Disentangling Intrinsic and Extrinsic Gene Expression Noise in Growing Cells

Jie Lin and Ariel Amir

Phys. Rev. Lett. **126**, 078101 — Published 16 February 2021

DOI: 10.1103/PhysRevLett.126.078101

Disentangling intrinsic and extrinsic gene expression noise in growing cells

Jie Lin[1,2] and Ariel Amir[2]

[1] *Center for Quantitative Biology and Peking-Tsinghua Joint Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China*
[2] *John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA*

(Dated: January 14, 2021)

Gene expression is a stochastic process. Despite the increase of protein numbers in growing cells, the protein concentrations are often found to be confined within small ranges throughout the cell cycle. Generally, the noise in protein concentration can be decomposed into an intrinsic and an extrinsic component, where the former vanishes for high expression levels. Considering the time trajectory of protein concentration as a random walker in the concentration space, an effective restoring force (with a corresponding "spring constant") must exist to prevent the divergence of concentration due to random fluctuations. In this work, we prove that the magnitude of the effective spring constant is directly related to the fraction of intrinsic noise in the total protein concentration noise. We show that one can infer the magnitude of intrinsic, extrinsic, and measurement noises of gene expression solely based on time-resolved data of protein concentration, without any a priori knowledge of the underlying gene expression dynamics. We apply this method to experimental data of single-cell bacterial gene expression. The results allow us to estimate the average copy numbers and the translation burst parameters of the studied proteins.

INTRODUCTION

Gene expression in all forms of life is subject to noise [1–7]. Experimentally, stochastic gene expression has been intensively studied, mostly in growing cells with exponentially growing cell volume [8–12] in which the copy numbers of mRNAs and proteins in general double on average during the cell cycle, as widely observed in bacterial and eukaryotic cells [8, 13–15]. To reduce cell cycle effects, a more biologically relevant protocol to quantify the stochastic degree of gene expression is to calculate the variability of *concentration* because most genes in proliferating cells exhibit approximately constant protein concentrations throughout the cell cycle over multiple generations [13, 16–21]. In yeast and mammalian cells, most genes also exhibit approximately constant mRNA concentrations throughout the cell cycle [14, 22, 23].

Considering the time trajectory of protein concentration as a one dimensional random walker in the space of concentration, it must be subject to an effective restoring force to prevent the divergence of concentration in the long time limit (note that cell growth contributes to this restoring force via the effect of dilution, as discussed extensively in Ref. [19]). However, little is known about how the strength of this restoring force is related to the stochastic nature of protein concentration. In this work we show that one can in fact infer the contribution of intrinsic and extrinsic noise (which we will define later) to the total gene expression noise from the properties of the restoring force. Previous works on solving this challenge often rely on particular models of the underlying dynamics of gene expression [24–27]. Here we develop a novel protocol which is, in contrast, insensitive to many of the details of the gene expression dynamics, and is thus applicable to a broad class of models. The protocol only relies on analysis of time-series data of protein concentrations. We expect it to be applicable to exponentially growing cells such as bacteria, yeast and cancer cells [8–12].

In the following, we first introduce a general framework to study the variability of mRNA and protein concentrations in growing cells. Within the framework, the initiation rates of transcription and translation can be age-dependent (here, we define age as the elapsed time since cell birth), *e.g.*, due to gene dosage effects as well as more complex cell cycle dependencies [15]. We show that independent of the details of the gene expression dynamics, the variances of mRNA and protein concentrations can always be decomposed into an extrinsic component and an intrinsic component. In the large cell volume limit, the intrinsic noise vanishes while the extrinsic noise remains finite [28]. We then introduce our protocol to extract the fraction of intrinsic noise, extrinsic noise and measurement noise in the total noise of protein concentrations and finally apply the method to experimental data of bacterial gene expression.

Decomposition of noise.—For simplicity, we consider a cell growing exponentially at a constant growth rate μ with a constant doubling time $T = \ln(2)/\mu$, see Fig. 1. When the cell divides, the cell volume divides symmetrically, therefore the molecules are assumed to be segregated binomially and symmetrically between the two daughter cells [3]. Since for both bacterial and eukaryotic cells the degradation times of many proteins are longer than the cell cycle duration [29], we consider a non-degradable protein in the main text. Our results are equally valid for proteins with a finite degradation rate after some slight modifications (Supplementary Information, SI A) [30]. Our results are also robust against

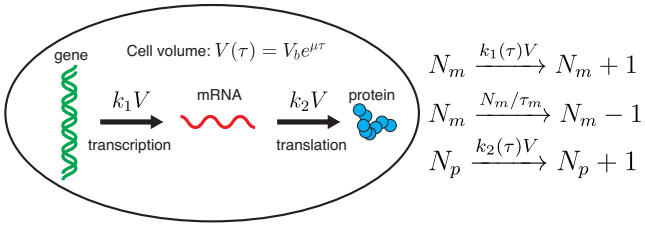


FIG. 1. The cell volume V grows exponentially in time with a growth rate μ and τ is the cell age. k_1 and k_2 are the transcription rate and translation rate per cell volume which can be age-dependent. The chemical reactions of gene expression are summarized on the right. N_m and N_p are the absolute mRNA and protein copy numbers respectively. τ_m is the lifetime of mRNA using which one can define the translation burst parameter $\beta\tau_m$ (the average number of proteins produced in the lifetime of a single mRNA).

fluctuating growth rates and doubling times as we show in SI B. We allow the initiation rates of transcription and translation per cell volume, k_1, k_2, to be time dependent and, for example, they can exhibit stochastic dynamics. One can further express $k_2 = \beta m$ where m is the mRNA concentration and β is the initiation rate of translation per mRNA. Mechanistically β is determined by the binding rate of ribosomes to mRNAs and largely determined by the concentration of ribosomes, which is roughly constant throughout the cell cycle [20].

Consider an experiment where one tracks a single lineage of cells over multiple generations, records the data of protein concentrations p uniformly in time with resolution Δt, and finally computes the resulting variance of concentrations based on all collected data. We find that the resulting variance of protein concentration σ_p^2 can be generally decomposed into three components (SI A):

$$\sigma_p^2 = \underbrace{\frac{\mathrm{cov}(k_2, p)}{\mu}}_{\text{Upstream noise}} + \underbrace{\left\langle \frac{k_2}{2\mu V} \right\rangle}_{\text{Poisson noise}} + \underbrace{\frac{\overline{p(T)}}{4\ln(2) V_b}}_{\text{Partitioning noise}} . \quad (1)$$

Here $\mathrm{cov}(k_2, p) = \langle k_2 p \rangle - \langle k_2 \rangle \langle p \rangle$ and $\langle \cdot\cdot \rangle$ represents average over time. The first part represents the noise due to a fluctuating upstream factor, namely, the initiation rate of translation per cell volume. One important source of upstream noise is the fluctuation in mRNA copy number [28]. The second term represents the noise due to the stochastic production process which we denote as Poisson noise here. The last term stems from the random partitioning during cell division where $T = \ln 2/\mu$ is the doubling time. The Poisson noise and the partitioning noise scale with the inverse of cell volume and their contributions to the square of the coefficient of variation (variance/mean^2) vanish for highly expressed proteins. In contrast, the upstream noise stems from the fluctuation in the translation rate per cell volume and it does not vanish in the large cell volume limit. We therefore define the sum of the Poisson noise and the partitioning noise as intrinsic and the upstream noise as extrinsic, consistent with previous works [28, 31]. We numerically confirm the validity of the noise decomposition for multiple gene expression dynamics including stochastic transcription and translation rate (SI B, Fig. S1).

We remark that the definition of extrinsic noise in our framework is different from the extrinsic noise inferred from the dual-reporter setup [1, 32], which is defined as the correlated noise of two identical genes controlled by the same promoters. The possible sources of extrinsic noise in the dual-reporter setup belong to a subset of those of the extrinsic noise in our framework which includes all possible upstream factors correlated or not across genes. Therefore, the extrinsic noise from the dual-reporter method is typically smaller than the extrinsic noise defined in our current framework, as we will discuss further later.

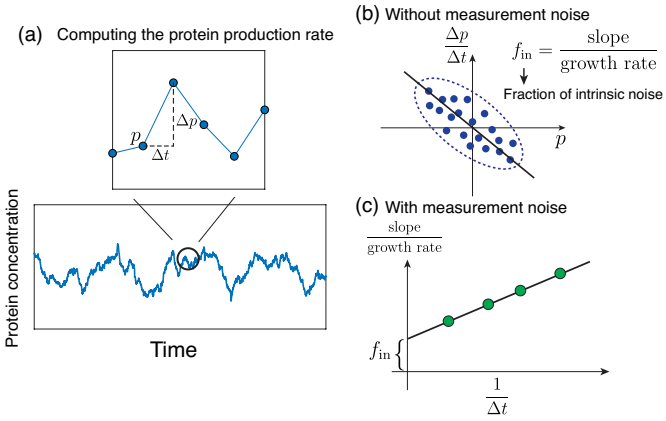


FIG. 2. (a) Given a time series of protein concentration, we first compute the discrete time derivative of protein concentration $\Delta p/\Delta t$ with a time interval Δt. (b) Next, we perform a linear fit of $\Delta p/\Delta t$ against the current protein concentration p and consider the absolute value of the fitted slope. In the case of negligible measurement noise, the fraction of intrinsic noise is the ratio between the slope and growth rate. (c) For experimental data with measurement noise, we compute $\Delta p/\Delta t$ for multiple time intervals Δt and repeat the protocol in (b) for each time interval. Finally, we perform a linear fit of the normalized slopes against $1/\Delta t$ and infer the fraction of intrinsic noise from the intercept.

Extracting the fraction of intrinsic and extrinsic noise.—In the following, we discuss a protocol to disentangle the contribution of intrinsic and extrinsic noise to the total noise based on the time trajectory of concentration [Fig. 2(a, b)]. We consider a discrete increment of protein concentration over a small time window, $\Delta p(t) = p(t + \Delta t) - p(t)$, which can be expressed as

$$\Delta p(t) = \frac{\Delta N_p(t)}{V(t)} - \mu p(t) \Delta t, \quad (2)$$

where $\Delta N_p(t)$ is a random variable from a Poisson distribution with mean $k_2(t) V \Delta t$ assumed constant within

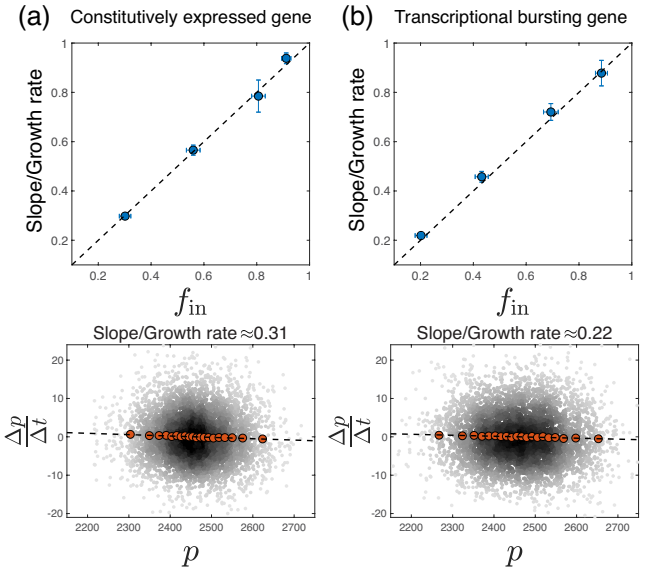


FIG. 3. (a) Simulation of a constitutively expressed gene. (Upper) We compare the predicted fraction of intrinsic noise (y axis) to the measured value (x axis). (Bottom) An example of the raw numerical data with the binned data shown as well (red circles). The dashed line is the linear fit of the raw data. The same analysis also applies to panel (b). Here $k_1 = 10$. (b) Simulation of a transcriptional bursting gene with $k_{\rm on} = 10$, $k_{\rm off} = 10$, $k_1 = 20$. In all upper panels, the doubling time $T = 60$, $\tau_m = 10$, and β is varied so that $\log_{10}\beta = -2, -1.5, -1, -0.5$. In all bottom panels, $\log_{10}\beta = -0.5$. We compute the time-derivative of protein concentration with a time interval $\Delta t = 0.5$. The errorbars are computed as the standard deviation of 5 independent simulations and in each simulation, 10^3 cell cycles are tracked.

the small time interval Δt. The second term on the right side arises from dilution due to cell growth. The protein concentration fluctuates but does not diverge in the long time limit, therefore we can make an analogy with a Brownian particle attracted to a fixed point with a linear restoring force equal to $-kx$ where k is the spring constant and x is the particle position relative to its equilibrium point. In the case of a Brownian particle, one can find the spring constant of the restoring force as the slope in the linear fitting of the discrete velocity $\Delta x/\Delta t$ *vs.* x. In the case of protein concentration, one can do a similar analysis by linearly fitting the discrete time derivative of protein concentration $\Delta p/\Delta t$ *vs.* p. Considering a least square linear fitting, the slope of the linear fitting is found to be

$$S \equiv -\frac{\mathrm{cov}\left(\frac{\Delta p(t)}{\Delta t}, p(t)\right)}{\sigma_p^2} = \mu - \frac{\mathrm{cov}(k_2(t), p(t))}{\sigma_p^2}. \qquad (3)$$

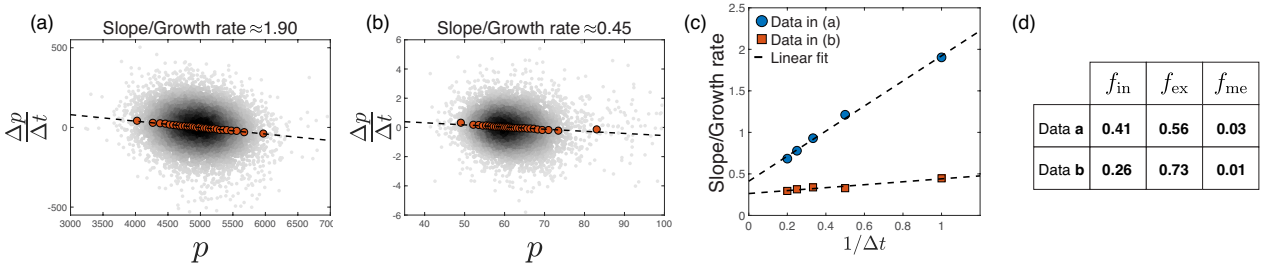
where we have used Eq. (2). If the covariance between the translation rate and protein concentration vanishes, the spring constant of the restoring force is simply the growth rate. Combined with Eq. (1), we find that the slope is proportional to the growth rate and the proportional constant is precisely the fraction of intrinsic noise in the total protein concentration noise variance:

$$S = \mu\left(1 - \frac{\mathrm{cov}(k_2(t), p(t))}{\mu\sigma_p^2}\right) = \mu f_{\rm in}. \qquad (4)$$

The above equation shows that we can extract the fraction of intrinsic noise $f_{\rm in}$ in the total noise by linearly fitting the time derivative of the protein concentration against the current protein concentration without any *a priori* knowledge of the underlying gene expression dynamics. Extrinsic noise reduces the slope in the linear fitting which precisely equals the growth rate μ in the absence of extrinsic noise. An extended discussion along with an intuitive argument on the effects of extrinsic noise based on a Langevin equation is provided in SI F. We remark that our protocols are also applicable to nongrowing cells with a constant cell volume given the lifetime of the studied protein is known (SI A).

Analysis of synthetic data.—We test Eq. (4) on synthetic data, first considering a constitutively expressed gene where the initiation rate of transcription per cell volume k_1 is constant as is the initiation rate of translation per mRNA β. This assumption corresponds to the case in which both RNA polymerase and ribosomes are limiting for gene expression, as discussed in detail in Ref. [19]. We compute $f_{\rm in}$ numerically using Eq. (1) and compare it with the prediction from Eq. (4), finding excellent agreement [Fig. 3(a)]. To test the robustness of our protocol, we also verify our theoretical results on various other gene expression dynamics: (1) the scenario of transcriptional bursting where a gene switches from "off" state to "on" state with rate $k_{\rm on}$ and vice versa with rate $k_{\rm off}$ [Fig. 3(b)]; (2) a gene with a constant transcription rate proportional to the gene number which doubles in the middle of the cell cycle [Fig. S2(a)]; this scenario corresponds to the situation when the gene copy number is the sole limiting factor of transcription [19]; (3) a gene with a transcription rate modulated throughout the cell cycle due to a finite period of DNA replication [Fig. S2(b), see details in SI E]; (4) a gene with a fluctuating transcription rate [Fig. S2(c)]; (5) a gene with a fluctuating translation rate per mRNA [Fig. S2(d)]. In all cases, the predicted fractions of intrinsic noise match the actual values well. We also find that in all cases increasing the translation rate per mRNA β increases the fraction of extrinsic noise as the effects of upstream noise are amplified, consistent with the analytical results of constitutively expressed genes (SI C, D). We have also confirmed the robustness of our results against the number of cell cycles sampled and the effects of fluctuating growth rates and division volumes (Fig. S3). Note that in the case of a fluctuating growth rate one also has to account for the correlation between the protein concentration and growth rate, as discussed in SI B.

In our framework the extrinsic noise is extracted from the time trajectory of the protein concentration of a sin-



	$f_{\rm in}$	$f_{\rm ex}$	$f_{\rm me}$
Data **a**	**0.41**	**0.56**	**0.03**
Data **b**	**0.26**	**0.73**	**0.01**

FIG. 4. (a) We compute the time derivative of protein concentration as a function of the current protein concentration using data from Ref. [33] and the measured slope normalized by the growth rate is 1.90. The time interval used is $\Delta t = 1$ min and the growth rate is $\mu = 0.0213$ min^{-1}. (b) We repeat the analysis using another data from Ref. [8] where the measured slope normalized by the growth rate is 0.45. Here $\Delta t = 1$ min and $\mu = 0.0327$ min^{-1}. (c) We adjust the time interval to compute the time derivative of protein concentration and compute the slope in the linear fit of $\Delta p/\Delta t$ *vs.* p. The normalized slope is linearly fitted as a function of the inverse of the time interval. The fraction of intrinsic noise in the total noise can be calculated from the intercept of the linear fit. We also infer the fraction of measurement noise in the total noise from the slope of the linear fit. (d) We summarize the calculated fractions of different noise for the two data sets. $f_{\rm in}$: the fraction of intrinsic noise. $f_{\rm ex}$: the fraction of extrinsic noise. $f_{\rm me}$: the fraction of measurement noise.

gle gene, which is distinct from that of the dual-reporter method. If the two genes in the dual-reporter setup share the same fluctuating translation rate $k_2(t)$, the two definitions of extrinsic noise will coincide [SI G, Fig. S6(a)]. However, if the correlated noise between the two genes is at the transcriptional level, the extrinsic noise inferred from the dual-reporter will be smaller than the one extracted from our protocol, which we confirm numerically [Fig. S6(b)].

Analysis of experimental data.—Experimentally, the measured protein concentration is always augmented by measurement noise. To model the effects of measurement noise, we assume the measured protein concentration at time t to equal

$$p(t) = p_0(t) + \eta(t) \tag{5}$$

where $p_0(t)$ is the actual protein concentration and $\eta(t)$ is the measurement noise term assumed uncorrelated between different measurements. We will revisit this assumption later on and show that the datasets we analyzed are consistent with it. The covariance between $\Delta p/\Delta t$ and p becomes $\mathrm{cov}(\frac{\Delta p}{\Delta t}, p) = \mathrm{cov}(\frac{\Delta p_0}{\Delta t}, p_0) - \sigma_\eta^2/\Delta t$. Compared with Eq. (4), the slope in the linear fitting of $\Delta p/\Delta t$ *vs.* p is modified to

$$S \equiv -\frac{\mathrm{cov}(\frac{\Delta p}{\Delta t}, p)}{\sigma_p^2} = \mu\big(f_{\rm in} + \frac{\sigma_\eta^2}{\mu\sigma_p^2\Delta t}\big). \tag{6}$$

We confirm Eq. (6) using numerical simulations with artificial measurement noise. In this case since σ_η^2 is assigned and $f_{\rm in}$ is known, we can directly compare the left and right sides of Eq. (6), obtaining good agreement (SI H, Fig. S7). Experimentally, the fluorescence level may not accurately reflect the instantaneous protein number due to a finite maturation time of the fluorescent protein. We have confirmed that the effects of a finite maturation time does not affect our results for experimentally relevant values of the maturation times [34] (SI I, Fig. S8).

We analyze two datasets of *E. coli* growth. In both, cells are exponentially growing and a fluorescent protein is constitutively expressed [8, 33]. A single lineage of cells is tracked for about 100 generations with cell volume and fluorescence level measured simultaneously. In both cases, the time interval between two consecutive data points is 1 min. To compute $f_{\rm in}$, we increase the time interval to compute $\Delta p/\Delta t$ and find the slopes in the linear fitting of $\Delta p/\Delta t$ *vs.* p for each time interval [see examples for $\Delta t = 1$ min in Fig. 4(a, b)]. We then linearly fit the resulting slopes as a function of $1/\Delta t$ [Fig. 2(c)] and the results agree well with the prediction of Eq. (6) [Fig. 4(c)]. Notably, this allows us to infer both $f_{\rm in}$ as the intercept of the linear fit, and the fraction of measurement noise from the slope. The results are summarized in Fig. 4(d). To justify the assumption of uncorrelated measurement noise, we show that the scaling with Δt in Eq. (6) is violated for correlated measurement noise (SI H, Fig. S7).

In this way we find that the ratio between the measurement noise and the total noise in the two data sets are respectively 17% and 10% in terms of their standard deviations, which are the square roots of the numbers in Figure 4(d). We can further use our analytic results for constitutively expressed genes as used in these experiments to estimate the average copy numbers of proteins at cell birth and the translation burst parameter $\beta\tau_m$ (see Eqs. S28, S29 in SI C) [31]. We find that $N_p \approx 230$ at cell birth, $\beta\tau_m \approx 1.37$ for Data in Fig. 4(a), and $N_p \approx 210$ at cell birth, $\beta\tau_m \approx 2.81$ for Data in Fig. 4(b). The differences between the two data sets are presumably due to the different strains and promoters. We note that if the normalization constant to convert the fluorescence level to protein number is known, one can also compute the partitioning noise based on Eq. (1) and the Poisson noise

as the remaining component of the intrinsic noise, which is confirmed using the synthetic data [Fig. S1(g)].

Summary and outlook.—In this work, we start from a general framework of stochastic gene expression in exponentially growing cells. Our approach allows us to take into account the cell growth and division explicitly and study the variability in protein *concentrations*, directly relevant to experiments on proliferating cells such as bacteria, yeast or cancer cells. We derive a broadly applicable decomposition of the protein concentration noise, finding that the total noise can be expressed as the sum of the noise due to upstream factors, the Poisson noise due to the random process of production and degradation, and the noise due to random partitioning during cell division. These results are independent of the underlying details of the particular dynamics of mRNA and protein synthesis. Given a time trajectory of protein concentration, one may linearly fit the discrete time derivative of protein concentration as a function of the protein concentration. We find that the slope of the fit, normalized by the growth rate, equals the fraction of intrinsic noise in the total protein concentration noise in the absence of measurement noise. We verify our theoretical framework on synthetic data of protein concentrations for genes with various underlying gene expression dynamics.

Importantly, we generalize our protocol to analyze experimental data of *E. coli* gene expression and show how a generalization of the method can simultaneously reveal the fraction of *measurement noise* in addition to that of intrinsic and extrinsic noise. Our framework predicts that the slope in the linear fitting of the time derivative of protein concentration *vs.* the current protein concentration has a linear dependence on the inverse of the time interval used to compute the time derivative, which agrees well with the experimental results. Assuming a model of a constitutively expressed protein as used in these experiments, our approach also allows us to infer the average copy numbers of proteins at cell birth as well as the translation burst parameter.

The generality of our approach and the agreement between experiments and theoretical predictions suggests that the method should be broadly applicable and will serve as a useful tool for gene expression analysis including mammalian cells and other non-microbial eukaryotes as long as a sufficient number of cell cycles are sampled. Our protocol to extract the intrinsic and extrinsic noise relies only on the time trajectory of protein concentration of a single gene, in contrast to the dual-reporter protocol which relies on measuring protein concentrations of two identical genes. Combing our method with the dual-reporter method, one can further decompose the extrinsic noise into correlated and uncorrelated components. Theoretically, our work paves the way to further studies on the nature of the widely-observed yet poorly understood extrinsic noise in gene expression.

We thank Ido Golding and Lydia Robert for useful discussions and feedback. We also thank the anonymous reviewers for their comments. A.A. was supported by NSF CAREER grant 1752024 and the Harvard Dean's Competitive Fund. A.A. and J.L. thank support from Harvard's MRSEC (DMR-1420570).

[1] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, Science **297**, 1183 (2002).
[2] J. M. Raser and E. K. O'shea, Science **304**, 1811 (2004).
[3] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, Cell **123**, 1025 (2005).
[4] J. R. Newman, S. Ghaemmaghami, J. Ihmels, D. K. Breslow, M. Noble, J. L. DeRisi, and J. S. Weissman, Nature **441**, 840 (2006).
[5] A. Raj and A. van Oudenaarden, Cell **135**, 216 (2008).
[6] Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie, Science **329**, 533 (2010).
[7] N. Eling, M. D. Morgan, and J. C. Marioni, Nature Reviews Genetics **20**, 536 (2019).
[8] P. Wang, L. Robert, J. Pelletier, W. L. Dang, F. Taddei, A. Wright, and S. Jun, Current Biology **20**, 1099 (2010).
[9] M. Godin, F. F. Delgado, S. Son, W. H. Grover, A. K. Bryan, A. Tzur, P. Jorgensen, K. Payer, A. D. Grossman, M. W. Kirschner, *et al.*, Nature Methods **7**, 387 (2010).
[10] M. Campos, I. V. Surovtsev, S. Kato, A. Paintdakhi, B. Beltran, S. E. Ebmeier, and C. Jacobs-Wagner, Cell **159**, 1433 (2014).
[11] S. Taheri-Araghi, S. Bradde, J. T. Sauls, N. S. Hill, P. A. Levin, J. Paulsson, M. Vergassola, and S. Jun, Current Biology **25**, 385 (2015).
[12] N. Cermak, S. Olcum, F. F. Delgado, S. C. Wasserman, K. R. Payer, M. A Murakami, S. M. Knudsen, R. J. Kimmerling, M. M. Stevens, Y. Kikuchi, A. Sandikci, M. Ogawa, V. Agache, F. Baleras, D. M. Weinstock, and S. R. Manalis, Nature Biotechnology **34**, 1052 (2016).
[13] H. A. Crissman and J. A. Steinkamp, The Journal of Cell Biology **59**, 766 (1973).
[14] O. Padovan-Merhar, G. P. Nair, A. G. Biaesch, A. Mayer, S. Scarfone, S. W. Foley, A. R. Wu, L. S. Churchman, A. Singh, and A. Raj, Molecular Cell **58**, 339 (2015).
[15] M. Wang, J. Zhang, H. Xu, and I. Golding, Nature microbiology **4**, 2118 (2019).
[16] S. Elliott and C. McLaughlin, Proceedings of the National Academy of Sciences **75**, 4384 (1978).
[17] N. Brenner, E. Braun, A. Yoney, L. Susman, J. Rotella, and H. Salman, The European Physical Journal E **38**, 1 (2015).
[18] N. Walker, P. Nghe, and S. J. Tans, BMC Biology **14**, 11 (2016).
[19] J. Lin and A. Amir, Nature Communications **9**, 4496 (2018).
[20] G. E. Neurohr, R. L. Terry, J. Lengefeld, M. Bonney, G. P. Brittingham, F. Moretto, T. P. Miettinen, L. P. Vaites, L. M. Soares, J. A. Paulo, *et al.*, Cell **176**, 1083 (2019).
[21] N. Nordholt, J. H. van Heerden, and F. J. Bruggeman, Current Biology **30**, 2238 (2020).
[22] H. Kempe, A. Schwabe, F. Crémazy, P. J. Verschure, and F. J. Bruggeman, Molecular Biology of the Cell **26**, 797

(2015).
[23] X.-M. Sun, A. Bowman, M. Priestman, F. Bertaux, A. Martinez-Segura, W. Tang, C. Whilding, D. Dormann, V. Shahrezaei, and S. Marguerat, Current Biology **30**, 1217 (2020).
[24] D. L. Jones, R. C. Brewster, and R. Phillips, Science **346**, 1533 (2014).
[25] R. D. Dar, B. S. Razooky, L. S. Weinberger, C. D. Cox, and M. L. Simpson, PLoS One **10**, 1 (2015).
[26] O. Lenive, P. D. Kirk, and M. P. Stumpf, BMC Systems Biology **10**, 81 (2016).
[27] M. Soltani, C. A. Vargas-Garcia, D. Antunes, and A. Singh, PLoS Comput Biol **12**, e1004972 (2016).
[28] J. Paulsson, Nature **427**, 415 (2004).
[29] R. Milo and R. Phillips, *Cell biology by the numbers* (Garland Science, 2015).
[30] See Supplemental Material [url] for detailed discussions, which includes Refs. [35–40].
[31] J. Paulsson, Physics of Life Reviews **2**, 157 (2005).
[32] P. S. Swain, M. B. Elowitz, and E. D. Siggia, Proceedings of the National Academy of Sciences **99**, 12795 (2002).
[33] Y. Tanouchi, A. Pai, H. Park, S. Huang, N. E. Buchler, and L. You, Scientific Data **4**, 170036 (2017).
[34] E. Balleza, J. M. Kim, and P. Cluzel, Nature Methods **15**, 47 (2018).
[35] D. Huh and J. Paulsson, Nature Genetics **43**, 95 (2011).
[36] E. Powell, Microbiology **15**, 492 (1956).
[37] J. Lin and A. Amir, Cell Systems **5**, 358 (2017).
[38] E. H. Simpson, Journal of the Royal Statistical Society: Series B (Methodological) **13**, 238 (1951).
[39] D. T. Gillespie, The Journal of Chemical Physics **113**, 297 (2000).
[40] J. Lin, M. Manhart, and A. Amir, Genetics **215** (2020).