). In this calculation, we use our hypergeometric function-based solutions for protein moments derived in the case of Poisson-distributed mRNA and fixed ribosome counts, and we then numerically average such moments over a given distribution for ribosome counts to derive the desired protein moments in the presence of fluctuating mRNA and ribosome counts. This quasi-steady-state calculation can thus be done by numerical evaluation of one-dimensional summations. It is assumed that the mRNA abundances,  $T_i$ , are independently Poisson distributed with common mean  $\tau$ , and the ribosome count  $R$  has an independent discrete Gaussian distribution given by [Eq. 26](#), with approximate mean  $R_0$  and approximate variance parameter  $\sigma_R^2$ . The rebinding probability is either (a)  $p = 0.0$  or (b)  $p = 0.5$ . Approximations based on a small-noise ansatz (see [Eqs. 22–23](#)) are also included (*corresponding dashed lines*). Other parameters are  $R_0 = 10$ ,  $\mu = 1$ , and  $\gamma = 0.1$ .

scale ( $\delta$  comparable in size to  $\gamma$ ) and still obtain reasonably accurate results. This approximation treats the mRNA counts as approximate stochastic processes and leverages the exact linear ordinary differential equation system derived for the moments of the model with fixed mRNA and ribosome counts (see [Supporting Material](#) Section C.4). Several representative comparisons between stochastic simulations and these approximations are given in [Fig. 6](#), where  $\delta \rightarrow 0$  results use the quasi-steady-state approximation, and we exclude the influence of ribosome fluctuations for simplicity. We find good agreement between these approximations and simulation results.

We have also extended our investigation by extensive stochastic simulations (using highly parallel GPU coprocessors) of models where the binding rates of mRNA molecules to ribosomes are finite rather than infinite, and where the binding rates may differ between mRNA species (see [Supporting Material](#)). These simulations show that the negative correlation resonance is a robust feature. The negative correlation extremum typically occurs near the balance point, where total mRNA count equals total ribosome count, except when the mRNA binding rate is very small, in which case the extremum occurs at higher levels of mRNA. An



**FIGURE 6** Plots of the steady-state protein correlation for fluctuating mRNA counts according to [Eqs. 19–21](#) (see the fourth section). Lower solid curves represent results obtained using the quasi-steady-state approximation in which mRNAs are independently Poisson distributed with identical means of  $\tau$ . These curves are labeled with  $\delta \rightarrow 0$ , since this approximation becomes exact in the limit of arbitrarily slowly varying mRNA copy number. Upper solid curves represent results from stochastic simulations when mRNA copy numbers fluctuate on a timescale commensurate with that of the protein dynamics,  $\delta \approx \gamma$ . Simulation parameters were  $\nu = 10^5$ ,  $\delta = 0.1$ , and  $\alpha_i = \tau\delta$ , with ensemble size = 12,800. Approximate solutions derived in the [Supporting Material](#), Sections C.3 and C.4, and discussed at the end of the fourth section of the main text, are included as nearby dashed lines. Results are plotted for (a)  $p = 0$ , (b)  $p = 0.25$ , (c)  $p = 0.5$ , and (d)  $p = 0.75$ . Other parameters are  $R = 10$ ,  $\mu = 1$ , and  $\gamma = 0.1$ .

intuitive explanation for the latter is that with small individual mRNA binding rates, higher mRNA numbers are required to achieve sufficient occupancy of the ribosomes to create the upstream molecular competition required for strong correlations in protein counts.

## DISCUSSION

Gene expression in cells is regulated through a variety of transcriptional, translational, and posttranslational control mechanisms. One of the key factors in translational regulation is a limited access of mRNAs to translational machinery (42). Under these conditions, different mRNAs naturally compete for the limited translational resources. More abundant mRNA transcripts of one type can displace less abundant or lower-affinity transcripts from ribosomes and thus reduce their translational activity, the effect of which may be partially responsible for the sometimes poor correlation between transcript and protein levels (43–45). Translational cross talk is difficult to identify in high-throughput data sets, and only recently have experimental efforts begun to seriously explore competition for ribosomes as an important factor in posttranscriptional regulation (27,46).

In response to the emerging experimental evidence of significant translational cross talk, here we introduced a stochastic molecular model to understand the statistical behavior of a translational bottleneck. In the simplest case, the model assumes constant numbers of ribosomes and mRNAs, but with stochastic events of mRNA-ribosome binding allowing for the possibility of rebinding after translation. However, we later generalized our treatment to a case in which both mRNAs and ribosomes are allowed to fluctuate in number. We found an analytical solution to our basic model, and utilizing concepts from the theory of multiclass queues, we described how the strength of the cross talk strongly depends on whether the ribosomes are underloaded (more ribosomes than mRNAs) or overloaded (more mRNAs than ribosomes). Our model in the underloaded case recovers much of the phenomenology predicted by standard models for protein production, including a lack of cross talk between the production rates of protein species, whereas our model in the overloaded state exhibited substantial cross talk. When the number of mRNAs is allowed to fluctuate slowly, using a formal quasi-steady-state approximation and building on our exact analysis of the basic model, we discovered that the system exhibits a negative correlation resonance (minimum) slightly above the balance point, where the number of ribosomes is equal to the total number of mRNAs. This resonance is analogous, though opposite in sign, to a positive correlation resonance we have found in degradation pathways and queueing systems upstream from a processing bottleneck (33). This correlation resonance appears robust and we anticipate it will be observed reliably when translational cross talk is significant. Finally, we extended our approximate analysis to address

more complex situations, including those where both the numbers of mRNAs and ribosomes are allowed to fluctuate.

Our analysis contains a number of simplifying assumptions that were used to arrive at analytically tractable solutions. One of the primary limitations is tied to a relatively simple description of the kinetics upstream of protein production. Beyond the situation where mRNA and ribosome copy numbers were constant as a function of time, we had to assume that these copy number fluctuations were either slow or relatively small in magnitude. Furthermore, we assumed that the binding between transcripts and ribosomes was instantaneous and that the binding affinities between different types of transcripts and ribosomes are equal. We also neglected polysomal modes of translation where multiple ribosomes can simultaneously translate a single transcript, however this can be straightforwardly incorporated into our description. Exploring these and other generalizations of the basic stochastic model are left for future studies.

## SUPPORTING MATERIAL

Supporting methods, three figures, and reference (47) are available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(13\)00517-1](http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)00517-1).

This work was supported by the National Institutes of Health (R01-GM079333 and the San Diego Center for Systems Biology: P50GM085764). L.T. was supported by National Institutes of Health grant R01-GM089976. R.J.W. was supported in part by National Science Foundation Grant DMS 1206772.

## REFERENCES

- Lodish, H., A. Berk, ..., M. P. Scott. 2012. *Molecular Cell Biology*, 7th ed. W. H. Freeman, New York.
- Grigorova, I. L., N. J. Phleger, ..., C. A. Gross. 2006. Insights into transcriptional regulation and  $\sigma$  competition from an equilibrium model of RNA polymerase binding to DNA. *Proc. Natl. Acad. Sci. USA*. 103:5332–5337.
- Klumpp, S., and T. Hwa. 2008. Growth-rate-dependent partitioning of RNA polymerases in bacteria. *Proc. Natl. Acad. Sci. USA*. 105:20245–20250.
- Voigt, C. A. 2006. Genetic parts to program bacteria. *Curr. Opin. Biotechnol.* 17:548–557.
- Purnick, P. E. M., and R. Weiss. 2009. The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* 10:410–422.
- Alterovitz, G., T. Muso, and M. F. Ramoni. 2010. The challenges of informatics in synthetic biology: from biomolecular networks to artificial organisms. *Brief. Bioinform.* 11:80–95.
- Lee, T. S., R. A. Krupa, ..., J. D. Keasling. 2011. BglBrick vectors and datasheets: a synthetic biology platform for gene expression. *J Biol Eng.* 5:12.
- Morey, K. J., M. S. Antunes, ..., J. Medford. 2012. Crosstalk between endogenous and synthetic components—synthetic signaling meets endogenous components. *Biotechnol. J.* 7:846–855.
- Cheng, A. A., and T. K. Lu. 2012. Synthetic biology: an emerging engineering discipline. *Annu. Rev. Biomed. Eng.* 14:155–178.
- Slusarczyk, A. L., A. Lin, and R. Weiss. 2012. Foundations for the design and implementation of synthetic genetic circuits. *Nat. Rev. Genet.* 13:406–420.

11. Del Vecchio, D., A. J. Ninfa, and E. D. Sontag. 2008. Modular cell biology: retroactivity and insulation. *Mol. Syst. Biol.* 4:161.
12. Kim, K. H., and H. M. Sauro. 2010. Fan-out in gene regulatory networks. *J. Biol. Eng.* 4:16.
13. Cookson, N. A., W. H. Mather, ..., J. Hasty. 2011. Queueing up for enzymatic processing: correlated signaling through coupled degradation. *Mol. Syst. Biol.* 7:561.
14. Kim, K. H., and H. M. Sauro. 2011. Measuring retroactivity from noise in gene regulatory networks. *Biophys. J.* 100:1167–1177.
15. Rondelez, Y. 2012. Competition for catalytic resources alters biological network dynamics. *Phys. Rev. Lett.* 108:018102.
16. Buchler, N. E., and M. Louis. 2008. Molecular titration and ultrasensitivity in regulatory networks. *J. Mol. Biol.* 384:1106–1119.
17. Chapman, S. A., and A. R. Asthagiri. 2009. Quantitative effect of scaffold abundance on signal propagation. *Mol. Syst. Biol.* 5:313.
18. Genot, A. J., T. Fujii, and Y. Rondelez. 2012. Computing with competition in biochemical networks. *Phys. Rev. Lett.* 109:208102.
19. Warner, J. R., J. Vilardell, and J. H. Sohn. 2001. Economics of ribosome biosynthesis. *Cold Spring Harb. Symp. Quant. Biol.* 66:567–574.
20. Mauro, V. P., and G. M. Edelman. 2002. The ribosome filter hypothesis. *Proc. Natl. Acad. Sci. USA.* 99:12031–12036.
21. Mauro, V. P., and G. M. Edelman. 2007. The ribosome filter redux. *Cell Cycle.* 6:2246–2251.
22. Shachrai, I., A. Zaslaver, ..., E. Dekel. 2010. Cost of unneeded proteins in *E. coli* is reduced after several generations in exponential growth. *Mol. Cell.* 38:758–767. <http://dx.doi.org/10.1016/j.molcel.2010.04.015>.
23. Scott, M., C. W. Gunderson, ..., T. Hwa. 2010. Interdependence of cell growth and gene expression: origins and consequences. *Science.* 330:1099–1102.
24. De Vos, D., F. J. Bruggeman, ..., B. M. Bakker. 2011. How molecular competition influences fluxes in gene expression networks. *PLoS ONE.* 6:e28494.
25. Chu, D., D. J. Barnes, and T. von der Haar. 2011. The role of tRNA and ribosome competition in coupling the expression of different mRNAs in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 39:6705–6714.
26. Bennett, M. R., W. L. Pang, ..., J. Hasty. 2008. Metabolic gene regulation in a dynamically changing environment. *Nature.* 454:1119–1122.
27. Baumgartner, B. L., M. R. Bennett, ..., J. Hasty. 2011. Antagonistic gene transcripts regulate adaptation to new growth environments. *Proc. Natl. Acad. Sci. USA.* 108:21087–21092.
28. Asmussen, S. 2003. *Applied Probability and Queues*, 2nd ed. Springer-Verlag, New York.
29. Arazi, A., E. Ben-Jacob, and U. Yechiali. 2004. Bridging genetic networks and queueing theory. *Physica A.* 332:585–616.
30. Levine, E., and T. Hwa. 2007. Stochastic fluctuations in metabolic pathways. *Proc. Natl. Acad. Sci. USA.* 104:9224–9229.
31. Fredriksson, A., M. Ballesteros, ..., T. Nyström. 2007. Decline in ribosomal fidelity contributes to the accumulation and stabilization of the master stress response regulator  $\sigma S$  upon carbon starvation. *Genes Dev.* 21:862–874.
32. Elgart, V., T. Jia, and R. V. Kulkarni. 2010. Applications of Little's Law to stochastic models of gene expression. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 82:021901.
33. Mather, W. H., N. A. Cookson, ..., R. J. Williams. 2010. Correlation resonance generated by coupled enzymatic processing. *Biophys. J.* 99:3172–3181.
34. Mather, W. H., J. Hasty, ..., R. J. Williams. 2011. Factorized time-dependent distributions for certain multiclass queueing networks and an application to enzymatic processing networks. *Queueing Syst.* 69:313–328.
35. Jia, T., and R. V. Kulkarni. 2011. Intrinsic noise in stochastic models of gene expression with molecular memory and bursting. *Phys. Rev. Lett.* 106:058102.
36. Kelly, F. P. 1991. Effective bandwidths at multi-class queues. *Queueing Syst.* 9:5–15.
37. Rajkowitzsch, L., C. Vilela, ..., J. E. McCarthy. 2004. Reinitiation and recycling are distinct processes occurring downstream of translation termination in yeast. *J. Mol. Biol.* 335:71–85.
38. Gadgil, C., C. H. Lee, and H. G. Othmer. 2005. A stochastic analysis of first-order reaction networks. *Bull. Math. Biol.* 67:901–946.
39. Raser, J. M., and E. K. O'Shea. 2004. Control of stochasticity in eukaryotic gene expression. *Science.* 304:1811–1814.
40. Taniguchi, Y., P. J. Choi, ..., X. S. Xie. 2010. Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science.* 329:533–538.
41. Munsky, B., G. Neuert, and A. van Oudenaarden. 2012. Using gene expression noise to understand gene regulation. *Science.* 336:183–187.
42. Gebauer, F., and M. W. Hentze. 2004. Molecular mechanisms of translational control. *Nat. Rev. Mol. Cell Biol.* 5:827–835.
43. Gygi, S. P., Y. Rochon, ..., R. Aebersold. 1999. Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* 19:1720–1730.
44. Guo, Y., P. Xiao, ..., H. Deng. 2008. How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes. *Acta Biochim. Biophys. Sin. (Shanghai).* 40:426–436.
45. Greenbaum, D., C. Colangelo, ..., M. Gerstein. 2003. Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol.* 4:117.
46. Lackner, D. H., M. W. Schmidt, ..., J. Bähler. 2012. Regulation of transcriptome, translation, and proteome in response to environmental stress in fission yeast. *Genome Biol.* 13:R25.
47. Gardiner, C. 2009. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer, Berlin.