SCIENTIFIC REPORTS
nature research

OPEN

Enhancement of gene expression noise from transcription factor binding to genomic decoy sites

Supravat Dey[1] ✉, Mohammad Soltani[1] & Abhyudai Singh[1,2,3,4] ✉

The genome contains several high-affinity non-functional binding sites for transcription factors (TFs) creating a hidden and unexplored layer of gene regulation. We investigate the role of such "decoy sites" in controlling noise (random fluctuations) in the level of a TF that is synthesized in stochastic bursts. Prior studies have assumed that decoy-bound TFs are protected from degradation, and in this case decoys function to buffer noise. Relaxing this assumption to consider arbitrary degradation rates for both bound/unbound TF states, we find rich noise behaviors. For low-affinity decoys, noise in the level of unbound TF always monotonically decreases to the Poisson limit with increasing decoy numbers. In contrast, for high-affinity decoys, noise levels first increase with increasing decoy numbers, before decreasing back to the Poisson limit. Interestingly, while protection of bound TFs from degradation slows the time-scale of fluctuations in the unbound TF levels, the decay of bound TFs leads to faster fluctuations and smaller noise propagation to downstream target proteins. In summary, our analysis reveals stochastic dynamics emerging from nonspecific binding of TFs and highlights the dual role of decoys as attenuators or amplifiers of gene expression noise depending on their binding affinity and stability of the bound TF.

The level of a gene product can show remarkable cell-to-cell differences within an isogenic cell population exposed to the same external environment[1–9]. This intercellular variation has been referred to as gene expression noise and is attributed to biochemical processes operating with low-copy number components, such as the number of the promoter, mRNA, and protein for a given gene. The noise in expression critically impacts the functioning of diverse cellular processes in both beneficial and detrimental ways. For example, on one hand the noise is detrimental for developing embryos[7,10,11], organismal fitness[12] and has been connected to disease phenotypes[13,14]. On the other hand, noisy expression drives phenotypic heterogeneity between otherwise genetically-identical cells, allowing them to hedge their bets against environmental uncertainties[5,8,15–24]. While there has been considerable work uncovering the role of transcription/translation processes together with molecular feedbacks in driving expression noise[25–28], how nonspecific binding of a protein shapes this noise remains poorly understood.

A fundamental step in gene regulation is the binding of a transcription factor (TF) to its target promoter[29–31]. Besides this specific binding, TFs also bind to other non-functional high-affinity sites spread across the genome. These spurious binding sites, known as transcription factor decoys, can be present in different abundances with various binding affinities[32–35]. The regularity role of decoys through molecular sequestration of TFs has been experimentally demonstrated via synthetic circuits in *Escherichia coli*[36] and *Saccharomyces cerevisiae*[37], and not surprisingly, the TF-inhibiting activity of decoys has tremendous therapeutic potential[38–41]. An immediate consequence of TF sequestration by decoys is the modulation of TF's stability. Depending on the context, decoy binding can either enhance or reduce the stability of the TFs[42]. For certain TFs, such as MyoD, the DNA binding provide increased stability against degradation[43,44]. On the other hand, for VP16 in *Saccharomyces cerevisiae*, the bound TFs become more prone to degradation by the ubiquitin-mediated proteolysis[45]. A key focus of this work is to investigate how the relative stability of bound vs. unbound TF affects both the extent and time-scale of fluctuations in the level of a given TF.

Prior theoretical work on this topic has highlighted the role of decoys as noise buffers, in the sense that, the presence of decoys attenuates random fluctuations in number of freely (unbound) available copies of the

[1]Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, 19716, USA. [2]Department of Biomedical Engineering, University of Delaware, Newark, DE, 19716, USA. [3]Department of Mathematical Sciences, University of Delaware, Newark, DE, 19716, USA. [4]Center for Bioinformatics and Computational Biology, University of Delaware, Newark, DE, 19716, USA. ✉e-mail: supravat.dey@gmail.com; absingh@udel.edu

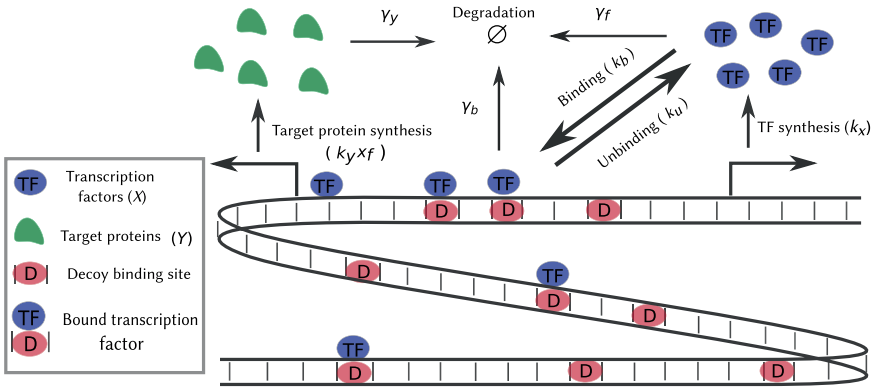

Figure 1. Model schematic for investigating the impact of nonspecific transcription factor binding on expression noise. A genome with several decoy binding sites where transcription factors (TFs) bind reversibly, is depicted. The synthesis of TFs occurs in stochastic bursts. Both the free and bound TFs are subject to degradation. The free TFs activate a target gene and regulate its bursty protein synthesis.

TF[42,46–55]. However, these results are based on the assumption that sequestration of TF at a decoy site protects the TF from degradation. Relaxing this assumption to consider an arbitrary decay rate of the bound TF, we uncover a novel role of decoys as both noise amplifiers and buffers. We systematically characterize the parameter regimes leading to these contrasting roles in terms of the bound vs. free-TF stability, number and affinity of decoy sites. Finally, we study noise transmission from the TF to a downstream target protein reporting counterintuitive effects in some cases, for example, decoys amplify noise in the level of a TF but reduces noise in the level of the TF's downstream target protein.

Results

Model formulation. To study the role of decoy binding sites, we consider a TF that is synthesized in stochastic bursts (Fig. 1). Such bursty expression of gene products has been experimentally observed in diverse cell types[56–63], and corresponds to distinct mechanisms at the transcriptional and translation level. For example, a promoter randomly switches from a transcriptionally inactive to an active state, and quickly turns off after producing a bursts of mRNAs[59,64–66]. Moreover, assuming short-lived mRNAs, each synthesized mRNA is rapidly degraded after translating a burst of proteins[67–69]. We phenomenologically model the combined effect of transcriptional and translational bursts by considering a Poisson arrival of burst events with rate k_x, and each event creates B_x molecules, where B_x is an independent and identically distributed non-negative random variable following an arbitrary distribution[70–75]. More specifically, $B_x = i$ with probability $\alpha_x(i)$ where $i \in \{0, 1, 2, \ldots\}$. We consider a general form for the burst size distribution $\alpha_x(i)$ throughout the paper, except for plotting and simulation purposes where $\alpha_x(i)$ follows a geometrical distribution[69,76–78].

Consider N decoy binding sites in the genome, with the TF binding/unbinding to each decoy site with rates are k_b and k_u, respectively. As motivated earlier, we allow for both the free (unbound) and bound TF to decay with rates γ_f and γ_b, respectively. Let $x_f(t)$ denote the number of free TF molecules, and $x_b(t)$ the number of bound TFs at time t inside a single cell. Then, the stochastic evolution of $x_f(t)$ and $x_b(t)$ is governed by the following probabilities

$$\text{TF synthesis: Probability}\{x_f(t+dt) = x_f(t) + i\} = k_x \alpha_x(i) dt, \tag{1a}$$

$$\begin{aligned}\text{TF binding: Probability}\{x_f(t+dt) = x_f(t) - 1, \;\; x_b(t+dt) = x_b(t) + 1\} \\ = k_b x_f(t)(N - x_b(t))dt,\end{aligned} \tag{1b}$$

$$\text{TF unbinding: Probability}\{x_f(t+dt) = x_f(t) + 1, \;\; x_b(t+dt) = x_b(t) - 1\} = k_u x_b(t) dt, \tag{1c}$$

$$\text{Free TF degradation: Probability}\{x_f(t+dt) = x_f(t) - 1\} = \gamma_f x_f(t) dt, \tag{1d}$$

$$\text{Bound TF degradation: Probability}\{x_b(t+dt) = x_b(t) - 1\} = \gamma_b x_b(t) dt. \tag{1e}$$

Each equation here defines a stochastic event that occurs with a certain probability in the small time interval $(t, t+dt]$, and whenever the event occurs, the population counts change by discrete integer amounts. Based on the underlying stochastic chemical kinetics, these occurrence probabilities are either independent (as in 1a), or linearly/nonlinearly dependent on the molecular counts. For readers convenience, we provide a list of model parameters along with their description in Table 1. Expressing γ_b in terms of γ_f as $\gamma_b = \beta\gamma_f$, $\beta = 0$ corresponds to no decay of bound TF, and $\beta = 1$ corresponds to equal degradation rates for both the free and bound TF. The key focus of this work is to understand the stochastic dynamics of $x_f(t)$ in different regimes of β.

Parameter	Description	Parameter	Description
x_f	Free TF count	x_b	Bound TF count
N	Total decoy binding sites	k_x	TF burst frequency
k_b	TF binding rate	k_u	TF unbinding rate
γ_b	Bound TF degradation rate	γ_f	Free TF degradation rate
β	γ_b/γ_f	k_d	Dissociation constant (k_u/k_b)
y	Target protein count	$k_y x_f$	Target protein burst frequency
γ_y	Target protein degradation rate	α_x (α_y)	Burst size distribution of $X(Y)$
$\langle B_x \rangle$ ($\langle B_y \rangle$)	Average burst size for X (Y)	$\langle B_x^2 \rangle$ ($\langle B_y^2 \rangle$)	Second moment of burst size distribution for X (Y)

Table 1. Summary of notation used.

Based on the above discrete-state continuous-time Markov model, one can write a corresponding Chemical Master Equation (CME) that provides the time evolution of the joint probability density function $p(x_f, x_b, t)$, for observing x_f free TF, and x_b bound TF molecules at time t[79,80]

$$\begin{aligned}\frac{\partial p(x_f, x_b, t)}{\partial t} &= k_x \sum_{i=0}^{x_f} \alpha_x(i) p(x_f - i, x_b, t) + k_u(x_b + 1)p(x_f - 1, x_b + 1, t) \\ &+ k_b(x_f + 1)(N - x_b + 1)p(x_f + 1, x_b - 1, t) \\ &+ \gamma_f(x_f + 1)p(x_f + 1, x_b, t) + \gamma_b(x_b + 1)p(x_f, x_b + 1, t) \\ &- [k_x + \gamma_f x_f + \gamma_b x_b + k_u x_b + k_b x_f(N - x_b)]p(x_f, x_b, t).\end{aligned} \tag{2}$$

As the CME is analytically intractable, it is typically solved numerically through either the Finite State Projection algorithm[81–83] or various Monte Carlo simulation techniques[84–88]. Taking an alternative approach, we focus on the statistical moments of the molecular counts and use the well-known Linear Noise Approximation[89–94] to obtain closed-form formulas for the mean and noise levels.

In addition to the LNA, we assume that the binding/unbinding reactions are very fast compared to protein synthesis and degradations. This fast binding/unbinding limit, which is also known as quasi-static equilibrium or adiabatic limit, is common in the gene expression modelling[29,95,96] and is supported by recent experiments[97]. This assumption simplifies our calculations to obtain the analytical expression for the Fano factors. As this limit implies negligible fluctuations due to binding kinetics, one may expect relaxing this limit can give rise to more variation in gene expression[95,97]. We have numerically verified that for our system, the qualitative behavior of results does not depend on this limit (see Fig. S1).

Noise in free TF counts in the absence of decoy sites. When there are no decoys ($N = 0$), the moments of the free TF count can be solved exactly from the CME. In particular, the steady-state mean level $\overline{\langle x_{f,0} \rangle}$, and the Fano factor $F_{x_f,0}$ (variance/mean) of $x_f(t)$ are given by[26,98],

$$\overline{\langle x_{f,0} \rangle} = \frac{k_x \langle B_x \rangle}{\gamma_f}, \quad \text{and} \quad F_{x_f,0} := \frac{\overline{\langle x_{f,0}^2 \rangle} - \overline{\langle x_{f,0} \rangle}^2}{\overline{\langle x_{f,0} \rangle}} = \frac{\langle B_x \rangle + \langle B_x^2 \rangle}{2 \langle B_x \rangle}, \tag{3}$$

respectively, where $\langle B_x \rangle$ and $\langle B_x^2 \rangle$ are the first and second-order moments of the burst size B_x. Throughout the paper we use angular brackets $\langle\ \rangle$ to denote the expected value operation, and $\overline{\langle\ \rangle}$ to represent the steady-state expected value. Note that the Fano factor is completely determined by the burst size distribution and is independent of the burst arrival rate and the protein decay rate. As expected, we recover the Poisson limit ($F_{x_f,0} = 1$) for non-bursty production $B_x = 1$ with probability one, and noise is super-Poissonian ($F_{x_f,0} > 1$) for bursty production. In the special case where B_x follows a geometric distribution

$$\alpha_x(i) = (1 - 1/\langle B_x \rangle)^{i-1}/\langle B_x \rangle \text{for } i \in \{1, 2, 3, \dots\}, \tag{4}$$

the steady-state Fano factor $F_{x_f,0} = \langle B_x \rangle$ is equal to the mean burst size.

Bound TF's degradation titrates the regulating activity of TF. In the presence of decoy ($N > 0$), exact solutions to the mean and noise levels are unavailable, and one has to resort to approximation techniques. Recall from (1), that the probability of the TF binding event is nonlinearly related to the molecular counts via the product term $x_f(t)x_b(t)$. This nonlinearity results in unclosed moment dynamics – the time evolution of lower-order moments depends on higher-order moments[99,100]. For example, the dynamics of the means $\langle x_f \rangle$, $\langle x_b \rangle$ depends on the second order moment $\langle x_f x_b \rangle$ (see SI, section 1). Typically approximate closure schemes are employed to solve for moments in such cases[101–111]. Here, we use the Linear Noise Approximation (LNA) method, where assuming small fluctuations in $x_f(t)$ and $x_b(t)$ around their respective mean values $\langle x_f \rangle$ and $\langle x_b \rangle$, the nonlinear term is linearized as $k_b x_f x_b \approx k_b(x_f \langle x_b \rangle + \langle x_f \rangle x_b - \langle x_b \rangle \langle x_f \rangle)$. Exploiting this linearization, the probability of all events in (1) become linear with respect to the molecular counts, resulting in closed moment dynamics (see SI, section 1). A

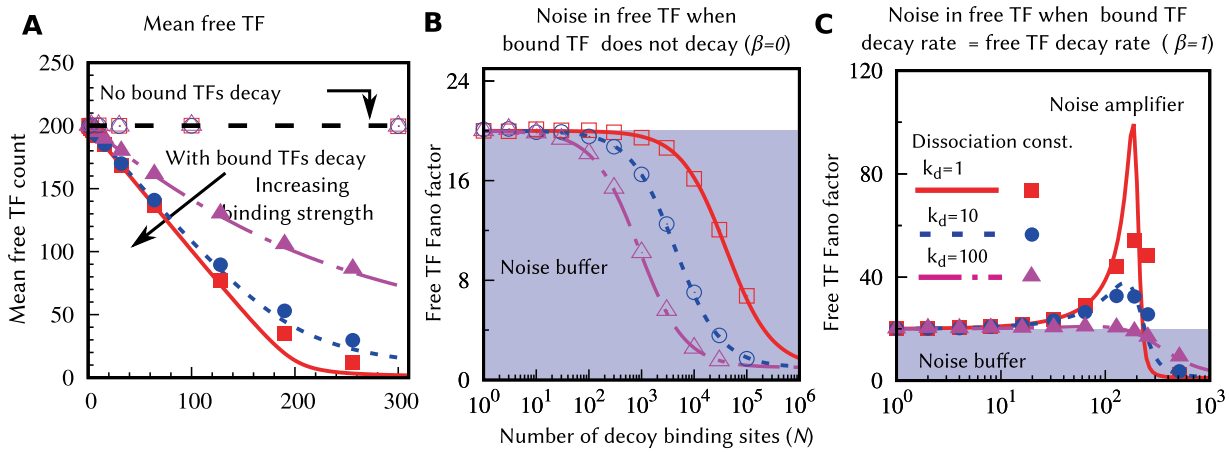

Figure 2. Degradation of bound TFs reduces the mean and drives non-monotonicity in the free TF noise level: The mean and the Fano factor for free TF counts are plotted against the total decoy binding sites N for different values of the dissociation constant ($k_d=1$, $k_d=10$ and $k_d=100$). Lines are plotted using the analytical formulas (6) and (9), and the results with symbols are obtained from stochastic simulations using Gillespie algorithm [112]. (**A**) When bound TFs are protected from degradation, the mean becomes independent of N and k_d. In the presence of bound TF degradation (results are shown for $\beta=1$), the mean free TF count decreases with N. This decay becomes faster for larger binding strengths. (**B**) If bound TFs are protected from the degradation, the Fano factor decreases monotonically with N, suggesting decoys role as a noise buffer. (**C**) For intermediate values of decoy sites, a large noise enhancement can be seen in the presence of bound TF degradation. This noise amplification becomes larger for smaller values of k_d. The Fano factors for both cases (**B**,**C**) approach to the Poissonian limit (Fano factor = 1) for large N. The parameters used for this figure: $\langle B_x\rangle=20$, $k_x=10\,hr^{-1}$, and $\gamma_f=1.0\,hr^{-1}$ per protein molecule. For simulations, $k_b=50\,hr^{-1}$ per pair of molecules.

direct result of using the LNA is that the time evolution of the means is identical to the deterministic chemical rate equations

$$\frac{d\langle x_f\rangle}{dt} = k_x\langle B_x\rangle + k_u\langle x_b\rangle - k_bN\langle x_f\rangle + k_b\langle x_b\rangle\langle x_f\rangle - \gamma_f\langle x_f\rangle, \tag{5a}$$

$$\frac{d\langle x_b\rangle}{dt} = -k_u\langle x_b\rangle + k_bN\langle x_f\rangle - k_b\langle x_b\rangle\langle x_f\rangle - \gamma_b\langle x_b\rangle. \tag{5b}$$

Solving the above equations at steady state, and further considering fast binding/unbinding rates ($k_b\to\infty$, $k_u\to\infty$) for a given dissociation constant $k_d=k_u/k_b$, yields the following mean levels for the unbound and bound TF

$$\overline{\langle x_f\rangle} = \frac{\overline{\langle x_{f,0}\rangle} - N\beta - k_d + \sqrt{(\overline{\langle x_{f,0}\rangle} - N\beta - k_d)^2 + 4k_d\overline{\langle x_{f,0}\rangle}}}{2}, \quad \overline{\langle x_b\rangle} = \frac{N\overline{\langle x_f\rangle}}{k_d + \overline{\langle x_f\rangle}}, \tag{6}$$

respectively. Here, $\overline{\langle x_{f,0}\rangle}$ given by (3) is the mean TF count in the absence of decoy sites. When bound TFs are protected from degradation $\beta=0$, the mean free TF count $\overline{\langle x_f\rangle} = \overline{\langle x_{f,0}\rangle} = k_x\langle B_x\rangle/\gamma_f$ becomes independent of decoy numbers[42,52]. In contrast, with bound TF degradation $\beta>0$, the mean free TF count monotonically decreases with increasing decoy numbers, with the decrease being faster for stronger binding affinity (or lower dissociation constant). This point is exemplified in Fig. 2(A) where we plot $\overline{\langle x_f\rangle}$ as a function of N for different dissociation constants. In the limit of a small number of decoys, (6) can be approximated as

$$\overline{\langle x_f\rangle} \approx \overline{\langle x_{f,0}\rangle} - \frac{\overline{\langle x_{f,0}\rangle}}{\overline{\langle x_{f,0}\rangle} + k_d}N\beta, \tag{7}$$

and the rate of decrease of $\overline{\langle x_f\rangle}$ with increasing N is inversely proportional to k_d. In the limit of large N and $\beta>0$,

$$\overline{\langle x_f\rangle} \approx \frac{k_d\overline{\langle x_{f,0}\rangle}}{N\beta}, \tag{8}$$

exhibiting a $1/N$ scaling of mean free TF levels with decoy abundance. Both these limits emphasize the point that when $\beta>0$, increasing decoy numbers titrate away the TF, leading to reduced levels of the free TF. These results are consistent with experimental data in *Saccharomyces cerevisiae* showing reduced activity of the TF as a result of decoy binding[37].

Bound TF's degradation enhances noise in the free TF count. Next, LNA is used to derive F_{x_f}, the steady-state Fano factor of the free TF level in the presence of decoys. Assuming fast binding and unbinding of TFs to decoy sites ($k_b \to \infty$, $k_u \to \infty$, fixed $k_d = k_u/k_b$) we obtain (see SI, section 1 for details)

$$\begin{aligned} F_{x_f} &= \frac{\overline{\langle x_f^2 \rangle} - \overline{\langle x_f \rangle}^2}{\overline{\langle x_f \rangle}} \\ &= 1 + \frac{\overline{\langle x_f \rangle}[(F_{x_f,0} - 1)\overline{\langle x_{f,0} \rangle} + \beta f^2 N]}{(\overline{\langle x_{f,0} \rangle} - \beta f^2 N)(Nf(1-f) + \overline{\langle x_f \rangle})}, \quad \text{with } f \\ &= \frac{\overline{\langle x_b \rangle}}{N} \\ &= \frac{\overline{\langle x_f \rangle}}{k_d + \overline{\langle x_f \rangle}}. \end{aligned} \tag{9}$$

Here, $F_{x_f,0}$ is the Fano factor in the absence of decoy binding sites (3), and f is the fraction of bound decoys. As expected, in the limit of no decoys ($N \to 0$) or weakly binding decoys ($k_d \to \infty$), $F_{x_f} \to F_{x_f,0}$.

When bound TFs are protected from degradation ($\beta = 0$), then $\overline{\langle x_f \rangle} = \overline{\langle x_{f,0} \rangle}$ and the fraction of bound decoy f becomes independent of N. In this scenario, (9) simplifies to

$$F_{x_f} = 1 + \frac{F_{x_f,0} - 1}{\frac{Nf(1-f)}{\overline{\langle x_{f,0} \rangle}} + 1}, \quad \text{with } f = \frac{\overline{\langle x_{f,0} \rangle}}{k_d + \overline{\langle x_{f,0} \rangle}}. \tag{10}$$

As f is independent of N, it is clear from (10) that F_{x_f} monotonically decreases from $F_{x_f,0}$ to 1 with increasing N (Fig. 2(B)). Thus, when the bound TFs are protected from degradation, the decoy sites function as a *noise buffer* with $F_{x_f} < F_{x_f,0}$[52].

An interesting finding that emerges form (9) is that when $\beta > 0$, F_{x_f} can vary non-monotonically with N. This point is illustrated in Fig. 2(C), where we plot F_{x_f} as a function of decoy abundance for $\beta = 1$. While F_{x_f} monotonically decreases to 1 for weak binding affinities (similar to the case of $\beta = 0$), for strong binding affinities F_{x_f} first increases with increasing N to reach a maxima, before decreasing back to the Poisson limit for large N. In essence, with degradation of bound TFs, decoys can function as a *noise amplifier* ($F_{x_f} > F_{x_f,0}$). Checking the sign of the derivative $dF_{x_f}/dN > 0$ in the limit $N \to 0$ leads to an analytical condition for noise enhancement – if the dissociation constant is below a critical threshold

$$k_d < k_d^{th} = \frac{F_{x_f,0}\beta\overline{\langle x_{f,0} \rangle}}{F_{x_f,0} - 1}, \tag{11}$$

then the Fano factor will increase starting from $F_{x_f,0}$ as decoy sites are introduced. For $F_{x_f,0} \gg 1$, the threshold value simplifies to $k_d^{th} = \beta\overline{\langle x_{f,0} \rangle}$, independent of $F_{x_f,0}$. It turns out that this condition for noise enhancement implies that the fraction of bound decoys $f > 1/(1+\beta)$ when a small number of decoys are present. In spite of the noise amplification, it is important to point out that $F_{x_f} \to 1$ as $N \to \infty$ irrespective of the value of β. Thus, strongly-binding decoys that function as a noise amplifier for small N, become a noise buffer for large N.

Overall these results suggest that decoys buffer noise in a variety of scenarios: $\beta = 0$ (irrespective of N and k_d), or for a large number of decoys (irrespective of β and k_d) or if decoys have sufficiently weak binding affinities (irrespective of β and N). In contrast, decoys function as a noise amplifier when $\beta > 0$ provided two other conditions hold – decoys have sufficiently strong binding affinities as per (11) *and* are not present in large numbers.

To check the validity of the linearization and fast binding/unbinding approximations, we perform kinetic Monte Carlo simulations using the Gillespie algorithm[112] to obtain numerically exact results. In Fig. 2, along with the analytical results (lines) we also plot the simulation results (symbols). The match between analytical and simulations results are quite well, especially for large and intermediate k_d values. For small k_d values with $\beta \neq 0$, there is a clear deviation from the analytical results. In this regime, being fluctuations very large, the linearization approximation based on absence of large fluctuations may not be justified. However, as can be seen, the qualitative features do not depend on these approximations.

As illustrated in Fig. 2(A), bound TFs degradation reduces the number of available TFs with increasing decoy abundance. This naturally leads to the question: is the noise behavior reported in Fig. 2(C) also seen when the mean free TF count is held constant at given desired level? In Fig. 3, we plot the noise as a function N and k_d for a given mean free TF count $\overline{\langle x_f \rangle}$ by simultaneously enhancing the production rate k_x as per (6). Our results show similar qualitative behaviors with decoys functioning as a noise buffer for $\beta = 0$ (Fig. 3(C)), becoming a noise amplifier when $\beta = 1$ for sufficiently small N and k_d (Fig. 3(B)). Interestingly, the region of noise amplification is greatly enhanced when bound TFs become more unstable compared to their free counterparts (Fig. 3(A)).

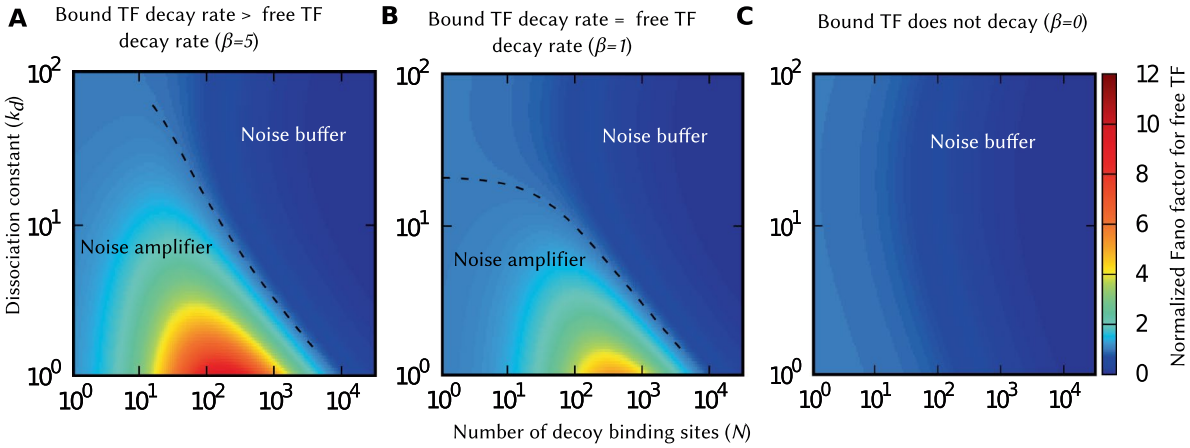

Figure 3. Decreasing stability of bound TFs expands the parameter space for decoy-mediated noise enhancement. The normalized Fano factors ($F_{x_f}/F_{x_f,0}$) for different degradation rates of bound TFs are plotted as a function of decoy numbers (N) and the dissociation constant (k_d), for a constant mean free TF count. The color box represents the scale for the normalized Fano factor, and its value larger than one implies noise enhancement. For smaller values of k_d and N, the decoy acts as a noise amplifier when bound TFs are unstable. The region of noise enhancement (demarcated by a dashed line representing $F_{x_f} = F_{x_f,0}$) increases with increasing degradation rate of bound TFs. Parameters used: $\overline{\langle x_f \rangle} = 20$ molecules, $\langle B_x \rangle = 20$, and $\gamma_f = 1\, hr^{-1}$ per protein molecule. These plots are generated by using (9). While to keep $\overline{\langle x_f \rangle}$ constant, we change $\overline{\langle x_{f,0} \rangle}$ accordingly by varying k_x and obeying (6).

Noise in free TF counts in a mixture of strong and weak decoy binding sites. Inside cells, TFs bind to various decoy sites with different affinities[33]. How do fluctuations in the TF count gets affected by a mixture of nonidentical decoys? To address this question, we study a system two decoy types D1 and D2 that are present in numbers N_1 and N_2, respectively. We assume that each free TF molecule can bind to both decoy types with the same (diffusion-limited) rate k_b, but unbinds with different rates k_{u1} and k_{u2}, respectively, leading to two different dissociation constants $k_{d1} = k_{u1}/k_b$ and $k_{d2} = k_{u2}/k_b$. As before, TFs are synthesized in random bursts, the free and bound TFs decay with rates γ_f and γ_b (the decay rates of TFs bound to D1 and D2 are assumed to be equal). Together with (1a) and (1d), the stochastic model is now governed by the following probabilities representing jumps in the population counts in the infinitesimal time interval $(t, t+dt]$

$$\begin{aligned}&\text{TF binding decoy D1: Probability}\{x_f(t+dt)\\&= x_f(t)-1,\, x_{b1}(t+dt)\\&= x_{b1}(t)+1\}\\&= k_b x_f(t)(N_1 - x_{b1}(t))dt,\end{aligned} \tag{12a}$$

$$\begin{aligned}&\text{TF binding decoy D2: Probability}\{x_f(t+dt)\\&= x_f(t)-1,\, x_{b2}(t+dt)\\&= x_{b2}(t)+1\}\\&= k_b x_f(t)(N_2 - x_{b2}(t))dt,\end{aligned} \tag{12b}$$

$$\begin{aligned}&\text{TF unbinding from D1: Probability}\{x_f(t+dt)\\&= x_f(t)+1,\, x_{b1}(t+dt)\\&= x_{b1}(t)-1\}\\&= k_{u1} x_{b1}(t)dt,\end{aligned} \tag{12c}$$

$$\begin{aligned}&\text{TF unbinding from D2: Probability}\{x_f(t+dt)\\&= x_f(t)+1,\, x_{b2}(t+dt)\\&= x_{b2}(t)-1\}\\&= k_{u2} x_{b2}(t)dt,\end{aligned} \tag{12d}$$

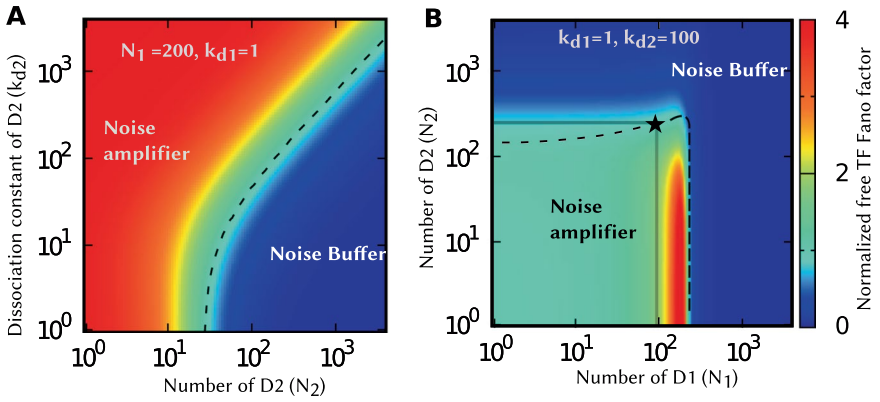

Figure 4. Noise-buffering decoys can mitigate the effects of noise-amplifying decoys. (**A**) The density plots of normalized Fano factor against N_2 and k_{d2} for $N_1 = 200$ and $k_{d1} = 1$. (**B**) The normalized Fano factor against N_1 and N_2 for $k_{d1} = 1$ and $k_{d2} = 100$. The noise enhancement region is separated from noise buffer region with dashed lines (correspond to $F_{x_f}/F_{x_f,0} = 1$). While individual decoy types act oppositely on free TF noise, presence of both can cancel this effect and maintain the same noise level as in the absence of any decoys. The point marked by the star is such a representative point. Parameter used: $\langle B_x \rangle = 20$, $\gamma_f = \gamma_b = 1\, hr^{-1}$ per molecule, $k_x = 10\, hr^{-1}$, and $k_b = 1000\, hr^{-1}$.

$$\text{D1} - \text{bound TF degradation: Probability}\{x_{b1}(t+dt) = x_{b1}(t) - 1\} = \gamma_b x_{b1}(t)dt, \tag{12e}$$

$$\text{D2} - \text{bound TF degradation: Probability}\{x_{b2}(t+dt) = x_{b2}(t) - 1\} = \gamma_b x_{b2}(t)dt. \tag{12f}$$

Here x_{b1} and x_{b2} denote the number TFs bound to decoys D1 and D2, respectively. We apply the linear noise approximation to obtain time evolution of statistical moments (presented in the SI, section 2), and solve the resulting moment equations at the steady to compute the Fano factor numerically as the analytical formula for the noise level becomes quite involved.

Noise in the free TF counts is investigated in two complementary ways. In Fig. 4A, we fix the number of decoys D1 and their binding affinity such that D1 itself functions as a noise amplifier. The noise in the free TF counts is then plotted as a function of the number of decoys D2 and its binding affinity. Recall that the dashed line represents $F_{x_f} = F_{x_f,0}$, i.e., the noise with decoys is the same as the noise in the absence of decoys. Having a large number of decoys D2 makes the overall decoy mixture a noise buffer, with the dashed line showing how large of a pool N_2 is needed as a function of k_{d2}. In Fig. 4B we fix the binding affinity of both decoys and plot the noise level as a function of decoys abundances. Here the noise enhancement is only observed when both decoys are present at small numbers, and the decoy mixture is a noise buffer even if one of the decoys is present in sufficiently large numbers. It is interesting to note that in Fig. 4B, D1 by itself is a noise amplifier (gray line intersection with the x-axis), D2 by itself is a noise buffer (gray line intersection with the y-axis), but their combined presence (star on the dashed line) mitigates each other's effect, and the noise level is similar to when there were no decoys.

Quantifying noise propagation from TF to downstream target proteins. Having quantified the magnitude of fluctuations in the free TFs counts, we next focus our attention on the time-scale of fluctuations. Given that the available TF pool activates downstream target proteins, the time scale at which fluctuations relax back to mean levels is key in understanding downstream noise propagation[113,114]. For example, for a given noise level, more prolonged fluctuations in the free TF counts will lead to higher noise in the expression of the target protein.

The time-scale of fluctuations is characterized using the autocorrelation function defined as,

$$R(\tau) = \frac{\langle x_f(\tau + t)x_f(t)\rangle - \overline{\langle x_f \rangle}^2}{\overline{\langle x_f^2 \rangle} - \overline{\langle x_f \rangle}^2}, \tag{13}$$

where time t is sufficiently large for the system to reach the steady-state. In the absence of any decoy binding sites, the decay in the autocorrelation function is given by $\exp(-\gamma_f \tau)$[52]. Note that this function only depends on the decay rate and independent of bursting parameters. Thus, the magnitude and time-scale of fluctuations can be independently tuned using (3) and (13) via the burst size and the decay rate. In the presence of decoy binding sites, we compute $R(\tau)$ numerically from stochastic realizations of $x_f(t)$ obtained via kinetic Monte Carlo simulations[112]. Figure 5 plots the decay of $R(\tau)$ as a function of lag-time τ, in the absence and presence of bound TF degradation. When the bound TFs are protected from degradations, the autocorrelation function shifts to the right leading to slower and more prolonged fluctuations in the free TF count. In contrast, when bound TFs are unstable, $R(\tau)$ shifts to the left, resulting in faster fluctuations, and the shift is more pronounced for larger values of β.

To understand how fluctuations in the free TF propagate downstream, we consider a case where the stochastic synthesis of target protein Y is activated by available TFs (see Fig. 1). Instead of incorporating the direct binding

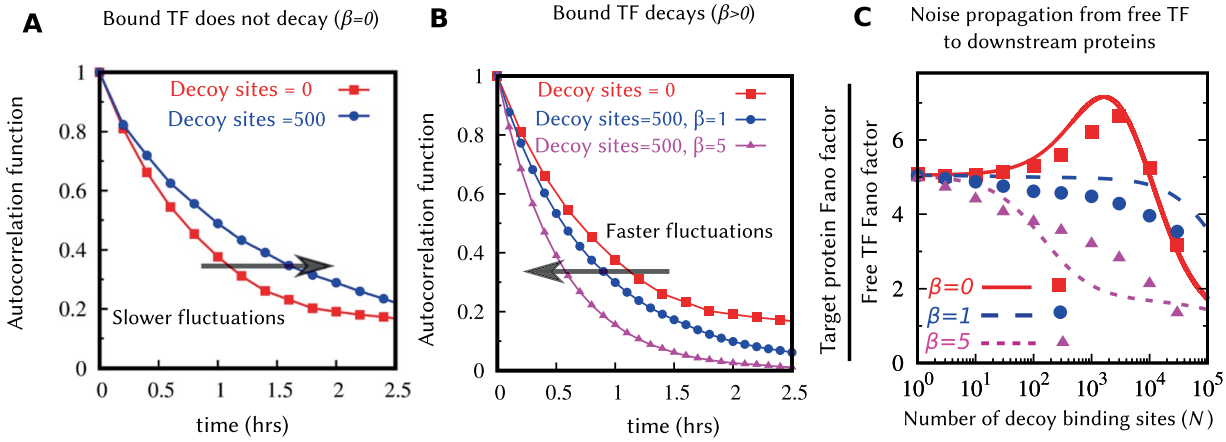

Figure 5. Bound TFs degradation leads to faster fluctuations in the free TF counts reducing noise propagation to downstream proteins. The simulation results of autocorrelation function $R(\tau)$ of the free TF count given by (13) for (**A**) Bound TFs are protected from degradation, and (**B**) Bound TFs are subject to degradation. The addition of decoy binding sites makes the autocorrelation decay slower in (**A**) but faster in (**B**). (**C**) Noise propagation from the free TF to the downstream target protein is measured by the ratio of the Fano factors F_y/F_{x_f}. Lines are the plotted using the analytical formula of F_y and F_{x_f}, given by (16) and (9) respectively. Symbols represent corresponding results with simulations. Noise propagation is enhanced when $\beta=0$ consistent with the shift in the autocorrelation function to the right in Fig. 5A. In contrast, noise propagation is reduced when $\beta>0$. Parameters used: $\overline{\langle x_f \rangle} = 20$ molecules, $\langle B_x \rangle = 20$, $k_d = 1$, $\langle B_y \rangle = 1$, $k_u = 10\,hr^{-1}$, and $\gamma_f = \gamma_y = 1\,hr^{-1}$ per protein molecule. For simulations, $k_b = 50$ per pair of molecules. To keep $\overline{\langle x_f \rangle}$ constant, we change $\overline{\langle x_{f,0} \rangle}$ accordingly by varying k_x and obeying (6).

of TFs to the promoter of the target gene, we model the activation via a linear dose-response by the TF — the synthesis rate of Y is proportional to the number of available TF. We note that this linearity assumption is more appropriate when the binding affinity of TFs to the target gene is relatively small. This limit provides a theoretical understanding of noise propagation in the target protein as the mathematical formula for the noise can be derived.

The target protein is assumed to be produced in bursts B_y with an arbitrary burst size distribution

$$\text{Probability}\{B_y = i\} = \alpha_y(i), \tag{14}$$

and burst frequency $x_f k_y$ that increases linearly with the free TF count. The probabilistic events governing the production and decay of target proteins are given by

$$\text{Probability}\{y(t+dt) = y(t) + i\} = k_y x_f \alpha_y(i) dt, \tag{15a}$$

$$\text{Probability}\{y(t+dt) = y(t) - 1\} = \gamma_y y(t) dt, \tag{15b}$$

where $y(t)$ is the population count of protein Y at time t. Applying LNA to the stochastic model (1) and (15) yields the following expression for the steady-state Fano factor of $y(t)$ in the limit of fast binding/unbinding ($k_b \to \infty$, $k_u \to \infty$ with $k_d = k_u/k_b$ being finite),

$$\begin{aligned} F_y &= \frac{\overline{\langle y^2 \rangle} - \overline{\langle y \rangle}^2}{\overline{\langle y \rangle}} \\ &= \frac{\langle B_y^2 \rangle + \langle B_y \rangle}{2\langle B_y \rangle} + \frac{F_{x_{f,0}} \overline{\langle y \rangle}\, \overline{\langle x_{f,0} \rangle} \gamma_y}{(\overline{\langle x_{f,0} \rangle} - \beta f^2 N)[(Nf(1-f) + \overline{\langle x_f \rangle})\gamma_y + (\overline{\langle x_{f,0} \rangle} - \beta f^2 N)\gamma_f]}, \end{aligned} \tag{16}$$

where $\overline{\langle y \rangle} = \overline{\langle x_f \rangle} k_y \langle B_y \rangle / \gamma_y$ is the steady-state mean level of the target protein, $\langle B_y \rangle$ is the average burst size, and $\langle B_y^2 \rangle$ is the second-order moment of the burst size. Note that this noise level is made up of two components – the first component on the right-hand-side of (16) represents the noise from bursting of the target protein, and the second component is the noise in Y due to upstream fluctuations in the free TF count. In the absence of any decoy, (16) reduces to

$$F_{y,0} = \frac{\langle B_y^2 \rangle + \langle B_y \rangle}{2\langle B_y \rangle} + \frac{F_{x_{f,0}} \langle B_y \rangle k_y}{\gamma_y + \gamma_f}. \tag{17}$$

Our analysis shows that when bound TFs are stable ($\beta \to 0$), the second noise component in F_y monotonically decreases to zero with increasing decoy numbers. For $\beta > 0$, similar to the free TF count, noise enhancement in the target protein occurs for a small number of high-affinity decoys (see Fig. S2). However, both the magnitude ($F_y/F_{y,0}$)and the region of the noise enhancement are smaller compared to that of the free TF count (see Fig. S2 and compare with Fig. 3).

Noise propagation can be measured by the ratio of the target protein noise to the free TF noise, F_y/F_{x_f}. Figure 5C shows the noise propagation as a function of decoy numbers for different decay rates of bound TF. Here the y-axis intercept quantifies noise propagation in the absence of decoys. When the bound TFs are protected from degradation, noise propagation is first enhanced as expected from the right-shift of the autocorrelation function (Fig. 5A), but them sharply decreases at higher decoy abundances. Note that the increase in noise propagation seen at intermediate decoy numbers does not imply an increase in the target protein noise level, as at the same time the free TF noise level decreases with increasing N for $\beta = 0$ (Fig. 3). When bound TFs are unstable, noise propagation is reduced (Fig. 5C) due to faster time-scale of fluctuations of the free TF count (Fig. 5B). This implies that the noise increase seen in the free TF level when $\beta > 0$ (Fig. 3) is buffered by a decreased noise propagation, and hence, the region of noise enhancement for the target protein is reduced compared to that of the free TF count. The qualitative feature of noise propagation agrees with the stochastic simulations results. However, the quantitative match between stochastic simulations and the LNA results is poor for $\beta > 0$ (Fig. 5C).

Discussion

The sequestration of transcription factors by genomic decoy sites can either protect it from degradation[43,44] or make it more facile for degradation[45]. Here, we have systematically investigated how the magnitude and time-scale of TF copy-number fluctuations are impacted by the stability of the bound TF, the number and affinity of decoy sites. While this contribution focuses on TFs binding to genomic decoy sites, these results are applicable to other classes of proteins, for example, RNA-binding proteins binding to sites on RNA[115–118].

Our results show that the degradation of decoy-bound TFs critically impacts both the mean and noise levels of the free TF pool. More specifically, while the average number of free (available) TFs monotonically decreases with increasing decoy abundance, the noise levels can sharply increase at low/intermediate decoy numbers before attenuating to Poisson levels as $N \to \infty$ (Fig. 2). This behavior can be contrasted to when bound TFs are protected from degradation, in which case the mean free TF count becomes invariant of N, and decoys always buffer noise (Figs. 2 and 3). When $\beta > 0$, high-affinity decoys can transition from being noise amplifiers to noise buffers as their numbers are increased (Fig. 2). This point is exemplified in Fig. 6 where high-affinity low-abundance decoys and low-affinity high-abundance decoys result in the same average number of freely available TFs, but with much higher noise in the former case. Moreover, a mixture of both decoy types can mitigate each other's effect leading to no noise enhancement or buffer (Fig. 4).

Closed-form formulas for the noise levels derived using the Linear Noise Approximation led to precise conditions for noise enhancement – the number of decoys is not large, and their binding affinity is strong enough such that the fraction of bound decoys is higher than $1/(\beta+1)$. For example, $\beta = 1$ for a stable TF whose turnover is primarily governed by dilution from cell growth, and this condition implies that more than 50% decoy sites must be occupied. The decoy-mediated noise enhancement is higher for large values of β and for more "sticky" or higher affinity decoys (Fig. 3). It is interesting to point out that the peak noise enhancement generally occurs when the number of decoy sites is comparable to the TF counts when both bound and free TFs decay with the same rate. For example, in Fig. 2C there is an average of 200 TF molecules in the absence of decoys and the noise peak is also seen around $N \approx 200$. Note that our results are restricted to TF production in stochastic bursts, and it remains to be seen if these results generalize to other noise mechanisms in gene expression, such as, extrinsic noise that arises from intercellular differences in cell size and the abundance of expression machinery[98,119–122].

The speed of fluctuation in free TF counts also shows distinct features in the presence and absence of bound TF degradation. While the fluctuations become faster in the presence of bound TF degradation, as indicated by the left-shift of the autocorrelation function (Fig. 5), we see an opposite result when bound TFs do not degrade. These results have important implications on how noise propagates from the TF to downstream target proteins. For example, when $\beta > 0$ decoys can amplify noise in the free TF levels, while at the same time make fluctuations in TF counts relax back faster attenuating downstream noise transmission. As a result of these two opposing forces, the noise in the target protein may not increase even though the noise in the free TF level has increased. This can be seen for $\beta = 5$ in Fig. 3 where for $N = 100$ and $k_d = 10$ the decoy is a noise amplifier, but as seen in Fig. S2, for the exact same parameter values the noise in the target protein has reduced compared to the no-decoy case.

The analytical formulas for the noise give theoretical insights into the role decoys on the noise enhancement. These are derived by linearizing the propensity for the binding event assuming small copy number fluctuations around the statistical mean and then taking fast bind/binding limit. Using numerical exact stochastic simulations, we have shown that the main results are good agreement with the theory. The quantitative match between theory and simulations can be poor where the fluctuations are large. This happens for higher decoy binding affinities and degradation rates for bound TFs. We numerically have found that results are not much sensitive to fast binding/unbinding limit (Fig. S1). The objective of obtaining more accurate theoretical results is a matter of future work.

Our novel finding of the dual role of decoys as noise amplifiers/buffers encourages more investigation into the regulatory function of decoys in complex gene networks. In our study, we have considered target gene dose-response is linear. In future, we want to investigate the role of decoy sites in gene networks that contain non-linearity. For example, in genetic/signalling circuits with oscillatory dynamics decoys can tune the oscillation frequency[48], enhance coherence[47], and it will be interesting to see if decoys can make biological clocks more robust to molecular noise. An exciting avenue of experimental research is to use decoys as manipulations of phenotypic

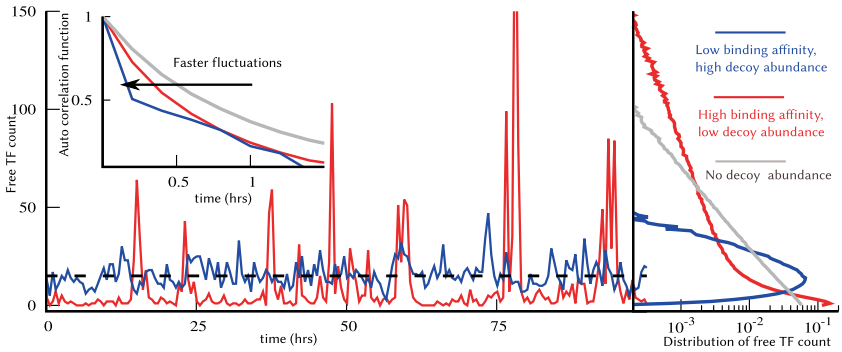

Figure 6. High-affinity low-abundance decoys, and low-affinity high-abundance decoys result in the same mean free TF counts with contrasting fluctuation dynamics. Stochastic realizations of the number of free TFs for high-affinity low-abundance decoys (red), and low-affinity high-abundance decoys (blue) when $\beta = 1$, along with their corresponding probability distributions on the right. Both scenarios yield the same average number of free TF molecules, but high-affinity decoys drive significantly enhanced noise levels. The instability of the bound TF leads to faster time-scale of fluctuations as illustrated by the left-shift of the autocorrelation function (inset). For comparison, the free TF count distribution and the autocorrelation function for the no-decoy case are also shown with grey lines. The parameters used are as follows: $\langle B_x \rangle = 20$, $k_x = 10\,hr^{-1}$, $\gamma_f = \gamma_b = 1.0\,hr^{-1}$ per protein molecule and $k_b = 50\,hr^{-1}$ per pair of molecules. For high affinity decoy $k_d = 1$ and $N = 245$, and low affinity decoy $k_d = 100$ and $N = 1400$. This choice produces the same mean TF count for both the cases, $\overline{\langle x_f \rangle} \simeq 15$. For the no-decoy case: $k_x = 0.75\,hr^{-1}$ to keep $\overline{\langle x_f \rangle} \simeq 15$.

heterogeneity. For example, aberrant expression of resistance markers in individual melanoma cells have been shown to be a driver of cancer drug resistance[5], and decoy-mediated noise buffering can play a therapeutic role in reducing such outlier cells. In the context of the Human Immunodeficiency Virus (HIV), noisy expression of a key viral protein, Tat, controls the cell-fate outcome between active viral replication and latency[16,123–126]. Latency is a dormant state of the virus and considered the biggest challenge in eradicating the virus from a patient since latent cells are long-lived and resistant to drug therapy[127]. Recently, small-molecule compounds have been identified that enhance noise in Tat expression for efficient reactivation of latent cells[128], and here the role of decoys as noise amplifiers may allow for a Tat-specific compound-free approach for latency reversal.

Received: 24 December 2019; Accepted: 8 May 2020;
Published online: 04 June 2020

References

1. McAdams, H. H. & Arkin, A. It's a noisy business! Genetic regulation at the nanomolar scale. *Trends in genetics* **15**, 65–9 (1999).
2. Arkin, A. P., Ross, J. & McAdams, H. H. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ-infected *Escherichia coli* cells. *Genetics* **149**, 1633–1648 (1998).
3. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).
4. Losick, R. & Desplan, C. Stochasticity and cell fate. *Science* **320**, 65–68 (2008).
5. Shaffer, S. M. *et al.* Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature* **546**, 431–435 (2017).
6. Billman, M., Rueda, D. & Bangham, C. Single-cell heterogeneity and cell-cycle-related viral gene bursts in the human leukaemia virus htlv-1. *Wellcome Open Research* **2**, 87 (2017).
7. Keskin, S. *et al.* Noise in the vertebrate segmentation clock is boosted by time delays but tamed by notch signaling. *Cell Rep* **23**, 2175–2185 (2018).
8. Urban, E. A. & Johnston, R. J. Buffering and amplifying transcriptional noise during cell fate specification. *Frontiers in Genetics* **9**, 591 (2018).
9. Raj, A. & van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216–226 (2008).
10. von Dassow, G., Meir, E., Munro, E. M. & Odell, G. M. The segment polarity network is a robust developmental module. *Nature* **406**, 188 (2000).
11. Queitsch, C., Sangster, T. A. & Lindquist, S. Hsp90 as a capacitor of phenotypic variation. *Nature* **417**, 618 (2002).
12. Schmiedel, J. M., Carey, L. B. & Lehner, B. Empirical mean-noise fitness landscapes reveal the fitness impact of gene expression noise. *Nature Communications* **10**, 3180 (2019).
13. Kemkemer, R., Schrank, S., Vogel, W., Gruler, H. & Kaufmann, D. Increased noise as an effect of haploinsufficiency of the tumor-suppressor gene neurofibromatosis type 1 *in vitro*. *Proceedings of the National Academy of Sciences* **99**, 13783–13788 (2002).
14. Bahar, R. *et al.* Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature* **441**, 1011–1014 (2006).
15. Kussell, E. & Leibler, S. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309**, 2075–2078 (2005).
16. Weinberger, L. S., Burnett, J., Toettcher, J., Arkin, A. & Schaffer, D. Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell* **122**, 169–182 (2005).
17. Kærn, M., Elston, T. C., Blake, W. J. & Collins, J. J. Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics* **6**, 451–464 (2005).
18. Blake, W. J. *et al.* Phenotypic consequences of promoter-mediated transcriptional noise. *Molecular Cell* **24**, 853–865 (2006).
19. Davidson, C. J. & Surette, M. G. Individuality in bacteria. *Annual Review of Genetics* **42**, 253–268 (2008).
20. Fraser, D. & Kaern, M. A chance at survival: gene expression noise and phenotypic diversification strategies. *Mol Microbiol* **71**, 1333–1340 (2009).

21. Rotem, E. *et al*. Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. *Proceedings of the National Academy of Sciences* **107**, 12541–12546 (2010).
22. Schreiber, F. *et al*. Phenotypic heterogeneity driven by nutrient limitation promotes growth in fluctuating environments. *Nature Microbiology* **1**, 16055 (2016).
23. Maamar, H., Raj, A. & Dubnau, D. Noise in gene expression determines cell fate in bacillus subtilis. *Science* **317**, 526–529 (2007).
24. Balázsi, G., van Oudenaarden, A. & Collins, J. J. Cellular decision making and biological noise: From microbes to mammals. *Cell* **144**, 910–925 (2014).
25. Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nature Genetics* **31**, 69–73 (2002).
26. Singh, A. & Hespanha, J. P. Optimal feedback strength for noise suppression in autoregulatory gene networks. *Biophysical Journal* **96**, 4013–4023 (2009).
27. Dublanche, Y., Michalodimitrakis, K., Kummerer, N., Foglierini, M. & Serrano, L. Noise in transcription negative feedback loops: simulation and experimental analysis. *Molecular Systems Biology* **2**, 41 (2006).
28. Blake, W. J., Kaern, M., Cantor, C. R. & Collins, J. J. Noise in eukaryotic gene expression. *Nature* **422**, 633–637 (2003).
29. Alon, U. An Introduction to Systems Biology: Design Principles of Biological Circuits (Chapman and Hall/CRC, New York, 2011).
30. Bintu, L. *et al*. Transcriptional regulation by the numbers: applications. *Current Opinion in Genetics & Development* **15**, 125–135 (2005).
31. Sánchez, Á. & Kondev, J. Transcriptional control of noise in gene expression. *Proceedings of the National Academy of Sciences* **105**, 5081–5086 (2008).
32. Wunderlich, Z. & Mirny, L. A. Different gene regulation strategies revealed by analysis of binding motifs. *Trends in genetics* **25**, 434–440 (2009).
33. Kemme, C. A., Nguyen, D., Chattopadhyay, A. & Iwahara, J. Regulation of transcription factors via natural decoys in genomic dna. *Transcription* **7**, 115–120 (2016).
34. Esadze, A., Kemme, C. A., Kolomeisky, A. B. & Iwahara, J. Positive and negative impacts of nonspecific sites during target location by a sequence-specific dna-binding protein: origin of the optimal search at physiological ionic strength. *Nucleic Acids Research* **42**, 7039–7046 (2014).
35. Kemme, C. A., Esadze, A. & Iwahara, J. Influence of quasi-specific sites on kinetics of target dna search by a sequence-specific dna-binding protein. *Biochemistry* **54**, 6684–6691 (2015).
36. Bakk, A. & Metzler, R. *In vivo* non-specific binding of λ ci and cro repressors is significant. *FEBS Letters* **563**, 66–68 (2004).
37. Lee, T. & Maheshri, N. A regulatory role for repeated decoy transcription factor binding sites in target gene expression. *Molecular systems biology* **8**, 576 (2012).
38. Morishita, R. *et al*. A gene therapy strategy using a transcription factor decoy of the e2f binding site inhibits smooth muscle proliferation *in vivo*. *Proceedings of the National Academy of Sciences* **92**, 5855–5859 (1995).
39. Mann, M. J. Transcription factor decoys: A new model for disease intervention. *Annals of the New York Academy of Sciences* **1058**, 128–139 (2005).
40. Hecker, M. & Wagner, A. H. Transcription factor decoy technology: A therapeutic update. *Biochemical Pharmacology* **144**, 29–34 (2017).
41. Francois, M., Donovan, P. & Fontaine, F. Modulating transcription factor activity: Interfering with protein-protein interaction networks. *Seminars in Cell and Developmental Biology* (2018).
42. Burger, A., Walczak, A. M. & Wolynes, P. G. Abduction and asylum in the lives of transcription factors. *Proceedings of the National Academy of Sciences* **107**, 4016–4021 (2010).
43. Abu Hatoum, O. *et al*. Degradation of myogenic transcription factor myod by the ubiquitin pathway *in vivo* and *in vitro*: Regulation by specific dna binding. *Molecular and Cellular Biology* **18**, 5670–5677 (1998).
44. Molinari, E., Gilman, M. & Natesan, S. Proteasome-mediated degradation of transcriptional activators correlates with activation domain potency *in vivo*. *EMBO J* **18**, 6439–6447 (1999).
45. Thomas, D. & Tyers, M. Transcriptional regulation: Kamikaze activators. *Current Biology* **10**, R341–R343 (2000).
46. Burger, A., Walczak, A. M. & Wolynes, P. G. Influence of decoys on the noise and dynamics of gene expression. *Physical Review E* **86**, 041920 (2012).
47. Wang, Z., Potoyan, D. A. & Wolynes, P. G. Molecular stripping, targets and decoys as modulators of oscillations in the nf-kb/ikb α/dna genetic network. *Journal of The Royal Society Interface* **13**, 20160606 (2016).
48. Jayanthi, S. & Del Vecchio, D. Tuning genetic clocks employing DNA binding sites. *PLOS ONE* **7**, e41019 (2012).
49. Jayanthi, S., Nilgiriwala, K. S. & Del Vecchio, D. Retroactivity controls the temporal dynamics of gene transcription. *ACS Synthetic Biology* **2**, 431–441 (2013).
50. Ricci, F., Vallée-Bélisle, A. & Plaxco, K. W. High-precision, *in vitro* validation of the sequestration mechanism for generating ultrasensitive dose-response curves in regulatory networks. *PLOS Computational Biology* **7**, e1002171 (2011).
51. Jones, D. L., Brewster, R. C. & Phillips, R. Promoter architecture dictates cell-to-cell variability in gene expression. *Science* **346**, 1533–1537 (2014).
52. Soltani, M., Bokes, P., Fox, Z. & Singh, A. Nonspecific transcription factor binding can reduce noise in the expression of downstream proteins. *Physical Biology* **12**, 055002 (2015).
53. Bokes, P. & Singh, A. Protein copy number distributions for a self-regulating gene in the presence of decoy binding sites. *PLOS ONE* **10**, e0120555 (2015).
54. Das, D., Dey, S., Brewster, R. C. & Choube, S. Effect of transcription factor resource sharing on gene expression noise. *PLoS Comput Biol* **13**, e1005491 (2017).
55. Razo-Mejia, M. *et al*. Tuning transcriptional regulation through signaling: A predictive theory of allosteric induction. *Cell Systems* **6**, 456–469 (2018).
56. Suter, D. M. *et al*. Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**, 472–474 (2011).
57. Dar, R. D. *et al*. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proceedings of the National Academy of Sciences* **109**, 17454–17459 (2012).
58. Fukaya, T., Lim, B. & Levine, M. Enhancer control of transcriptional bursting. *Cell* **166**, 358–368 (2015).
59. Bartman, C. R., Hsu, S. C., Hsiung, C. C.-S., Raj, A. & Blobel, G. A. Enhancer regulation of transcriptional bursting parameters revealed by forced chromatin looping. *Molecular Cell* **62**, 237–247 (2016).
60. Corrigan, A. M., Tunnacliffe, E., Cannon, D. & Chubb, J. R. A continuum model of transcriptional bursting. *eLife* **5**, e13051 (2016).
61. Chong, S., Chen, C., Ge, H. & Xie, X. S. Mechanism of transcriptional bursting in bacteria. *Cell* **158**, 314–326 (2014).
62. Singh, A., Razooky, B., Cox, C. D., Simpson, M. L. & Weinberger, L. S. Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression. *Biophysical Journal* **98**, L32–L34 (2010).
63. Dar, R. D. *et al*. Transcriptional bursting explains the noise–versus–mean relationship in mRNA and protein levels. *PLOS ONE* **11**, e0158298 (2016).
64. Golding, I., Paulsson, J., Zawilski, S. & Cox, E. Real-time kinetics of gene activity in individual bacteria. *Cell* **123**, 1025–1036 (2005).
65. Raj, A., Peskin, C., Tranchina, D., Vargas, D. & Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLOS Biology* **4**, e309 (2006).

66. Singh, A., Razooky, B. S., Dar, R. D. & Weinberger, L. S. Dynamics of protein noise can distinguish between alternate sources of gene-expression variability. *Molecular Systems Biology* **8**, 607 (2012).
67. Thattai, M. & van Oudenaarden, A. Intrinsic noise in gene regulatory networks. *Proceedings of the National Academy of Sciences* **98**, 8614–8619 (2001).
68. Friedman, N., Cai, L. & Xie, X. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Physical Review Letters* **97**, 168302 (2006).
69. Shahrezaei, V. & Swain, P. S. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences* **105**, 17256–17261 (2008).
70. Pedraza, J. M. & Paulsson, J. Effects of molecular memory and bursting on fluctuations in gene expression. *Science* **319**, 339–343 (2008).
71. Jia, T. & Kulkarni, R. V. Intrinsic noise in stochastic models of gene expression with molecular memory and bursting. *Journal of Mathematical Biology* **106**, 058102 (2011).
72. Kumar, N., Singh, A. & Kulkarni, R. V. Transcriptional bursting in gene expression: Analytical results for genera stochastic models. *PLOS Computational Biology* **11**, e1004292 (2015).
73. Bokes, P. & Singh, A. Gene expression noise is affected deferentially by feedback in burst frequency and burst size. *Journal of Mathematical Biology* **74**, 1483–1509 (2017).
74. Singh, A. & Soltani, M. Quantifying intrinsic and extrinsic variability in stochastic gene expression models. *PLOS ONE* **8**, e84301 (2013).
75. Soltani, M., Vargas-Garcia, C. A., Antunes, D. & Singh, A. Intercellular variability in protein levels from stochastic expression and noisy cell cycle processes. *PLOS Computational Biology* e1004972 (2016).
76. Yu, J., Xiao, J., Ren, X., Lao, K. & Xie, X. S. Probing gene expression in live cells, one protein molecule at a time. *Science* **311**, 1600–1603 (2006).
77. Paulsson, J. Model of stochastic gene expression. *Physics of Life Reviews* **2**, 157–175 (2005).
78. Elgart, V., Jia, T., Fenley, A. T. & Kulkarni, R. Connecting protein and mrna burst distributions for stochastic models of gene expression. *Physical biology* **8**, 046001 (2011).
79. Wilkinson, D. J. *Stochastic Modelling for Systems Biology* (Chapman and Hall/CRC, 2011).
80. McQuarrie, D. A. Stochastic approach to chemical kinetics. *Journal of Applied Probability* **4**, 413–478 (1967).
81. Munsky, B. & Khammash, M. The finite state projection algorithm for the solution of the chemical master equation. *Journal of Chemical Physics* **124**, 044104 (2006).
82. Gupta, A., Mikelson, J. & Khammash, M. A finite state projection algorithm for the stationary solution of the chemical master equation. *The Journal of Chemical Physics* **147**, 154101 (2017).
83. Dinh, K. N. & Sidje, R. B. Understanding the finite state projection and related methods for solving the chemical master equation. *Physical Biology* **13**, 035003 (2016).
84. Gillespie, D. T. Approximate accelerated stochastic simulation of chemically reacting systems. *Journal of Chemical Physics* **115**, 1716–1733 (2001).
85. Gibson, M. A. & Bruck, J. Efficient exact stochastic simulation of chemical systems with many species and many channels. *Journal of Physical Chemistry A* **104**, 1876–1889 (2000).
86. Cao, Y., Li, H. & Petzold, L. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *Journal of Chemical Physics* **121**, 4059–4067 (2004).
87. Anderson, D. F. A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *Journal of Chemical Physics* **127**, 214107 (2007).
88. Daigle, B., Soltani, M., Petzold, L. & Singh, A. Inferring single-cell gene expression mechanisms using stochastic simulation. *Bioinformatics* **31**, 1428–1435 (2015).
89. Van Kampen, N. *Stochastic Processes in Physics and Chemistry* (Elsevier, 2011).
90. Elf, J. & Ehrenberg, M. Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome Research* **13**, 2475–2484 (2003).
91. Lestas, I., Paulsson, J., Ross, N. E. & Vinnicombe, G. Noise in gene regulatory networks. *IEEE Transactions on Automatic Control* **53**, 189–200 (2008).
92. Modi, S., Soltani, M. & Singh, A. Linear noise approximation for a class of piecewise deterministic markov processes. In *2018 Annual American Control Conference (ACC)*, 1993–1998 (2018).
93. Munsky, B., Hlavacek, W. S. & Tsimring, L. S. *Quantitative biology: theory, computational methods, and models* (The MIT Press, 2018).
94. Thomas, P., Straube, A. V. & Grima, R. The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions. *BMC Systems Biology* **6**, 39 (2012).
95. Ochab-Marcinek, A. J. E., Drak, J. & Tabaka, M. Hill kinetics as a noise filter: the role of transcription factor autoregulation in gene cascades. *Phys. Chem. Chem. Phys.* **19**, 22580–22591 (2017).
96. Czuppon, P. & Pfaffelhuber, P. Limits of noise for autoregulated gene expression. *Journal of Mathematical Biology* **77**, 1153–1191 (2018).
97. Sepúlveda, L. A., Xu, H., Zhang, J., Wang, M. & Golding, I. Measurement of gene regulation in individual cells reveals rapid switching between promoter states. *Science* **351**, 1218–1222 (2016).
98. Singh, A. Transient changes in intercellular protein variability identify sources of noise in gene expression. *Biophysical Journal* **107**, 2214–2220 (2014).
99. Singh, A. & Hespanha, J. P. Stochastic hybrid systems for studying biochemical processes. *Philosophical Transactions of the Royal Society A* **368**, 4995–5011 (2010).
100. Singh, A. & Hespanha, J. P. Approximate moment dynamics for chemically reacting systems. *IEEE Transactions on Automatic Control* **56**, 414–418 (2011).
101. Gomez-Uribe, C. A. & Verghese, G. C. Mass fluctuation kinetics: Capturing stochastic effects in systems of chemical reactions through coupled mean-variance computations. *Journal of Chemical Physics* **126**, 024109 (2007).
102. Lee, C. H., Kim, K. & Kim, P. A moment closure method for stochastic reaction networks. *Journal of Chemical Physics* **130**, 134107 (2009).
103. Goutsias, J. Classical versus stochastic kinetics modeling of biochemical reaction systems. *Biophysical Journal* **92**, 2350–2365 (2007).
104. Gillespie, C. S. Moment closure approximations for mass-action models. *IET Systems Biology* **3**, 52–58 (2009).
105. Soltani, M., Vargas, C. & Singh, A. Conditional moment closure schemes for studying stochastic dynamics of genetic circuits. *IEEE Transactions on Biomedical Systems and Circuits* **9**, 518–526 (2015).
106. Zhang, J., DeVille, L., Dhople, S. & Dominguez-Garcia, A. A maximum entropy approach to the moment closure problem for stochastic hybrid systems at equilibrium. In *IEEE Conference on Decision and Control*, 747–752 (2014).
107. Smadbeck, P. & Kaznessis, Y. N. A closure scheme for chemical master equations. *Proceedings of the National Academy of Sciences* **110**, 14261–14265 (2013).
108. Schnoerr, D., Sanguinetti, G. & Grima, R. Validity conditions for moment closure approximations in stochastic chemical kinetics. *The Journal of Chemical Physicsl* **141**, 084103 (2014).

109. Lakatos, E., Ale, A., Kirk, P. D. W. & Stumpf, M. P. H. Multivariate moment closure techniques for stochastic kinetic models. *The Journal of Chemical Physics* **143**, 094107 (2015).
110. Lamperski, A., Ghusinga, K. R. & Singh, A. Stochastic optimal control using semidefinite programming for moment dynamics. *Proc. of the 55th IEEE Conf. on Decision and Control, Las Vegas* 1990–1995 (2016).
111. Ghusinga, K. R., Vargas-Garcia, C. A., Lamperski, A. & Singh, A. Exact lower and upper bounds on stationary moments in stochastic biochemical systems. *Physical Biology* **14**, 04LT01 (2017).
112. Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics* **22**, 403–434 (1976).
113. Paulsson, J. Summing up the noise in gene networks. *Nature (London)* **427**, 415–418 (2004).
114. Singh, A. & Bokes, P. Consequences of mRNA transport on stochastic variability in protein levels. *Biophysical Journal* **103**, 1087–1096 (2012).
115. Hooykaas, M. J. G. *et al.* Rna accessibility impacts potency of tough decoy microrna inhibitors. *RNA Biology* **15**, 1410–1419 (2018).
116. Parra, M. *et al.* An important class of intron retention events in human erythroblasts is regulated by cryptic exons proposed to function as splicing decoys. *RNA* **24**, 1255–1265 (2018).
117. Howard, J. M. *et al.* HNRNPA1 promotes recognition of splice site decoys by U2AF2 *in vivo*. *Genome Research* **28**, 689–698 (2018).
118. Denichenko, P. *et al.* Specific inhibition of splicing factor activity by decoy RNA oligonucleotides. *Nature Communications* **10**, 1590 (2019).
119. Padovan-Merhar, O. *et al.* Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. *Molecular Cell* **58**, 339–352 (2015).
120. Johnston, I. G. *et al.* Mitochondrial variability as a source of extrinsic cellular noise. *PLOS Computational Biology* **8**, e1002416 (2012).
121. Shahrezaei, V., Ollivier, J. F. & Swain, P. S. Colored extrinsic fluctuations and stochastic gene expression. *Molecular Systems Biology* **4** (2008).
122. Hilfinger, A. & Paulsson, J. Separating intrinsic from extrinsic fluctuations in dynamic biological systems. *Proceedings of the National Academy of Sciences* **108**, 12167–12172 (2011).
123. Razooky, B. S., Pai, A., Aull, K., Rouzine, I. M. & Weinberger, L. S. A hardwired HIV latency program. *Cell* **160**, 990–1001 (2015).
124. Chavez, L., Calvanese, V. & Verdin, E. HIV latency is established directly and early in both resting and activated primary CD4 T cells. *PLOS Pathogens* **11**, e1004955 (2015).
125. Singh, A. & Weinberger, L. S. Stochastic gene expression as a molecular switch for viral latency. *Current Opinion in Microbiology* **12**, 460–466 (2009).
126. Singh, A. Stochastic analysis of genetic feedback circuit controlling HIV cell-fate decision. *Proc. of the 51st IEEE Conf. on Decision and Control, Maui, Hawaii* 4918–4923 (2012).
127. Richman, D. D. *et al.* The challenge of finding a cure for HIV infection. *Science* **323**, 1304–1307 (2009).
128. Dar, R. D., Hosmane, N. N., Arkin, M. R., Siliciano, R. F. & Weinberger, L. S. Screening for noise in gene expression identifies drug synergies. *Science* **344**, 1392–1396 (2014).

Acknowledgements

A.S. is supported by the NSF grant ECCS-1711548, and NIH grants 5R01GM124446 and 5R01GM126557.

Author contributions

A.S. defined and supervised the project. S.D. and M.S. did the mathematical derivations. S.D. performed the necessary computations. The authors discussed the results and collaborated on the writing, read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-65750-2.

Correspondence and requests for materials should be addressed to S.D. or A.S.

Reprints and permissions information is available at www.nature.com/reprints.

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