Universal protein distributions in a model of cell growth and division

Naama Brenner
Department of Chemical Engineering and Laboratory of Network Biology, Technion, Haifa 32000, Israel

C. M. Newman
Courant Institute of Mathematical Sciences, New York, New York 10012 USA
and NYU-ECNU Institute of Mathematical Sciences at NYU Shanghai, 3663 Zhongshan Road North, Shanghai 200062, China

Dino Osmanović and Yitzhak Rabin
Department of Physics and Institute of Nanotechnology and Advanced Materials, Bar-Ilan University, Ramat Gan 52900, Israel

Hanna Salman
Department of Physics and Astronomy, Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

D. L. Stein
Department of Physics and Courant Institute of Mathematical Sciences, New York University, New York, New York 10012 USA
and NYU-ECNU Institutes of Physics and Mathematical Sciences at NYU Shanghai, 3663 Zhongshan Road North, Shanghai, 200062, China
(Received 2 March 2015; published 23 October 2015)

Protein distributions measured under a broad set of conditions in bacteria and yeast were shown to exhibit a common skewed shape, with variances depending quadratically on means. For bacteria these properties were reproduced by temporal measurements of protein content, showing accumulation and division across generations. Here we present a stochastic growth-and-division model with feedback which captures these observed properties. The limiting copy number distribution is calculated exactly, and a single parameter is found to determine the distribution shape and the variance-to-mean relation. Estimating this parameter from bacterial temporal data reproduces the measured distribution shape with high accuracy and leads to predictions for future experiments.

DOI: 10.1103/PhysRevE.92.042713 PACS number(s): 87.10.−e, 87.15.A−, 87.15.Ya, 87.17.Ee

I. INTRODUCTION

The phenotype of a biological cell—in particular, the types and copy numbers of its expressed proteins—fluctuates from cell to cell, even among those whose genotypes and growth environments are identical (reviewed in Refs. [1–3]). Protein content depends on a complex interplay of genetic, epigenetic, and metabolic processes, with numerous cell-specific regulatory mechanisms and feedback loops. However, recent experiments [4] have demonstrated that for two different types of microorganism (yeast and bacteria), each under a broad range of conditions, the distribution of highly expressed protein copy number appears *universal*: Under rescaling by mean and standard deviation, all such distributions collapse onto a single skewed curve [5]. In the same experiments variances were found to depend quadratically on their means, a trend displayed also by all highly expressed proteins in *Eschericia coli* in a genome-wide study [6] (see also Ref. [7]).

A recent study following the protein content in individual *E. coli* bacteria over roughly 70 generations has revealed that, under the same scaling criteria, the shape of the distribution of protein copy number sampled over time in an individual converges to the one observed in large populations [8]. While analogous temporal data are currently unavailable for yeast, this is an important property that reflects the ergodicity of the relative fluctuations in protein expression in bacteria. These results can serve as a basis for constructing a model relating bacterial temporal protein dynamics to their distributions.

II. MODEL WITHOUT FEEDBACK

Given that the universal statistical properties described above were found for a range of experimental conditions for various proteins in bacteria [4], such a model should rely only on general coarse-grained processes. We therefore start by assuming as little as possible given the experimental data:

(i) Protein number increases as $e^{k_i t}$ during the i^{th} generation, where the exponential growth rate k_i fluctuates with i [8].

(ii) The time T_i of the i^{th} generation [i.e., the time between cell division at the $(i-1)^{\text{st}}$ generation and that at the i^{th}] is also random [9–13].

(iii) The product $X_i = k_i T_i$ is a random variable, with (positive) mean μ and variance σ^2. We will refer to X_i as the *accumulation exponent*.

(iv) Protein number is conserved at cell division, and protein degradation is much slower than a typical interdivision time [14].

Let N_i denote the copy number of a given type of protein in the cell and f_i the copy number ratio between the daughter and parent cells, both at the end of the i^{th} generation. Incorporating the features listed above gives rise to the recursion relation

$$N_{i+1} = f_i N_i \exp(X_{i+1}). \tag{1}$$

In bacteria f_i is narrowly distributed about $1/2$ [15]. We take $f_i = 1/2$ for now and discuss deviations from this assumption

later. The solution of (1) for arbitrary generation number n is

$$N_n = 2^{-n} N_0 \exp\left(\sum_{j=1}^{n} X_j\right), \tag{2}$$

where N_0 is the initial copy number. Then

$$\ln(N_n/N_0) = -n \ln 2 + \sum_{j=1}^{n} X_j. \tag{3}$$

The mean of X_j should compensate for the decrease in protein number in the daughter cells caused by division; otherwise copy numbers would be unstable, running off to unsustainably large numbers or falling to zero within a few generations. Equation (3) can be rewritten as

$$[\ln(N_n/N_0) - n(\mu - \ln 2)]/\sqrt{n}\sigma = \left[\sum_{j=1}^{n} X_j - n\mu\right] \Big/ \sqrt{n}\sigma. \tag{4}$$

If the accumulation exponents X_j are independent, then the central limit theorem gives

$$[\ln(N_n/N_0) - n(\mu - \ln 2)]/\sqrt{n}\sigma \to \mathcal{N}(0,1), \tag{5}$$

that is, the left-hand side converges in distribution to the normal distribution with mean zero and variance 1. There is no stationary distribution for this process; the mean and variance of N_n vary with time (or n), even when $\mu = \ln 2$ exactly. This conclusion holds independently of the various distributions used; all that matters is that fluctuations are independent between generations. This analysis demonstrates that a stationary distribution, as experimentally observed, can result only if some negative feedback is present. Given this, we next introduce and analyze a modified model with effective feedback regulating protein accumulation, and, following that, we discuss its experimental justification and consequences.

III. MODEL WITH FEEDBACK

A given protein type in an individual cell has a well-defined typical copy number $\overline{N}$. Its value is nonuniversal, depending on protein type, growth conditions, and possibly other biological factors [8,16]. The stationary distribution shape must therefore be independent of $\overline{N}$.

A natural extension of the growth-and-division model consistent with observations is the introduction of an accumulation exponent that is negatively correlated with protein number at the start of the cycle. The experimental requirement of universality constrains the form of the feedback term: A change in scale of $\overline{N}$ cannot alter the functional form of the recursion relation. The only function with this property of scale invariance is the power law; the modified recursion relation is therefore

$$N_{i+1} = f_i N_i [\exp(\xi_{i+1})](N_i/\overline{N})^{-\alpha}, \tag{6}$$

with f_i defined as before, ξ_{i+1} (which we will call the residual accumulation exponent) the component of the accumulation that fluctuates independently from generation to generation, and a new phenomenological parameter $0 < \alpha < 1$; $\alpha = 0$ is the case without feedback [17].

The recursion relation of the modified model is

$$\ln N_{i+1} = \ln f_i + \xi_{i+1} + (1-\alpha) \ln N_i + \alpha \ln \overline{N}. \tag{7}$$

It is not hard to check that, first, there is now a limiting stationary distribution, with $\langle N \rangle \approx \overline{N}$, and, second, that $\overline{N}$ can be scaled out of the growth equations.

The introduction of a nonzero α makes the specific form of the limiting distribution dependent (though not too sensitively; see below) on the distributions of f_i and ξ_i. Experiments on bacteria indicate that ξ_i is approximately normally distributed [see Fig. 2(b)]. Using this, we can solve for the limiting distribution exactly when the division ratio is fixed. The limiting distribution is again lognormal:

$$P(N) = \frac{1}{N\Sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln N - \mathcal{M})^2}{2\Sigma^2}\right], \tag{8}$$

with $\mathcal{M} = \ln \overline{N} + (\mu - \ln 2)/\alpha$ and $\Sigma = \sigma/\sqrt{2\alpha - \alpha^2}$. These two parameters together determine the mean and variance of the distribution: specifically, $\langle N \rangle = \exp\{\mathcal{M} + \Sigma^2/2\}$ and $\langle N^2 \rangle - \langle N \rangle^2 = (e^{\Sigma^2} - 1)\exp\{2\mathcal{M} + \Sigma^2\}$. However, only Σ determines the shape of the distribution; different values of $\mathcal{M}$ collapse on one another following scaling by a linear transformation. Moreover, for fixed Σ the variance scales quadratically with the mean.

In terms of the model, Eq. (8) has several important features. First, it preserves universality under scaling with respect to all variables that appear only in the parameter $\mathcal{M}$, because the distribution shape is independent of this additive term. Consequently, all values of $\overline{N}$, μ, and division ratio yield the same distribution shape.

Second, as noted above, a single composite parameter Σ determines the shape of the distribution. Σ characterizes the balance between the *variance* of accumulation exponents, which tends to drive the process to diverge, and the effective feedback parameter α, which provides a "restoring force." Once Σ is determined, both properties—collapse of scaled distributions and quadratic dependence of variance on mean—are preserved.

Third, it should be noted that in the setting of our model, the limiting steady-state distribution equally well represents the time average over many generations of a single individual or the average at a single large time over a large population in which the individuals are evolving independently.

The analysis above assumed a Gaussian distribution of the accumulation exponents and a fixed value of the division ratio. We now explore the robustness of our conclusions. Without these assumptions, the lognormal solution will no longer be exact but will not be significantly altered for a variety of unimodal distributions for both variables. Moreover, the scaling properties within classes characterized by Σ (defined as before and which is now close to but not exactly equal to the standard deviation of $\ln N$) still hold. Figure 1(a) shows examples of means and variances computed from many simulations, in which the f_i's were drawn from a Gaussian distribution and the ξ_i's from a gamma distribution. For each simulation, α and $\overline{N}$ were chosen randomly, and the variance σ^2 of the gamma distribution was adjusted to give a shape

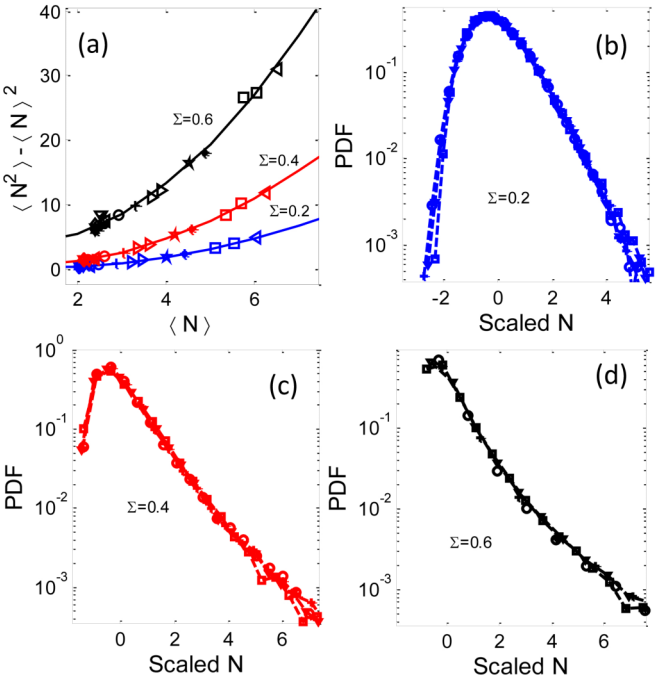

FIG. 1. (Color online) Model simulations. The stochastic process described by Eq. (6) was simulated over 20 000 generations for each run. Values for α and $\overline{N}$ were chosen uniformly at random in the range [0,1] and [0.25,0.75], respectively. The division ratio f_i was drawn from a Gaussian distribution with mean 1/2 and standard deviation 0.1 and ξ_i from a gamma distribution whose variance was adjusted to α to obtain one of three values of the shape parameter Σ. (a) Variance vs mean of 18 simulations for each shape parameter show a collapse on three corresponding parabolas. [(b)–(d)] Limiting distributions of three simulations for each class are all well fit by lognormal but have different shape parameters and so span a range of shapes, from approximately Gaussian (b) to an exponential-like tail (c) to a highly skewed distribution (d). In each class the distributions collapse onto one another to high accuracy.

parameter Σ equal to one of three values: 0.2, 0.4, or 0.6. The resulting means and variances are seen in Fig. 1(a) to collapse onto three parabolas corresponding to the three classes defined by the value of Σ. Limiting distributions are shown in Figs. 1(b)–1(d); while they are all very close to lognormal, the different Σ's lead to a range of shapes, from nearly Gaussian [Fig. 1(b) $\Sigma = 0.2$] to skewed with exponential-like tail [Fig. 1(c) $\Sigma = 0.4$] and, finally, to highly skewed [Fig. 1(d) $\Sigma = 0.6$]. Within each class the distributions from all simulations collapse after rescaling onto a single curve. Thus, both properties of distribution collapse and the quadratic dependence of variance on mean hold once Σ is fixed, even though the conditions of the exactly solvable model are relaxed.

In addition, the assumption of symmetric division can be relaxed as long as the average accumulation compensates for the loss at division. The assumption of stable proteins can also be relaxed to include first-order protein degradation; this would require an additional parameter and would modify only the accumulation exponent.

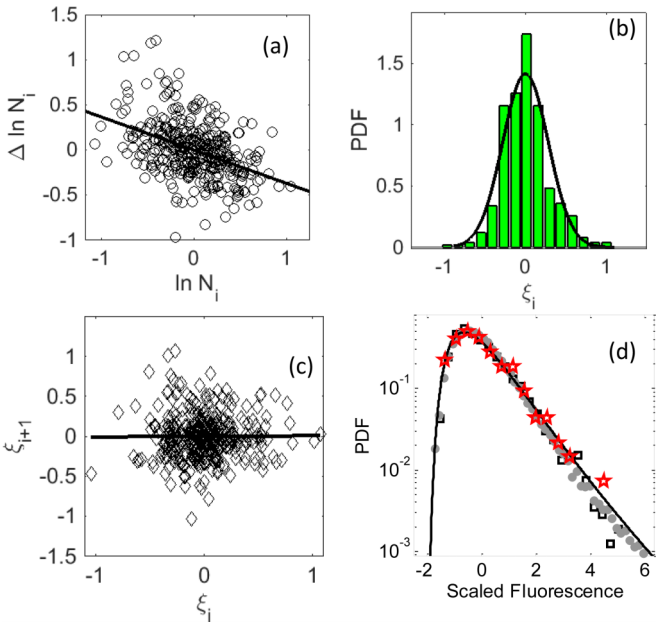

FIG. 2. (Color online) Comparison with data. (a) $\Delta \ln N_i \equiv \ln N_{i+1} - \ln N_i$ plotted vs $\ln N_i$ in units of $\overline{N}$. Solid line is $y = -0.37x$. Accumulation exponents in consecutive generations are (b) approximately normally distributed with average ln2 (subtracted out in the figure) and (c) are independent between generations. Solid line is $y = 0.0075x - 0.00036$. (d) Estimating the universality class parameter Σ for these data from (a) and (b); the distribution shape is predicted by Eq. (8) (black line) and compared to data. Gray circles: large population snapshot; black squares: protein trajectories of individual trapped bacteria; red stars: sampled points at the end of each cell cycle. Data from Refs. [4,8].

IV. COMPARISON WITH DATA

To test the assumption of negative correlation we plotted experimental values of $\Delta \ln N_i = \ln N_{i+1} - \ln N_i$ vs $\ln N_i$, as measured for bacteria, in Fig. 2(a). The data points were collected from six individual trajectories normalized to unit average (data from Ref. [8]).

In agreement with Eq. (7), the data are consistent with random scatter about an overall linear dependence, with negative slope determining α to be approximately 0.37 ± 0.04. Using this value the residual accumulation exponents ξ_i can be extracted from the data using Eq. (7) and measured values of N_i. The approximately Gaussian distribution of these exponents is shown in Fig. 2(b), and their independence between consecutive generations is evident from Fig. 2(c).

The parameter determining the distribution shape in our model is $\Sigma = \sigma/\sqrt{2\alpha - \alpha^2}$. Estimating α and σ from Figs. 2(a) and 2(b), respectively, we find $\Sigma \approx 0.4 \pm 0.02$. Figure 2(d) shows the lognormal distribution of Eq. (8) corresponding to this parameter in rescaled units (black line), together with data from a large bacterial population (gray circles), and single cell trajectories (black squares and red stars) that exhibits the measured universal distribution shape over several decades of probability. We note that this is not a fit but a model prediction with no adjustable parameters: The single parameter determining the distribution

shape is computed separately using the single-cell dynamic measurements.

V. DISCUSSION

Motivated by experiments that found universal protein distributions under various conditions in yeast and bacteria, and by single-cell measurements of protein accumulation and division in bacteria across multiple generations, we have presented a model based on the premise that the combined processes of growth, division, and feedback set the distribution shape. With fixed division ratio and Gaussian randomness the model is exactly solvable. The solution identifies a single parameter Σ [Eq. (8)] defining the distribution shape: It quantifies the balance between growth of variance and feedback that stabilizes protein numbers. With Σ fixed, a rescaling by mean and standard deviation collapses these distributions onto a single curve and displays a quadratic relation between variance and mean.

Thus, our model predicts that populations in the same class—i.e., which share the same shape parameter—exhibit similar set-point balance between the opposing forces in the dynamics of their protein content across time, i.e., between the *variance* of accumulation exponents (σ) that drive the process to diverge and the feedback parameter (α) that prevents divergence. Therefore, if the variance of the exponents ξ_i changes, then the feedback parameter α should change in a correlated manner. To test this prediction, single-cell dynamical trajectories need to be measured over a variety of conditions that span these parameters. Another possibility consistent with our model is that both σ and α are fixed. At the moment, experimental perturbations—for example, changing medium or temperature—can change the mean, for example, by modifying the mean cell cycle time, but their effect on the variance of exponents is unknown [10].

Our approach shares some features with previous theoretical work but differs in other respects. Earlier work focused on protein accumulation and division [18–22] or protein accumulation and continuous dissipation [23,24]. The recent data on protein content over multiple generations [8] shows that, due to the exponential nature of protein accumulation, division or dissipation alone cannot stabilize copy numbers and reveals a correlation between variables across generations.

The classes of proteins of interest are those consisting of high-copy-number molecules, characterized by exponential accumulation between successive cell divisions. The exponential accumulation of protein during a cell cycle suggests that protein production reflects a coherent integration of many correlated processes in the cell. Exponential growth of the cell size between divisions, as well as negative correlation analogous to the one reported here, were measured in several recent experiments [10,11,13,25]. Moreover, results on trapped bacteria show explicitly that the exponents of cell-size growth and protein accumulation are strongly correlated on a cycle-by-cycle basis [8]. This suggests a picture where highly expressed proteins that are strongly coupled to cell metabolism are components of multidimensional phenotypes that covary between individual cells. This view is supported by a model recently proposed to explain exponential biomass growth as resulting from an interacting network of reactions [26,27]. Furthermore, our model is mathematically related to a recently proposed model of cell-size regulation [12,28], which finds under similar assumptions a lognormal distribution with the same compound parameter governing its shape. For highly expressed proteins, this may be expected since protein production and cell growth are tightly coupled [8]. However, there are also important differences between the two models, which we address in detail in the Appendix.

Our model addresses directly the universal behavior of bacterial protein distribution among different biological realizations, including expression regulation mechanisms, growth conditions, and types of microorganism. Its ingredients are independent of specific biological mechanisms and rely on those general aspects of cellular events—exponential protein accumulation, division, and feedback—that are likely to be common to all dividing cell populations. This marks a significant departure from the main current line of research on protein number variation, which investigates synthetically produced proteins while experimentally isolating the contribution of specific microscopic mechanisms [29–34].

In particular, we have observed that feedback must be present, because without it the mean and variance necessarily drift to larger values as time increases. Moreover, regardless of the specific processes leading to feedback (which may differ for different protein types and organisms), the mathematical form of the feedback in a growth-and-division model must be power law to be consistent with universality.

ACKNOWLEDGMENTS

We thank Erez Braun for many useful discussions. This work was supported by the US-Israel Binational Science Foundation (N.B. and H.S.). D.L.S. thanks the John Simon Guggenheim Foundation for a fellowship that partially supported this research and Bar-Ilan University for its hospitality while this work was initiated. Y.R. acknowledges support by the I-CORE Program of the Planning and Budgeting committee and the Israel Science Foundation and by the US-Israel Binational Science Foundation. N.B acknowledges support of the Israel Science Foundation (Grant No. 1566/11). H.S. acknowledges the support of NSF Grant No. PHY-1401576. C.M.N. acknowledges the support of NSF Grant No. DMS-1007524 and D.L.S. acknowledges the support of NSF Grant No. DMS-1207678.

APPENDIX: UNIVERSAL PROTEIN DISTRIBUTIONS AND CELL GROWTH

In this Appendix, we detail the differences between our model and a mathematically similar model of cell-size regulation recently proposed by Amir [12,28]. However, before describing those differences, we first discuss some nontrivial consequences following from the *similarity* between the models. The fact that similar mathematical descriptions, albeit with different biological interpretations, may capture the dynamics of two distinct phenotypes is a potentially significant mathematical unification resulting from a strong coupling among disparate biological processes. As noted above, cell size and protein copy number are two separate, but strongly correlated, phenotypes of the cell. Which one controls the

other, or whether both are regulated together, is not known at this time. One interesting implication of the work presented here is that not only cell size but also the total content of highly expressed proteins is under the control of what appears to be a global cellular feedback, supporting the viewpoint that protein copy number variation is a *global* variable. This marks a significant departure from the main current line of research on protein number variation.

We now turn to a discussion of some important differences in the interpretation and consequences of the two models. First, the requirement that the average must scale out of the distribution shape in any model of universal distributions necessitates mathematically the power-law form of the feedback. Second, the parameter Σ that governs the "universality class" of the family of distributions has been identified, and its constancy leads both to the collapse of all curves under linear scaling and to the observed quadratic relation between variance and mean. The latter is an especially important consequence and has no analog in cell-size distributions. We have also seen numerically that the division into such classes extends beyond the conditions of the analytically solvable case, rendering this result robust with respect to a wide class of distributions of the underlying random variables.

Perhaps most importantly, our model makes specific predictions on the constrained changes allowed in protein trace parameters under varying conditions. Similar predictions cannot currently be made for cell-size distributions.

Both our model and that of Refs. [12,28] are also easily modified to handle asymmetric division, as in yeast. However, until data are available that relate temporal to population statistics, it remains to be seen to what extent the dynamics of proteins across generations in yeast can be described by the approach outlined in the main text.

A final important difference between the approach described in this paper and that in [28] concerns the nature of the feedback itself. Analysis of *E. coli* data led to the conclusion that cell-size feedback is characterized by $\alpha = 1/2$ [28], corresponding (in leading order) to the proposal that the feedback arises from constant addition of volume over the cell cycle. In contrast, a different mechanism(s) may apply for copy numbers, and α can in principle vary among protein types: Our current results are consistent with various values of α with an average of $\alpha \approx 0.37$. However, at this stage there is no direct evidence that can determine which phenotype (cell size, protein copy number, etc.) controls the division point of the cell and thus the feedback mechanism that controls it. In principle, it could also be a cellular state that is defined by several phenotypes simultaneously.

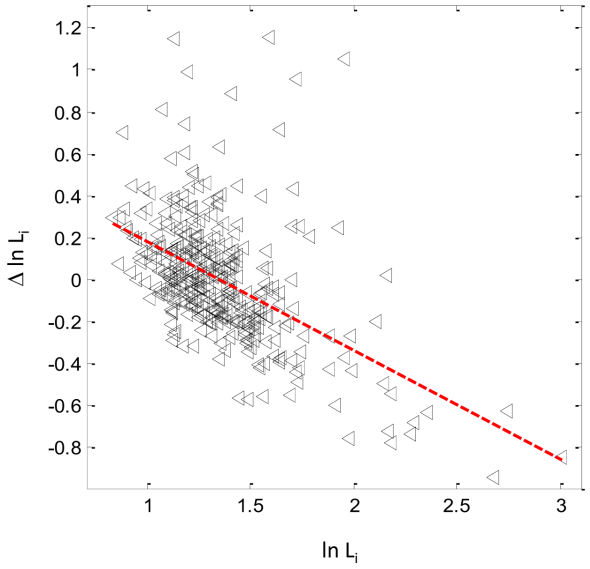

FIG. 3. (Color online) Cell-size analysis. Similar analysis of that carried out for protein copy number [Fig. 2(a)] was carried out for cell-size data, quantified here by the length L of the bacterial cell in the trapping channel. $\Delta \ln L_i = \ln L_{i+1} - \ln L_i$, computed from the data in Ref. [8], is plotted vs $\ln L_i$. The red dashed line is the linear best fit given by $\Delta \ln L_i = -0.5185 \ln L_i + 0.702$. The slope, which represents the feedback parameter α, is approximately $1/2$ as predicted by the constant volume addition model described in Ref. [28].

We show this difference explicitly in the figure above. Figure 3 uses cell-size data from Ref. [8] to compute the feedback parameter α, in a manner similar to that used in Fig. 2(a), for the cell-size phenotype. This results in $\alpha \approx 0.5$, in agreement with Ref. [28], but differing from the value $\alpha \approx 0.37$ shown in Fig. 2(a).

[1] N. Maheshri and E. K. O'Shea, Ann. Rev. Biophys. Biomol. Struct. **36**, 413 (2007).
[2] A. Raj and A. van Oudenaarden, Cell **135**, 216 (2008).
[3] A. Sanchez and I. Golding, Science **342**, 1188 (2013).
[4] H. Salman, N. Brenner, C.-K. Tung, N. Elyahu, E. Stolovicki, L. Moore, A. Libchaber, and E. Braun, Phys. Rev. Lett. **108**, 238105 (2012).
[5] Note that we use the term *universal* to describe these statistical properties of cellular protein content. This should not be confused with the controversy over the dependence of mRNA production on promoter architecture (see, e.g., L.-H. So, A. Ghosh, C. Zong, L. A. Sepulveda, R. Segev, and I. Golding, Nat. Genet. **43**, 554 (2011); and D. L. Jones, R. C. Brewster, and R. Phillips, Science **346**, 1533 (2014); indeed, quantitative measurements have shown that there is almost no correlation between mRNA and protein copy numbers in the cell [6].
[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] C. Furasawa, T. Suzuki, A. Kashiwagi, T. Yomo, and K. Kaneko, Biophysics **1**, 25 (2005).
[8] N. Brenner, E. Braun, A. Yoney, L. Susman, J. Rotella, and H. Salman, Eur. Phys. J. E **38**, 102 (2015).
[9] E. O. Powell and F. P. Errington, J. Gen. Microbiol. **31**, 315 (1963).
[10] S. Iyer-Biswas, C. S. Wrighta, J. T. Henryb, K. Loa, S. Burov, Y. Linc, G. E. Crooksd, S. Crossonb, A. R. Dinner, and N. F. Scherera, Proc. Natl. Acad. USA **111**, 15912 (2014).

[11] M. Osella, E. Nugent, and M. C. Lagomarsino, Proc. Natl. Acad. Sci. USA **111**, 3431 (2014).
[12] I. Soifer, L. Robert, N. Barkai, and A. Amir, arXiv:1410.4771v2 (2014).
[13] S. Taheri-Araghi, S. Bradde, J. T. Sauls, N. S. Hill, P. A. Levin, J. Paulsson, M. Vergassola, and S. Jun, Curr. Biol. **25**, 385 (2015).
[14] A degradation rate that is faster than a typical cell interdivisional time would likely simply modify the accumulation exponent, so such proteins would also display exponential increase of copy number during a cell cycle and be covered equally well by our model. However, we exclude such proteins for now to avoid complicating factors.
[15] J. Männik, Jaan, F. Wu, F. J. Hol, P. Bisicchia, D. J. Sherratt, J. E. Keymer, and C. Dekker, Proc. Acad. Natl. Sci. USA **109**, 6957 (2012).
[16] $\overline{N}$ could in principle be modulated on long time scales by processes not included in our model. If they occur, such modulations would be sufficiently slow to be decoupled from the fast growth-and-division processes described here. It is therefore reasonable, and consistent with observations, to take $\overline{N}$ to be constant over a large number of generations.
[17] *A priori*, it is possible that α is not a constant but varies between protein types or individuals, or it could even fluctuate from generation to generation for a given protein in a single individual. We defer consideration of these possibilities and analyze here the case of constant α specific to a given protein in a given cell.
[18] O. G. Berg, J. Theoret. Biol. **71**, 587 (1978).
[19] N. Brenner, K. Farkash, and E. Braun, Phys. Biol. **3**, 172 (2006).
[20] N. Brenner and Y. Shokef, Phys. Rev. Lett. **99**, 138102 (2007).
[21] T. Friedlander and N. Brenner, Phys. Rev. Lett. **101**, 018104 (2008).
[22] K. Hosoda, T. Matsuura, H. Suzuki, and T. Yomo, Phys. Rev. E **83**, 031118 (2011).
[23] J. Paulsson, Phys. Life Rev. **2**, 157 (2005).
[24] N. Friedman, L. Cai, and X. S. Xie, Phys. Rev. Lett. **97**, 168302 (2006).
[25] M. Campos, I. V. Surovtsev, S. Kato, A. Paintdakhi, B. Beltran, S. E. Ebmeier, and C. Jacobs-Wagner, Cell **159**, 1433 (2014).
[26] S. Iyer-Biswas, G. E. Crooks, N. F. Scherer, and A. R. Dinner, Phys. Rev. Lett. **113**, 028101 (2014).
[27] R. Pugatch, Proc. Natl. Acad. Sci. U.S.A. **112**, 2611 (2015).
[28] A. Amir, Phys. Rev. Lett. **112**, 208102 (2014).
[29] A. Bar-Even, J. Paulsson, N. Maheshri, M. Carmi, E. O'Shea, Y. Pilpel, and N. Barkai, Nat. Gen. **38**, 636 (2006).
[30] K. F. Murphy, G. Balazsi, and J. J. Collins, Proc. Natl. Acad. Sci. USA **104**, 12726 (2007).
[31] G. Hornung, R. Bar-Ziv, D. Rosin, N. Tokurkik, D. S. Tawfik, M. Oren, and N. Barkai, Genome Res. **22**, 2409 (2012).
[32] M. Dadiani, D. van Dijk, B. Segal, Y. Field, G. Ben-Artzi, T. Raveh-Sadka, M. Levo, I. Kaplow, A. Weinberger, and E. Segal, Genome Res. **23**, 966 (2013).
[33] L. B. Carey, D. van Dijk, P. M. A. Sloot, J. A. Kaandrop, and E. Segal, PLoS Biol. **11**, e1001528 (2013).
[34] E. Sharon, D. van Dijk, Y. Kalma, L. Keren, O. Manor, Z. Yakhini, and E. Segal, Genome Res. **24**, 1698 (2014).