

Construction of an Effective Landscape for Multistate Genetic Switches

Mingyang Lu,¹ José Onuchic,^{1,*} and Eshel Ben-Jacob^{1,2,†}

¹*Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005-1827, USA*

²*School of Physics and Astronomy and The Sagol School of Neuroscience, Tel-Aviv University, Tel-Aviv 69978, Israel*

(Received 13 February 2014; published 13 August 2014)

Multistate genetic switches play a crucial role during embryonic development and tumorigenesis. An archetypical example is the three-way switch regulating epithelial-hybrid-mesenchymal transitions. We devise a special WKB-based approach to investigate white Gaussian and shot noise effects on three-way switches, and construct an effective landscape in good quantitative agreement with stochastic simulations. This approach allows efficient analytical or numerical calculation of the landscape contours, the optimal path, and the state relative stability for general multicomponent multistate switches.

DOI: 10.1103/PhysRevLett.113.078102

PACS numbers: 87.10.Mn, 87.16.Yc, 87.18.Cf, 87.18.Tt

One of the most challenging tasks in biological physics is to understand the underlying principles governing cell fate decision during embryonic development and tumorigenesis [1]. The cellular decisions are typically regulated by coupled genetic circuits that operate as multistate switches, i.e., circuits whose dynamics permits the coexistence of multiple states [2,3], each of which corresponds to a distinct cell fate or phenotype.

A typical example of cell decisions is the bidirectional transitions between epithelial and mesenchymal phenotypes (EMT and MET), which are involved in embryonic development and tissue repair [4]. Aberrant regulation of these transitions is a hallmark of cancer metastasis [5]. It was proposed that the genetic network that regulates these transitions acts as a three-way switch [6,7]. In addition to the epithelial and mesenchymal phenotypes, this network also allows for transitions into a hybrid epithelial-mesenchymal phenotype [8].

Noise associated with gene expression often plays crucial roles in cellular functions and in particular in phenotypic stability and transitions [9,10]. Therefore, a sound understanding of the operational principles of multistate switches requires assessing the effect that different types of noise have on the transitions and on the relative stability of the switch states. Circuit stochasticity has gained considerable theoretical investigations in the context of the function of genetic networks [11,12]. Of interest here are previous attempts to apply transition rate theory, a method commonly used to investigate various kinds of problems in physics [13], such as protein folding [14].

Motivated by the Waddington epigenetic landscape, many rely on the notion of an effective potential to describe cell fate determination in the presence of noise [15]. In the noise-free limit, the circuit dynamics can be modeled deterministically by the chemical rate equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$. The challenge is that its deterministic dynamics is not derived from a potential function. The dynamics can be mapped to the overdamped dynamics of a particle driven by

a force \mathbf{f} . In the limit that one of the processes is considerably slower than the other degrees of freedom, the dynamics can be reduced to a one-dimensional equation, while the effective potential can be defined as $U(\mathbf{x}) = -\int_0^{\mathbf{x}} \mathbf{f}(\mathbf{x}') \cdot d\mathbf{x}'$ [16]. Clearly, this definition is not valid for general multidimensional systems (e.g., gene circuits with two or more components). The corresponding force field defined by multidimensional \mathbf{f} is not necessarily conservative; thus, the integral of \mathbf{f} is path dependent and a potential function cannot be defined.

A common approach is to define an effective potential $U(\mathbf{x})$ that captures the effect of noise to be

$$U(\mathbf{x}) = -\ln P(\mathbf{x}), \quad (1)$$

where $P(\mathbf{x})$ is the steady state probability for the state \mathbf{x} in the presence of noise. This approach captures the effects of the noise on gene circuit dynamics since random fluctuations are prevalent in modulating the circuit functions [12,17,18]. Many studies rely on stochastic simulations, such as the Gillespie algorithm [19], to compute the steady state probabilities, the transition rates, and the relative stability between different phenotypes. While stochastic simulations can provide this information, they are slow to converge and therefore extensively time consuming and less efficient in investigating the effect of the circuit architecture and dynamics in the presence of noise.

Alternatively, steady state probabilities and transition rate properties have been computed analytically or numerically by applying the WKB approximation [20]. The method was originally proposed to study the transition rate theory of nonequilibrium systems [21–23], and has been recently applied to a classical toggle switch with two coexisting states [24–26]. In a pioneering work by Wang *et al.* a similar path integral based method has been applied to genetic switches [27]. Here we address the challenge posed by three-way genetic switches that requires non-trivial generalization of the current WKB-based approaches

to generate a self-consistent landscape for systems with multiple steady states. While the method is presented in the context of genetic switches, it is applicable to general nonequilibrium systems with multiple steady states.

To meet the challenge, we introduce a new analytical approach to study the effect of white Gaussian noise and of shot noise on the dynamics of generic multistate switches. To simplify the presentation, we introduce the approach by considering the specific example of a three-way switch. The stochastic dynamics is described by a state-dependent diffusion problem, formulated in terms of a multivariate Fokker-Planck equation (FPE). Using the WKB method [28], we approximate the FPE by an eikonal equation, which is readily solved by the method of characteristics. Moreover, we develop a new contour extension method to construct the landscape contours, compute the optimal path, and calculate the relative stability by integration of the effective potential along the optimal path. We discuss how the method can be generalized for multicomponent multistate cases.

A classical toggle switch composed of two mutually inhibiting transcription factors (TFs) X and Y [inset of Fig. 1(a)] typically acts as a two-way switch [29]. A symmetric self-activating toggle switch [inset of Fig. 1(b)] acts as a three-way switch [6,30]. The deterministic dynamics of these circuits in the absence of noise (the limit of a large number of molecules and fast TF-DNA binding and unbinding) are described by the rate equations for the protein levels x and y given by [6]

$$\begin{aligned}\dot{x} &= f_x(x, y) = g_x H_{xy}^S(y) H_{xx}^S(x) - k_x x, \\ \dot{y} &= f_y(x, y) = g_y H_{yx}^S(x) H_{yy}^S(y) - k_y y,\end{aligned}\quad (2)$$

where $H_I^S(x)$ is the shifted Hill function (labeled by I) defined as $H_I^-(x) + \lambda_I H_I^+(x)$, $H_I^-(x) = 1/[1 + (x/x_I)^{n_I}]$, $H_I^+(x) = 1 - H_I^-(x)$. $\lambda_I > 1$ means activation, and $\lambda_I < 1$ means inhibition. For a classical toggle switch, the self-activation terms H_{xx}^S and H_{yy}^S are not included (see the

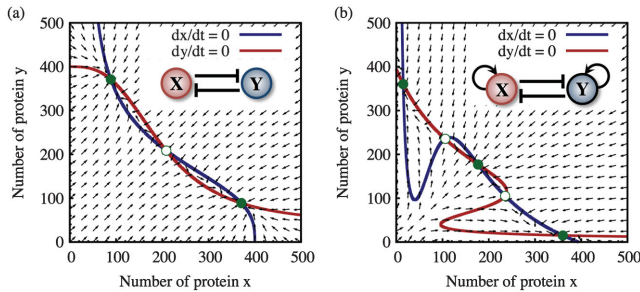


FIG. 1 (color). Deterministic dynamics of multistate genetic switches. The plots show the nullclines (navy and brown lines), stable steady states (solid green circles), saddle points (hollow green circles), and deterministic flow field (black arrows). (a) Bistable toggle switch and (b) tristable self-activating toggle switch (parameters in the Supplemental Material [31]).

Supplemental Material [31] for parameters). From Fig. 1, the two-way switch has two stable fixed points (solid green circles) and a saddle point (hollow green circle), and the three-way switch has three stable fixed points and two saddle points.

In the presence of noise originated from external signals, the stochastic dynamics can be approximated by the following nonlinear Langevin equations:

$$\dot{x} = f_x(x, y) + L_x(x, y, t), \quad \dot{y} = f_y(x, y) + L_y(x, y, t), \quad (3)$$

where $\langle L_i(x, y, t) L_j(x, y, t') \rangle = 2\epsilon D_{ij}(x, y) \delta(t - t')$ for $i, j \in x, y$, ϵ is a small scaling constant, and $\epsilon \mathbf{D}$ is the matrix of effective diffusion. Here, we consider the case of white Gaussian noise (\mathbf{D} is a unity matrix) and the case of shot noise (\mathbf{D} is a diagonal matrix with $D_{ii} \sim i$).

Following Itô's approach [33], the corresponding FPE for the probability $P(x, y)$ in this overdamped dynamics, also known as the Smoluchowski equation, is given by

$$\begin{aligned}\frac{\partial P}{\partial t} &= -\frac{\partial}{\partial x}(f_x P) - \frac{\partial}{\partial y}(f_y P) + \frac{\partial^2}{\partial x^2}(\epsilon D_{xx} P) \\ &+ \frac{\partial^2}{\partial y^2}(\epsilon D_{yy} P) + \frac{\partial^2}{\partial x \partial y}(2\epsilon D_{xy} P).\end{aligned}\quad (4)$$

For some other cases, in which the dynamics are not apparently associated with a Langevin equation, similar FPEs can still be derived in the limit of small noise. That includes the case of birth-death internal noise and gene switching noise in the limit of a large number of molecules and fast binding and unbinding [34,35] (see the Supplemental Material [31]). Thus, in general, the probability $P(x, y, t)$ is the solution of a FPE with a drift term that equals the deterministic rate $\mathbf{f}(x, y)$, and a noise-type dependent diffusion term. A typical internal noise resembles a mixture of both white Gaussian noise and shot noise. We calculate the steady state probability distribution $P(x, y)$ employing the WKB approximation [20] and use it to construct the genetic landscape according to Eq. (1).

We follow the WKB approximation method proposed by Schuss *et al.* [36], and express the stationary probability $P(x, y)$ as

$$P(x, y) = P_0 e^{-[W(x, y)/\epsilon]}, \quad (5)$$

where P_0 is a normalizing constant, and ϵ is the same small variable in the FPE. From Eq. (1), the effective potential U is proportional to W . $W(x, y)$ is the solution of the following eikonal equation that is obtained by expanding W to zero order in ϵ :

$$D_{xx} p_x^2 + 2D_{xy} p_x p_y + D_{yy} p_y^2 + f_x p_x + f_y p_y = 0, \quad (6)$$

where \mathbf{p} is an auxiliary vector representing the partial derivatives of W : $p_x = \partial W / \partial x$ and $p_y = \partial W / \partial y$. The eikonal equation can be solved by using Cauchy's method

of characteristics [28], which converts the eikonal equation and hence the original FPE into a set of parametric ordinary differential equations (ODEs) for $[x(s), y(s), p_x(s), p_y(s), W(s)]$ (with parameter s , see the Supplemental Material [31]).

These equations define characteristic curves along which W can be computed within the basin of attraction of each stable state in two stages. First, the basin of attraction is covered by a sheaf of characteristic curves beamed away from the corresponding fixed point. Second, the value of W along each characteristic curve is calculated by integrating the equation for dW/ds along the characteristic curves. The initial conditions of the ODEs are selected from a contour of constant W close to the fixed point, where $(x, y, p_x, p_y, W) = (x_0, y_0, 0, 0, c)$, and c is a constant that could be determined later. By approximating W to a quadratic form, the initial conditions can be solved numerically by the generalized eigenvalue method [37] described in the Supplemental Material [31].

The computed characteristic curves, along which W values monotonically increase when the parameter s increases, are shown in Fig. 2. The characteristics from a stable fixed point should in principle not cross the separatrix. However, we observe some violations for the regions with high W values, presumably because of the necessity for more than zero order approximation in Eq. (6). Sometimes, dramatically different characteristics can be obtained by slightly changing the initial conditions along the initial contour. We therefore devise an adaptive method (see the Supplemental Material [31]) to sample the initial conditions in a special way so that the characteristic curves cover the whole phase plane as uniformly as possible. The regimes that are not covered by the characteristics (Fig. 2) require more refined sampling and higher order approximation in the locations of high W . Of particular interest is the characteristic curve that connects

the stable fixed point with the saddle point and is associated with the optimal path [24]. This characteristic curve, which is calculated by a new method we developed (see below), enables us to calculate ΔW , the difference between the value of W at the saddle point and at the stable fixed point, and hence the relative stability of the different states.

Construction of the global effective potential requires calculating the values of W for the different stable fixed points, i.e., the constants c defined above. This is done by first defining the level of W to be zero for a fixed point, and then, for each pair of saddle point and stable point, calculating the integral of dW/ds along the optimal path. Figure 2 shows the computed characteristic curves and W for both the two-way and the three-way switches. This figure illustrates the efficiency of the characteristic curves in mapping the effective landscape.

In principle, once the sheaf of characteristic curves is mapped, the contours of constant W are well defined and can be constructed. Yet, doing so requires a special method to handle the ultrasensitivity of the characteristic curves to the initial conditions. Therefore, we devise a contour extension (CE) method, to efficiently compute the landscape contours for a two-dimensional system. The idea is to utilize the method of characteristics in a special iterative way. At each iteration, a new contour is calculated from the previous contours by interpolating the contour points from the characteristic curves initiated from points along the previous contour. More specifically, for any point (x', y', p'_x, p'_y, W') , we have $0 \equiv dW = p'_x dx + p'_y dy$ along the W' contour. So the slope of the contour at this point is $dy/dx = -p'_x/p'_y$, and the tangent line is $y = y' - (x - x')p'_x/p'_y$. For any two adjacent points A and B on the same contour [Fig. 3(a)], we construct the contour line between A and B by interpolation. First, we find an auxiliary point C , which is the intersection of the contour tangent lines that go through the points A and B . Second, the contour line can be approximated by the quadratic Bezier curve from A , B , and C . Any point D on the contour is given by

$$D(\alpha) = (1 - \alpha)^2 \mathbf{A} + 2(1 - \alpha)\alpha \mathbf{C} + \alpha^2 \mathbf{B}, \quad (7)$$

where \mathbf{A} , \mathbf{B} , and \mathbf{C} are the (x, y) vectors for points A , B , and C , respectively, and α is a parameter from 0 to 1.

By using this interpolation technique, we construct the landscape contours by the following iterative steps. First, several characteristics curves are computed starting from the points on the initial contour until the W values exceed a certain threshold value or the contour curvature between any two adjacent points is too large. Here, the contour curvature can be approximated by the difference between the contour slopes (in angle) of the two contour points. Second, the contours are constructed from the contour points by the above interpolation method. Third, the contour points are resampled along the whole contour

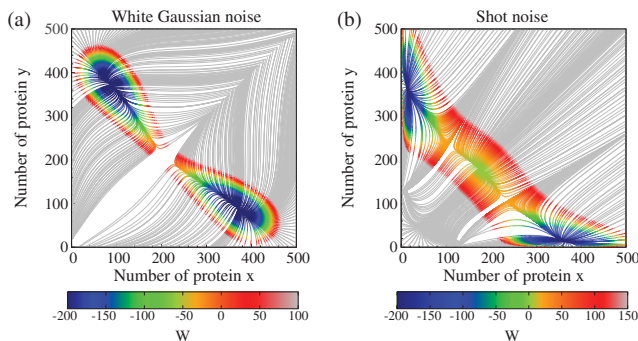


FIG. 2 (color). Construction of effective landscape by the WKB-based approach. Characteristic curves (solid lines, colors representing the W values) are generated by the method of characteristics, starting from the initial contours around each stable fixed point. Some white spaces are not covered by the curves because of under sampling. (a) Two-way switch with white Gaussian noise and (b) three-way switch with shot noise.

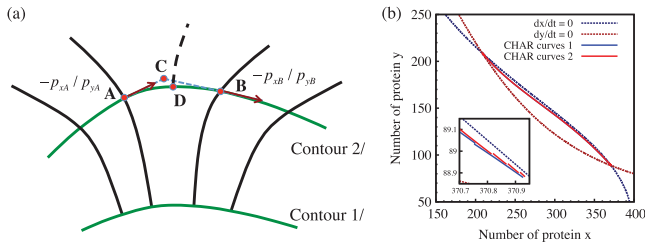


FIG. 3 (color). Illustration of the CE method for computing landscape contours and optimal paths. (a) The interpolation method to construct contour lines. Starting from points on contour 1, characteristic curves are generated until they reach contour 2. The contour line between two adjacent contour points A and B can be interpolated by the quadratic Bezier curves constructed from the points A , B , and C , where the point C is the intersection of the two contour slope lines through A and B . For the next iteration, the initial conditions are selected from points on contour 2. (b) The optimal path for the two-way switch. The plot shows the nullclines (navy and brown dotted lines) and the optimal path connecting the stable point and the saddle point. Our numerical methods generate two sets of line segments of the characteristic curves, shown in blue and red solid lines (which almost overlap, see inset).

according to the point density and the contour curvature [for more details, see Fig. 3(a) and the Supplemental Material [31]].

We also introduce a new method to compute the optimal path by the CE and a special iterative bisection method [Fig. 3(b)]. First, we select two points on the initial contour on either side of the crossing of the optimal path and the contour. Second, we advance the two points for a short distance along the two characteristic curves that pass through these points, while computing the W values and making sure to advance the two points such that the W values are kept equal. Third, once the two points are far away, a contour segment is constructed between the two new points and the middle point ($\alpha = 0.5$) is calculated from Eq. (7). Fourth, the middle point replaces one of the previous two contour points, so that the optimal path is still between the two characteristic curves starting from these two new points. The bisection process is repeated until the two points reach to the proximity of the saddle points.

Figure 4 shows an example of the landscape contours, optimal path, and relative stability (in terms of an effective potential along the optimal path) for the three-way switch in the presence of white Gaussian and shot noise. From the figure, the white Gaussian noise can stabilize the intermediate state while the shot noise destabilizes it, presumably because the shot noise level in the intermediate state is relatively higher than that in the other two states. Since the gradient of W for any point can be calculated from the contours, we also check the steepest descent path of W , which starts from a saddle point and ends at a stable point (red dotted line). From Fig. 4, the steepest descent path is different from the optimal path, suggesting that the

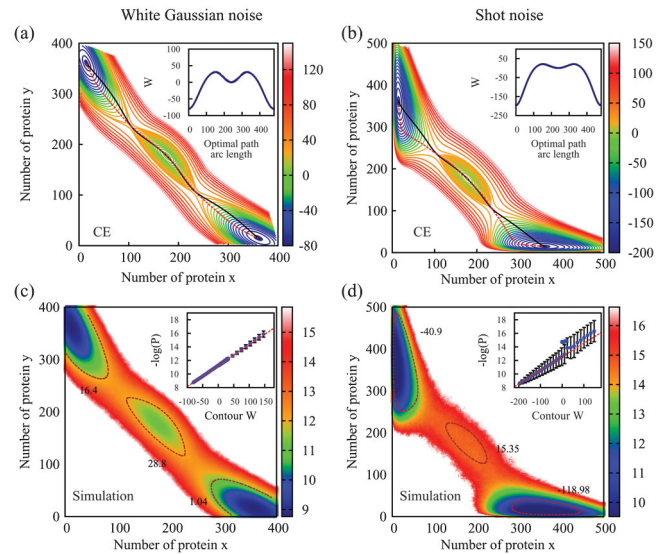


FIG. 4 (color). Comparison of the landscape contours with stochastic simulations for the three-way switch. Left panels: Gaussian noise; right panels: shot noise. Panels (a) and (b) show the CE contours (colors represent W values), optimal paths (black solid lines), and steepest descent paths (dotted red lines). The insets show W along the optimal path. Panels (c) and (d) show $-\log(P)$ from Langevin simulations (colors selected to better match the W values; see the Supplemental Material [31]). The insets show the mean (blue dots) and standard deviations (error bars) of $-\log(P)$ along each CE contour (e.g., contours illustrated as brown dashed lines, see Sec. 8 of the Supplemental Material [31]). Linear fitting was performed on the data points for the contours within the top left and the bottom right basins. The slope of the linear fitting coincides with ϵ^{-1} , as expected from the theory. In panel (d), some contour lines are shifted and the middle basin is undersampled because of large shot noise.

probability flux also plays an important role in the non-equilibrium dynamics, and the transition path is irreversible [27]. Quantitative comparison with the results from stochastic simulations reveals an excellent agreement between the two methods [Figs. 4(c) and 4(d) and Sec. 8 of the Supplemental Material [31]], except for the middle basin for the three-way switch with shot noise [Fig. 4(d)], where the shallow basin is undersampled because of large noise. As shown in detail in the Supplemental Material [31], some deviations of the stochastic simulations from the analytical or computational results reflect the ineffectiveness of the stochastic simulations in constructing an accurate landscape due to the need for consistently large sampling of the individual basins. Therefore, the deviations manifest the advantage of having an analytical or computational method for the landscape construction. According to the simulations (Fig. S7 [31]), the linear relationship $W \sim \log P$ is valid for a wide range of ϵ values, while larger deviations were observed for larger ϵ and W values (Figs. S5 and S6 [31]). Such deviations at the states with high W values may have a negligible effect, as they are usually far away from the steady state conditions.

To conclude, the method presented here can be readily extended to systems with more than two components. For example, in the study of a micro-RNA-based circuit, the system has at least three components, including micro-RNA, mRNA, and protein molecules [6,7]. The WKB-based method can be extended for multidimensional systems (see the Supplemental Material [31]). However, since sampling the initial conditions requires extensive computations, an efficient approach will be to first define one or two reaction coordinates, and then to apply the construction method. One possible solution would be the method in which the slow variables are first obtained from simulations by a diffusion map [38,39], and the corresponding FPE is established by the equation-free approach [40]. Some mean field approximation methods [41–44] also have recently been proposed to deal with multidimensional cases. It would be also interesting to compare the different approaches on a large-scale network.

We thank Zeev Schuss, Herbert Levine, and Sui Huang for useful discussions. This work was supported by the NSF Center for Theoretical Biological Physics (NSF Grant No. PHY-1427654), and by the Cancer Prevention and Research Institute of Texas (CPRIT).

*jonuchic@rice.edu

†eshelbj@gmail.com

- [1] G. Balázsi, A. van Oudenaarden, and J. J. Collins, *Cell* **144**, 910 (2011).
- [2] S. Huang, I. Ernberg, and S. Kauffman, *Semin. Cell Dev. Biol.* **20**, 869 (2009).
- [3] J. X. Zhou, M. D. S. Aliyu, E. Aurell, and S. Huang, *J. R. Soc. Interface* **9**, 3539 (2012).
- [4] J. P. Thiery, H. Acloque, R. Y. J. Huang, and M. A. Nieto, *Cell* **139**, 871 (2009).
- [5] P. Mehlen and A. Puisieux, *Nat. Rev. Cancer* **6**, 449 (2006).
- [6] M. Lu, M. K. Jolly, R. Gomoto, B. Huang, J. Onuchic, and E. Ben-Jacob, *J. Phys. Chem. B* **117**, 13164 (2013).
- [7] M. Lu, M. K. Jolly, H. Levine, J. N. Onuchic, and E. Ben-Jacob, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 18144 (2013).
- [8] P. Friedl and D. Gilmour, *Nat. Rev. Mol. Cell Biol.* **10**, 445 (2009).
- [9] D. Schultz, M. Lu, T. Stavropoulos, J. Onuchic, and E. Ben-Jacob, *Sci. Rep.* **3**, 1668 (2013).
- [10] H. Maamar, A. Raj, and D. Dubnau, *Science* **317**, 526 (2007).
- [11] B. Munsky, G. Neuert, and A. van Oudenaarden, *Science* **336**, 183 (2012).
- [12] J. Paulsson, *Nature (London)* **427**, 415 (2004).
- [13] P. Hänggi, P. Talkner, and M. Borkovec, *Rev. Mod. Phys.* **62**, 251 (1990).
- [14] J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, *Annu. Rev. Phys. Chem.* **48**, 545 (1997).
- [15] J. Wang, K. Zhang, L. Xu, and E. Wang, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 8257 (2011).
- [16] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (Westview, Boulder, CO, 1994).
- [17] M. Kittisopikul and G. M. Süel, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13300 (2010).
- [18] G. M. Süel, R. P. Kulkarni, J. Dworkin, J. Garcia-Ojalvo, and M. B. Elowitz, *Science* **315**, 1716 (2007).
- [19] D. T. Gillespie, *J. Phys. Chem.* **81**, 2340 (1977).
- [20] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory* (Springer, New York, 1999).
- [21] E. Ben-Jacob, D. J. Bergman, B. J. Matkowsky, and Z. Schuss, *Phys. Rev. A* **26**, 2805 (1982).
- [22] R. Kupferman, M. Kaiser, Z. Schuss, and E. Ben-Jacob, *Phys. Rev. A* **45**, 745 (1992).
- [23] R. S. Maier and D. L. Stein, *Phys. Rev. E* **48**, 931 (1993).
- [24] D. M. Roma, R. A. O’Flanagan, A. E. Ruckenstein, A. M. Sengupta, and R. Mukhopadhyay, *Phys. Rev. E* **71**, 011902 (2005).
- [25] M. Assaf, E. Roberts, and Z. Luthey-Schulten, *Phys. Rev. Lett.* **106**, 248102 (2011).
- [26] C. Lv, X. Li, F. Li, and T. Li, *PLoS One* **9**, e88167 (2014).
- [27] J. Wang, K. Zhang, and E. Wang, *J. Chem. Phys.* **133**, 125103 (2010).
- [28] L. C. Evans, *Partial Differential Equations* (American Mathematical Society, Providence, 2010).
- [29] T. S. Gardner, C. R. Cantor, and J. J. Collins, *Nature (London)* **403**, 339 (2000).
- [30] R. Guantes and J. F. Poyatos, *PLoS Comput. Biol.* **4**, e1000235 (2008).
- [31] See Supplemental Material <http://link.aps.org/supplemental/10.1103/PhysRevLett.113.078102>, which includes Ref. [32], for definition of the Kullback-Leibler distance.
- [32] S. Kullback and R. A. Leibler, *Ann. Math. Stat.* **22**, 79 (1951).
- [33] N. G. V. Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd. ed. (Elsevier, New York, 2007).
- [34] T. B. Kepler and T. C. Elston, *Biophys. J.* **81**, 3116 (2001).
- [35] A. M. Walczak, J. N. Onuchic, and P. G. Wolynes, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18926 (2005).
- [36] Z. Schuss, *Theory and Applications of Stochastic Processes: An Analytical Approach* (Springer-Verlag, Berlin, 2012).
- [37] W. F. Arnold III and A. J. Laub, *Proc. IEEE* **72**, 1746 (1984).
- [38] R. Erban, T. A. Frewen, X. Wang, T. C. Elston, R. Coifman, B. Nadler, and I. G. Kevrekidis, *J. Chem. Phys.* **126**, 155103 (2007).
- [39] A. Singer, R. Erban, I. G. Kevrekidis, and R. R. Coifman, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 16090 (2009).
- [40] R. Erban, I. G. Kevrekidis, D. Adalsteinsson, and T. C. Elston, *J. Chem. Phys.* **124**, 084106 (2006).
- [41] A. M. Walczak, M. Sasai, and P. G. Wolynes, *Biophys. J.* **88**, 828 (2005).
- [42] C. Li and J. Wang, *PLoS Comput. Biol.* **9**, e1003165 (2013).
- [43] H. Yan, L. Zhao, L. Hu, X. Wang, E. Wang, and J. Wang, *Proc. Natl. Acad. Sci. U.S.A.* **110**, E4185 (2013).
- [44] Z. Zhang and P. G. Wolynes, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 10185 (2014).