Transitions in genetic toggle switches driven by dynamic disorder in rate coefficients

Hang Chen,[a] Peter Thill, and Jianshu Cao
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

(Received 15 January 2016; accepted 20 April 2016; published online 5 May 2016)

In biochemical systems, intrinsic noise may drive the system switch from one stable state to another. We investigate how kinetic switching between stable states in a bistable network is influenced by dynamic disorder, i.e., fluctuations in the rate coefficients. Using the geometric minimum action method, we first investigate the optimal transition paths and the corresponding minimum actions based on a genetic toggle switch model in which reaction coefficients draw from a discrete probability distribution. For the continuous probability distribution of the rate coefficient, we then consider two models of dynamic disorder in which reaction coefficients undergo different stochastic processes with the same stationary distribution. In one, the kinetic parameters follow a discrete Markov process and in the other they follow continuous Langevin dynamics. We find that regulation of the parameters modulating the dynamic disorder, as has been demonstrated to occur through allosteric control in bistable networks in the immune system, can be crucial in shaping the statistics of optimal transition paths, transition probabilities, and the stationary probability distribution of the network. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4948461]

I. INTRODUCTION

Single molecule studies have demonstrated that enzymatic behavior is different from the static picture gleaned from ensemble averaged experiments,[1–3] and a growing body of theoretical work has examined the consequences of fluctuating chemical kinetics in protein interaction networks.[4,5] Enzymatic rates of catalysis can fluctuate over several orders of magnitude.[6,7] Recent single molecule studies from Iversen *et al.* have shown that this type of behavior is even exhibited in bistable networks that are involved in the earliest stages of T-cell signaling in the immune system.[8] The particular behavior outlined in that work shows that enzymatic rate constants can remain at a fixed value for an interval of time before jumping and resampling from a heavy tailed distribution.[8,9] Throughout this paper, we will model such a behavior by a stochastic process in which the time intervals between switching events are exponentially distributed with a parameter λ s^{-1} that we call the switching rate. The single molecule studies find that, interestingly, allosteric regulation can affect both the switching rate λ, and the shape of the probability distribution $p(k)$ that the rate parameters draw from. It was shown through computational studies that the presence of dynamic disorder[10] can make the network more stable than it would be if all kinetic parameters acted at a single average value and did not fluctuate, in the sense that a stronger signal input would be required for the network to switch from one basin to the other.

In this paper, we investigate the effects of dynamic disorder in a general toggle switch model with bistability.[11–15] We first investigate a relatively simple model where the rate parameter k draws from a discrete probability distribution. This discrete distribution model is simple enough for us to examine the optimal transition paths between stable states and the corresponding actions by using the geometric minimum action method (gMAM).[16,17] The results of the optimal transition paths illustrate a deviation between the optimal path and the average transition path obtained from chemical rate equations. And, as expected, the optimal paths converge to the average path when we increase the switching rate. The increasing switching rate also dramatically reduces the minimum action, which means that the dynamic disorder makes the system more stable.

We then consider a more complex genetic toggle switch model in which the rate parameter k draws from a continuous heavy tailed distribution. We see that by tuning the shape of the distribution $p(k)$ from which a particular rate constant samples, or by tuning the switching rate λ, the equilibrium probability mass function (PMF) is significantly changed and therefore one stable basin can be greatly stabilized relative to the other. This means that the dynamic disorder (and specifically allosteric control of the parameters of the model) can dramatically influence the mean first passage time for transition from one stable basin to the other. We investigate this behavior both in the small system limit by looking at Gillespie simulations of the toggle switch network, and also in the large volume, fixed concentration limit by using numerical procedures to investigate how the minimum action path is influenced by the presence of dynamic disorder. We find that the results are qualitatively unchanged when the stochastic rate parameter follows a Markov jump process as opposed to stochastic Langevin dynamics.

[a]Electronic mail: hangchen@mit.edu

II. GENETIC TOGGLE SWITCHES WITH THREE ISOMERS

In genetic toggle switches, a pair of genes mutually represses each other.[11–13] A pair of genes encodes proteins A and B, respectively, called transcription factors. Protein A and B in turn form homodimers A_2 and B_2 which bind to regulatory regions of DNA, called operators (O_a and O_b), of the respective other gene. The binding of A_2 to the operator O_b represses the production of protein B and vice versa. We count the total copy numbers of the transcription factors A (N_A) and B (N_B) which include those in homodimers and those bound to the operators. Within a region in parameter space, the genetic toggle switches have two possible stable states: a state with a large concentration of c_A and a small concentration of c_B, or the other state with small c_A and large c_B.[14]

To incorporate dynamic disorder into the model, we investigate a model in which, due to fluctuating conformational changes, the unbinding rate of A_2 from the operator O_b is not fixed but instead fluctuates. We study an asymmetric model in which only polymer A_2 has a fluctuating unbinding rate. To examine the optimal transition paths and find the corresponding minimum actions, we begin with a relatively simple model where the unbinding rate of O_bA_2 samples from a discrete probability distribution, $P_1 = 0.1$, $P_2 = 0.8$, and $P_3 = 0.1$. In other words, there are three conformations of polymer A_2, A_2^1, A_2^2, and A_2^3 with different unbinding rates, and these homodimers switch among each other governed by a rule of transition probability matrix,

$$R = \begin{array}{c} \\ A_2^1 \\ A_2^2 \\ A_2^3 \end{array} \begin{array}{c} \begin{array}{ccc} A_2^1 & A_2^2 & A_2^3 \end{array} \\ \begin{vmatrix} R_1 & R_2 & R_3 \\ R_1 & R_2 & R_3 \\ R_1 & R_2 & R_3 \end{vmatrix} \end{array},$$

where $R_1 = 0.1, R_2 = 0.8, R_3 = 0.1$.

We model the genetic toggle switches as a discrete Markov jump process with the reaction scheme in Figure 1 and Table I. The values of the rate constants are chosen based on Warren's paper.[14] We assume that there is only one copy of the genome in the genetic toggle switch.

A. Optimal transition paths

As described in Sec. I, genetic switches normally have bimodality and the intrinsic noise can drive the genetic switches from one stable state to the other. From the large deviation theory, the probability of a rare transition between stable states is dominated by the probability of the optimal transition path with the minimum action.[18,19] In this section, we investigate the most probable transition paths for the genetic switches with different switching rates (λ) by implementing the geometric minimum action method (gMAM).[17]

We first obtain the stable states by solving ordinary differential equations (ODEs) with different switching rates based on the reaction scheme in Table I. We plot these stable states in Figure 2. The x and y axes represent the concentrations of species A ($C_A = c_A + 2(c_{A_2^1} + c_{A_2^2} + c_{A_2^3} + c_{O_bA_2^1} + c_{O_bA_2^2} + c_{O_bA_2^3})$) and B ($C_B = c_B + 2(c_{B_2} + c_{O_aB_2})$), respectively. Note that stable states, for both high A (labeled as a in Figure 2) and low A (labeled as b), of the system with different switching rates are distinct from each other. As the switching rate increases, the stable states shift in the direction of less A and more B (top left in Figure 2). The reason is that polymers A_2 switch among each other on operators and the overall behavior will be closer to the situation with mean value of the unbinding rate, $\bar{k}_u = 5$, which is larger than the most probable value of the unbinding rate, $k_u = 4.625$. And the larger unbinding rate of A_2 leads to more production of B.

By using the gMAM method, we then find the forward (from a to b) and backward (from b to a) optimal paths between these stable states, which are drawn in solid lines and dashed lines, respectively, in Figure 2. The forward and backward optimal paths do not coincide, which implies that the system does not obey detailed balance. More interestingly, when we increase the switching rate, both forward and backward paths converge to the middle green lines. The reason for the convergence is probably because the large switching rate diminishes the stochastic effect of the system with fluctuating conformational changes. To validate this argument, we then investigate the mean paths of the transitions. We choose two different state points, c and d, which are close

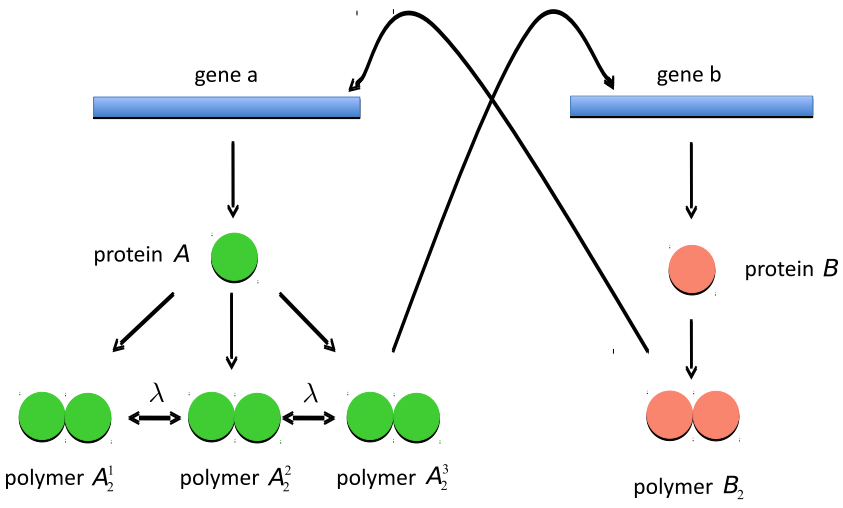

FIG. 1. The model of the genetic toggle switches. Genes a and b encode proteins A and B, respectively. Protein A and B in turn form homodimers A_2, with three conformations, and B_2 which bind to regulatory regions of DNA of the respective other gene. The binding of A_2 to the operator O_b represses the production of protein B and vice versa.

TABLE I. Reaction scheme of genetic toggle switches with three isomers of A_2. (i) k_f and k_b are the forward and backward rate constants of the corresponding reactions. (ii) $P_{1,2,3}$ are the distribution probabilities of three polymers A_2. (iii) $R_{1,2,3}$ are the switching probabilities of three polymers A_2. (iv) λ is the switching rate.

	Reactions		k_f	k_b
Production of proteins	$O_a \to O_a + A$	$O_b \to O_b + B$	1	
Degradation of proteins	$A \to \emptyset$	$B \to \emptyset$	0.8	
Formation of dimers	$A + A \rightleftharpoons A_2^1$	$B + B \rightleftharpoons B_2$	$10P_1$	5
	$A + A \rightleftharpoons A_2^2$		$10P_2$	5
	$A + A \rightleftharpoons A_2^3$		$10P_3$	5
Binding/unbinding to operators	$O_b + A_2^1 \rightleftharpoons O_bA_2^1$	$O_a + B_2 \rightleftharpoons O_aB_2$	10	3
	$O_b + A_2^2 \rightleftharpoons O_bA_2^2$		10	4.625
	$O_b + A_2^3 \rightleftharpoons O_bA_2^3$		10	10
Switching among dimers	$A_2^1 \rightleftharpoons A_2^2$		λR_2	λR_1
	$A_2^1 \rightleftharpoons A_2^3$		λR_3	λR_1
	$A_2^2 \rightleftharpoons A_2^3$		λR_3	λR_2
	$O_bA_2^1 \rightleftharpoons O_bA_2^2$		λR_2	λR_1
	$O_bA_2^1 \rightleftharpoons O_bA_2^3$		λR_3	λR_1
	$O_bA_2^2 \rightleftharpoons O_bA_2^3$		λR_3	λR_2

to a and b, respectively. Because they are unstable states, they will move to the stable states a and b, respectively.

We obtain the average transition paths by solving ODE where the system evolves with deterministic characteristics. The results in Figure 3(a) show that, from c to b, the average transition path is very different from the optimal paths when the switching rate is relatively small. However, when the switching rate is large, the optimal path coincides with the average path obtained by solving ODE. This observation demonstrates the decline of the stochastic characteristics due to the enhanced switching behavior among isomers.

To clearly see the difference between the optimal paths and the average path, we plot the deviation between them as a function of the concentration of A in Figure 3(b). We pick three switching rates 0.5, 1, and 10 and plot the deviations correspondingly. When the concentration of A is smaller, the deviation is larger due to the fluctuation. Also, a smaller switching rate results in a larger deviation.

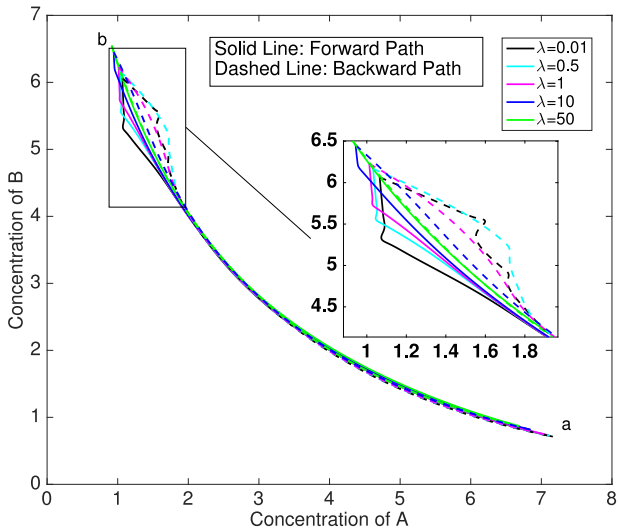

FIG. 2. The optimal paths with the minimum actions. The solid lines and dashed lines are forward (from a to b) and backward (from b to a) transition paths, respectively. The x and y axes are the concentrations of species A and B, respectively. Different colored lines represent systems with different switching rates. The inset highlights the difference and convergence of the optimal paths around state point b.

B. The minimum action

According to the large deviation theory, the transition rate of a rare event is a negative exponential function of the action of the optimal path, $K \approx A\exp(-VS^*)$, where S^* is the minimum action and V is the volume of the system.[20] If the minimum action is smaller, the probability of the transition will be larger. Thus, we use the minimum action to describe the probability of the transition between stable states.

In Figure 4, we plot the minimum action as a function of the switching rate among isomers A_2. The action decreases dramatically as the switching rate increases, and then plateaus for large switching rate. The decrease of the action indicates that a larger switching rate makes this system easier to shift from high A state to low A state. The reason for the plateau is that, at the large switching rate, the isomers A_2 are well mixed and the system behaves as the average, where the action reaches the lower-bound limit. The results demonstrate that the fluctuations of conformational polymers slow down the rare transition between stable states of the system.

III. CONTINUOUS SPACE MODEL

A. Discrete jumps

We then extend the simple genetic toggle switch model to a more complex one where rate coefficient samples from a continuous heavy tailed distribution instead of a discrete distribution. We begin by investigating the behavior of this toggle switch network in the small copy number limit using the Gillespie algorithm.[21–23] We first carry out the simulations

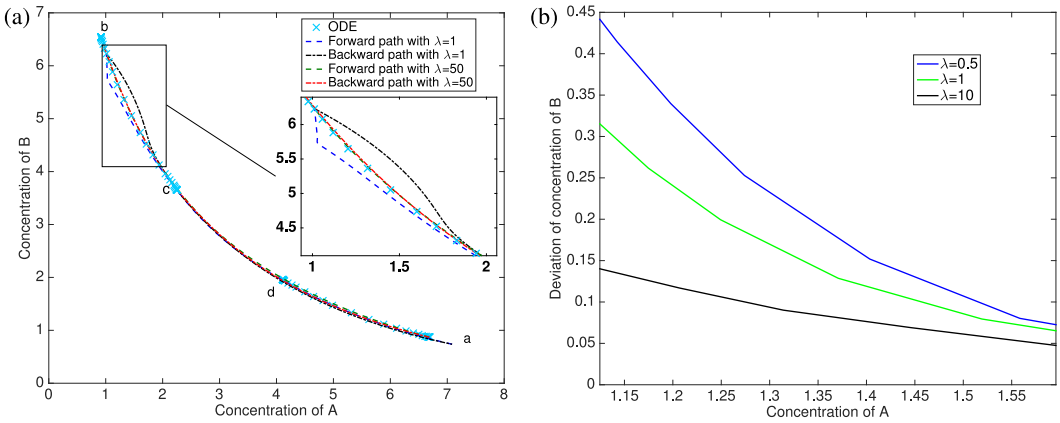

FIG. 3. (a) The optimal path with the minimum action is different from the average paths. The forward and backward optimal transition paths are indicated by dashed black and green lines, respectively. The average paths are indicated by cyan cross lines. The inset highlights the difference between the optimal paths and the average path between states b and c. (b) The deviation between the optimal paths and the average path as a function of the concentration of A.

for a model of dynamic disorder in which individual rate parameters remain at one value for a random, exponentially distributed amount of time (parameterized by a switching rate λ s^{-1}), before redrawing a value from a prescribed distribution. Following experimental evidence which shows fluctuating rate constants draw from heavy tailed distribution,[3,7,8] we allow the rate constants to sample from a log normal distribution. We carry out these simulations with only one copy of genome in the simulation box like the discrete distribution case in Sec. IV. For this model, the state space is continuous as an individual rate constant samples from a continuous probability distribution, but the jumps are discrete.

Because of the fluctuating conformational changes of polymer A_2, the unbinding rate of A_2 to the operator O_b samples from a lognormal distribution. We fix the mean value of the lognormal distribution at a value known to yield a bistable network and look at the effect of changing the variance/skewness of the distribution. For the lognormal distribution, changing the variance while keeping the mean fixed requires us to change both the μ and σ parameters of the exponentiated normal. We first investigate a model in which the fluctuation in rate parameter occurs only when the dimer A_2 is unbound from the operator O_b, but after dimer A_2 binds to the operator, it is fixed in one conformational change.

Throughout, the average value that an individual rate constant draws from is chosen to be 5. We find that, upon increasing the standard deviation in the distribution that a rate constant k chooses from over a range from less than one to ten, the mean first passage time to transition from a state of high A to low A (including those in homodimers and those bound to the operators) is significantly increased (Figure 5(a)).

The physical explanation behind the dramatic increase in mean first passage time is that, as we increase the standard deviation of the underlying lognormal distribution that an individual rate parameter draws from, we change the skew of the distribution. The consequence is that the most probable value (mode) of the probability distribution an off rate of A_2 is drawn from decreases to lower and lower values. This in turn means that a polymer A_2 will remain bound for a longer time on average, and thus does a better job at repressing the production protein B. Mathematically, if we specify a mean $\mathbb{E}[k]$ and variance $\mathrm{Var}(k)$ for an individual off rate k, we determine the parameters of the lognormal distribution by the inversion,

$$\sigma = \sqrt{\ln\left(\frac{\mathrm{Var}(k)}{\mathbb{E}[k]^2} + 1\right)}, \quad \mu = \ln(\mathbb{E}[k]) - \frac{\sigma^2}{2}. \tag{1}$$

The parameter σ increases monotonically with $\mathrm{Var}(k)$ and the skewness of the lognormal distribution (given by $(e^{\sigma^2} - 1)\sqrt{e^{\sigma^2} - 1}$) grows monotonically with σ. The parameter μ decreases monotonically with increasing σ, and the mode of the lognormal distribution, given by $e^{\mu - \sigma^2}$ decreases monotonically with $\mathrm{Var}(k)$. Therefore, mathematically we shift the most probable values of the off rate to values that are much lower than the average $\mathbb{E}[k]$.

As expected, initializing the system in the low A basin (using Monte Carlo sampling of the network to pick the initial point) and measuring the mean first passage time to transition to a state of high A follows the opposite trend noted above. Namely, increasing the variance of the lognormal distribution serves to decrease the mean first passage time for this reverse transition (Figure 5(b)).

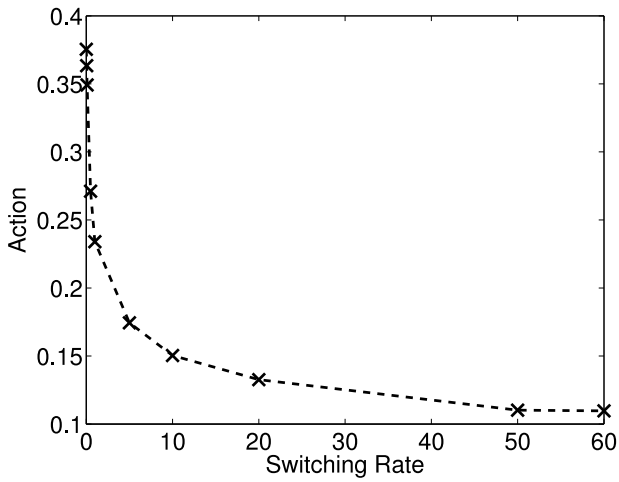

FIG. 4. The minimum action as a function of switching rate. The data points are marked by crosses. The action decreases dramatically as the switching rate increases, and then plateaus for large switching rate.

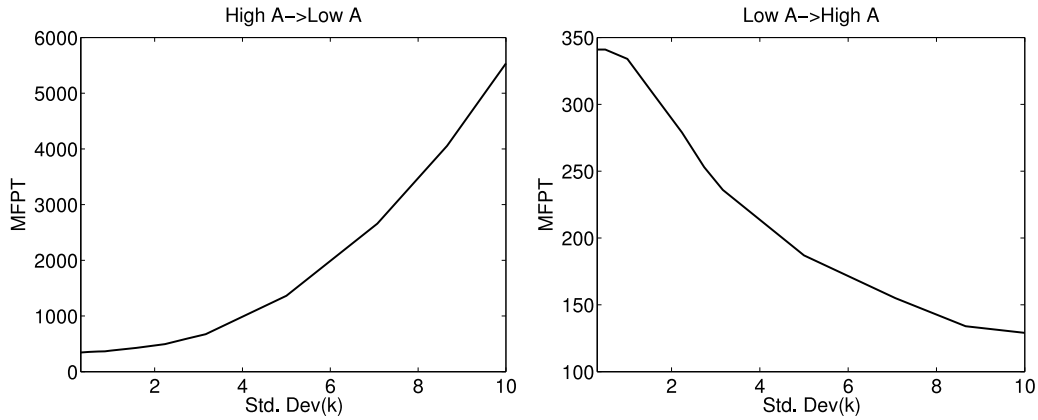

FIG. 5. (a) Here we plot the mean first passage time to transition from a state of high A to a state of low A. The stable basins are found by Monte Carlo simulation, and we use $A = 0, B \geq 7$ for low A state and $B = 0, A \geq 7$ for high A state. The x axis is the standard deviation in the lognormal distribution for k, not the parameter σ of the lognormal. (b) Here we plot the mean first passage time to transition from a state of low A to a state of high A as a function of increasing the standard deviation in the distribution of off rates for an individual polymer A from the operator. The trend is the reverse of that from Figure 5(a).

Indeed, studying the probability mass function (PMF) for the Toggle switch model with dynamic disorder shows that the basin of low A and high B is destabilized upon increasing the variance in the off rate of an individual polymer A_2 (Figure 6).

Of course, for a model in which the polymer does not switch when it is bound to the operator, we cannot see much interesting behavior with respect to the parameter λ. As long as the average time for an individual polymer to rebind is larger than $1/\lambda$, the polymer will on average have resampled its kinetic off rate before rebinding. Upon increasing λ any further, the statistics of the network should be unchanged since it will not matter whether the polymer has resampled its kinetic off rate once or multiple times.

To examine a model that is more sensitive to λ, we simulated the same network but instead let the polymer conformation fluctuate when it is bound to the operator. What we see in this case is also fairly intuitive. When the switching rate is very low and σ is large, a polymer that has just sampled a kinetic off rate that is far below the average value will get frozen at that particular value for the duration of its time on the operator. It will then, on average, remain bound for a long time and do a better job at inhibiting the production of protein B. If we increase the switching rate λ, we expect that the system will begin to behave as if it was acting at its average value. This can be shown in more mathematically precise language using renewal theory.[24] Therefore, as we increase the switching rate λ, the mean first passage time for transition from high A to low A decreases (Figure 7).

B. Continuous jumps

To demonstrate that the results are robust to the type of model we consider, we also look at a continuous model in which the kinetic off rate evolves continuously following a Langevin dynamics of the following form:

$$k_t = \exp(\tilde{k}_t), \quad \mathrm{d}\tilde{k}_t = \lambda(\mu - \tilde{k}_t)\mathrm{d}t + \left(\sqrt{2\sigma^2\lambda}\right)\mathrm{d}Z_t. \quad (2)$$

That is, the rate parameter is an exponential of an Ornstein-Uhlenbeck (OU) process. We let $\tilde{k}_t$ denote the value of the OU process at time t, and k_t the value of the actual kinetic parameter. This process is chosen since the steady state distribution of the OU process is a normal distribution, and thus the stationary distribution of the k_t is lognormal,

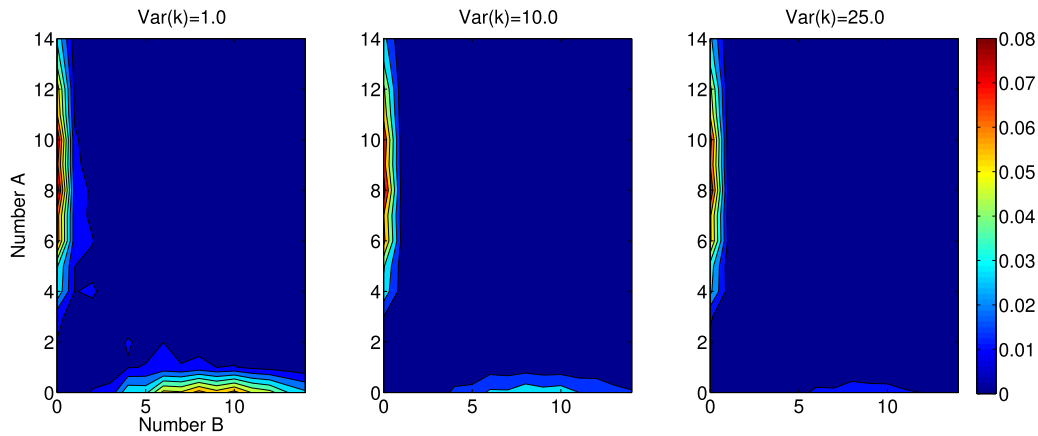

FIG. 6. Joint PMF's for the (number of A, number of B) molecules are plotted here (for $\lambda = 1.0$). These PMF's are plotted for the model in which kinetic off rates vary continuously on the DNA (though the location of peaks is unchanged for the discrete model). By increasing the value of var(k) of the underlying stochastic process, we see that the stable basin corresponding to high B and low A is destabilized. The colors vary between 0 and 0.08. These are calculated by running a single trajectory of simulation length 10^5 s with a time increment of $\mathrm{d}t = 10^{-3}$ and sampling every 1 s.

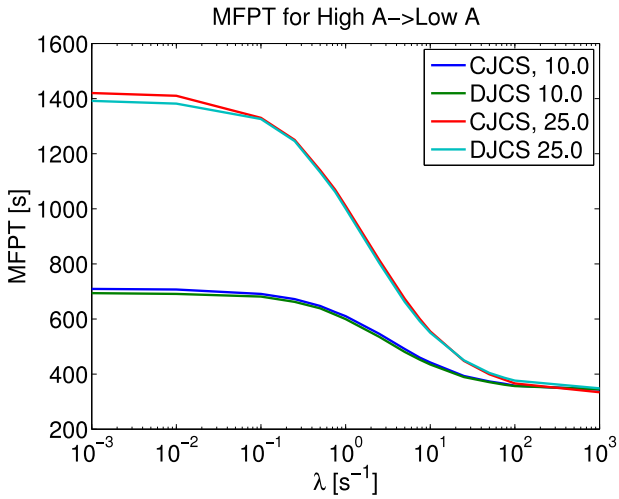

FIG. 7. Mean first passage times to transition from $A \geq 7, B = 0$ to $A = 0, B \geq 7$ are computed for various values of λ and $\mathrm{Var}(k) = 10, 25$. We do the calculation for both a discrete jump model and a continuous jump model, labeled DJCS, CJCS, respectively. In these parameter regimes, the mean first passage times are comparable for both models. Importantly, increasing λ allows faster transitions as the system behaves closer to the mean value. Increasing $\mathrm{Var}(k)$ increases the skewness and therefore decreases the MFPT. For the $\lambda = 10^3$ value in the continuous jump model, a time discretization of 10^{-4} is used.

which will mimic the stationary distribution sampled in the previous model. The value λ plays functionally the same role as it did in the previous case. Though it does not appear in the overall stationary distribution of the OU process, it does modulate how rapidly the value k_t changes in an instant of time dt.

Simulating the above process is much more computationally expensive. Now, we must discretize time by picking a dt such that the sum of all rates for various reaction events in the continuous time Markov process simulated by the Gillespie process above is small relative to dt. Let the rate for reaction event i be ρ_i. Then, in an instant of time dt, we first ask whether a reaction occurred or not, where the probability for the reaction to occur is the cumulative distribution function (CDF) of the exponential distribution with parameter $\sum_i \rho_i$. If a reaction did occur, we choose its value from the multinomial defined by the ρ_i. Importantly, dt must be sufficiently small such that the probability for more than one reaction event to occur is essentially zero.

We find qualitatively the same results as before. Increasing the value of $\mathrm{Var}(k)$ at a fixed λ will increase the mean first passage time from high A to low A, while increasing λ at a fixed value of $\mathrm{Var}(k)$ will decrease the mean first passage time from high A to low A (Figure 7). Quantitatively, the mean first passage time for both the continuous and discrete models is very similar, with the continuous model appearing to have a slightly larger MFPT for low λ. We explore this in more detail in the Appendix.

IV. CONCLUSION

In this paper, we have studied how kinetic switching between stable states in biological networks is influenced by dynamic disorder or conformational fluctuations in the rate coefficients. We carried out simulation and analysis based on a general genetic toggle switch model. In the model, the unbinding rate of polymers from the operators fluctuates with the switching rate, λ.

First, we have studied a model where the rate parameter samples from a discrete probability distribution. The model is simple enough for us to study the optimal transition paths and the minimum actions by using large deviations theory and the geometric minimum action method (gMAM). We varied the switching rate over three orders of magnitude to reflect fluctuations of enzymatic rates of catalysis observed in single molecule experiments. Under different conditions with different switching rates, we obtained two stable states for each case by solving chemical rate equations. As the switching rate increases, the stable states shift in the direction of less A and more B. We demonstrate that the optimal transition paths converge to the average path obtained by solving mass action ODEs. At a fixed switching rate, the smaller the concentration of A is, the larger is the deviation. The reason for this is fewer A makes the system more stochastic. Furthermore, according to the large deviation theory, the transition probability is a natural exponential function of the negative minimum action. We calculated minimum actions for different cases with different switching rates. The minimum action exponentially decays with respect to switching rate. The results indicate that dynamic disorder makes the system more stable.

We then considered the genetic toggle switch model in which the rate parameter draws from a continuous heavy tailed distribution. For this case, we have modeled the system first as a discrete Markov jump process and then under continuous Langevin dynamics. Both models produce qualitatively similar results. Under the condition of fixed switching rate λ, the mean first passage time from high A to low A increases with the variance $\mathrm{Var}(k)$ in the distribution of off rates for an individual polymer A_2 from the operator. On the other hand, when we fix the variance $\mathrm{Var}(k)$, the mean first passage time from high A to low A decreases as the switching rate λ increases.

Overall, the regulation of the parameters modulating the dynamic disorder plays an important role in shaping the statistics of optimal transition paths, transition probabilities, and the stationary probability distribution of the network.

ACKNOWLEDGMENTS

We thank Arup K. Chakraborty (A.K.C.) for helpful discussions. This research was supported by NIH Grant No. P01-AI091580 to A.K.C. and collaborators and SMART grant to Jianshu Cao and collaborators.

APPENDIX: TIME CORRELATION FUNCTIONS FOR DISCRETE AND CONTINUOUS MODELS

As was shown in the main text, both the discrete and continuous models have quantitatively very similar mean first passage times. Here, we compare the two models further to understand why they behave so similarly. To compare the discrete and continuous models further, we will look here

at the time correlation function $C(k_t, k_0)$ calculated as the correlation between the kinetic parameter k_t and the time 0 value k_0. The continuous model follows an exponential of an OU process. We denote the OU process by $\tilde{k}_t$ such that $k_t = \exp(\tilde{k}_t)$. The transition density for the OU process is given by

$$\begin{aligned} f(\tilde{k}_t, |\tilde{k}_0) &= \sqrt{\frac{1}{2\pi\sigma^2(1-e^{-2\lambda t})}}\exp \\ &\quad \times \left\{-\frac{1}{2\sigma^2}\left[\frac{(x-\mu(1-e^{-\lambda t})-\tilde{k}_0 e^{-\lambda t})^2}{1-e^{-2\lambda t}}\right]\right\} \\ &= \mathcal{N}\left(\mu(1-e^{-\lambda t})+\tilde{k}_0 e^{-\lambda t}, \sigma^2(1-e^{-2\lambda t})\right) \\ &\equiv \mathcal{N}(\mu(t), \sigma(t)^2) \end{aligned} \tag{A1}$$

(we will make use of the time dependent mean and variance of the $\tilde{k}_t$ later in calculating the correlation function). The stationary density of $\tilde{k}_0$ is of course $\mathcal{N}(\mu, \sigma^2)$. Using the mean and variance of a lognormal distribution allows us to calculate the time correlation function as

$$\begin{aligned} C(k_t, k_0) &= \frac{\mathbb{E}[\exp(\tilde{k}_t)\exp(\tilde{k}_0)] - \mathbb{E}[\exp(\tilde{k}_0)]^2}{\mathrm{Var}(\exp(\tilde{k}_0))} \\ &= \frac{\mathbb{E}[\exp(\tilde{k}_t)\exp(\tilde{k}_0)] - e^{2\mu+\sigma^2}}{(e^{\sigma^2}-1)e^{2\mu+\sigma^2}}. \end{aligned} \tag{A2}$$

Note that we can actually calculate the integral $\mathbb{E}[e^{\tilde{k}_t}e^{\tilde{k}_0}]$ making use of a simple identity which can be proven easily by completing the square. For any random variable $x \sim \mathcal{N}(\mu, \sigma^2)$ and any constant λ, one can show that

$$\mathbb{E}[e^{\lambda t}] = e^{\lambda\mu + \lambda^2\sigma^2/2}. \tag{A3}$$

Breaking the integral up and carrying out the $\tilde{k}_t$ integrals first, we get

$$\begin{aligned} \mathbb{E}[e^{\tilde{k}_t}e^{\tilde{k}_0}] &= \int\left(\int e^{\tilde{k}_t}\mathcal{N}(\tilde{k}_t; \mu(t), \sigma(t)^2)\mathrm{d}\tilde{k}_t\right)e^{\tilde{k}_0}\mathcal{N}(\tilde{k}_0; \mu, \sigma^2)\mathrm{d}\tilde{k}_0 \\ &= \exp\left\{\mu(1-e^{-\lambda t}) + \frac{\sigma^2(1-e^{-2\lambda t})}{2}\right\}\int e^{\tilde{k}_0(\exp(-\lambda t)+1)}\mathcal{N}(\tilde{k}_0; \mu, \sigma^2)\mathrm{d}\tilde{k}_0 \\ &= \exp\left\{\mu(1-e^{-\lambda t}) + \frac{\sigma^2(1-e^{-2\lambda t})}{2} + \mu(e^{-\lambda t}+1) + \frac{\sigma^2(e^{-\lambda t}+1)^2}{2}\right\} \\ &= \exp\left\{2\mu + \sigma^2 + e^{-\lambda t}\sigma^2\right\}. \end{aligned} \tag{A4}$$

This leads to the simple expression for the time correlation function given by

$$C(k_t, k_0) = \frac{\exp\left\{e^{-\lambda t}\sigma^2\right\} - 1}{e^{\sigma^2} - 1}, \tag{A5}$$

which we will denote by $C_{OU}(t)$.

We now compare the above model with the discrete jump model in which each rate k_t remains at its value for a random exponentially distributed interval and then redraws from a prescribed distribution. In this case, the time correlation function is very simple to compute and does not depend on the parameters of the distribution, which we call below $\mathbb{E}[k]$, $\mathbb{E}[k^2]$,

$$\begin{aligned} \mathbb{E}[k_t k_0] &= \mathbb{E}_{k_0}[\mathbb{E}[k_t|k_0]k_0] \\ &= \mathbb{E}_{k_0}[k_0(k_0 e^{-\lambda t} + \mathbb{E}[k](1-e^{-\lambda t}))] \\ &= \mathbb{E}[k^2]e^{-\lambda t} + \mathbb{E}[k]^2(1-e^{-\lambda t}). \end{aligned} \tag{A6}$$

This demonstrates that the time correlation function of this behavior is $C(k_t, k_0) = e^{-\lambda t}$. For small σ^2, the exponentiated OU correlation function approaches

$$\frac{1 + e^{-\lambda t}\sigma^2 - 1}{1 + \sigma^2 - 1} = e^{-\lambda t} \tag{A7}$$

and the two correlation functions are approximately equal. We expect the most difference between the two processes for large σ^2, when the OU correlation function is

$$\frac{\exp\{e^{-\lambda t}\sigma^2\}}{e^{\sigma^2}} = \exp\{\sigma^2(e^{-\lambda t} - 1)\} \tag{A8}$$

and the correlations die off with a large magnitude negative exponent for the exponentiated OU process. The results are intuitive. Because the correlation function for the continuous process is a double exponential in λ, it drops off more rapidly. For small values of λ, we know that the discrete model will get frozen at a single parameter value for the duration of its dwell time, while the continuous model will still explore some of the surrounding values of its probability distribution $p(k)$. This explains some of the differences seen in the mean first passage times from the continuous process and the discrete process (Figure 7).

To explain why the stochastic processes in the kinetic off rate do not lead to significantly big differences between the MFPT behavior for both networks (even at large values of σ when the above analysis demonstrates the two models to be most different), we compute the average dwell time for a polymer following both above behaviors (by Monte Carlo sampling). For each sample, we initialize the system with a single polymer on the operator and draw the initial kinetic parameter from the lognormal with μ, σ. For the exponentiated OU, we discretize time and in each instant $\mathrm{d}t$ update the value of k_t, and also ask if unbinding occurred in the last increment, which does so with probability $1 - e^{-k_t \mathrm{d}t}$.

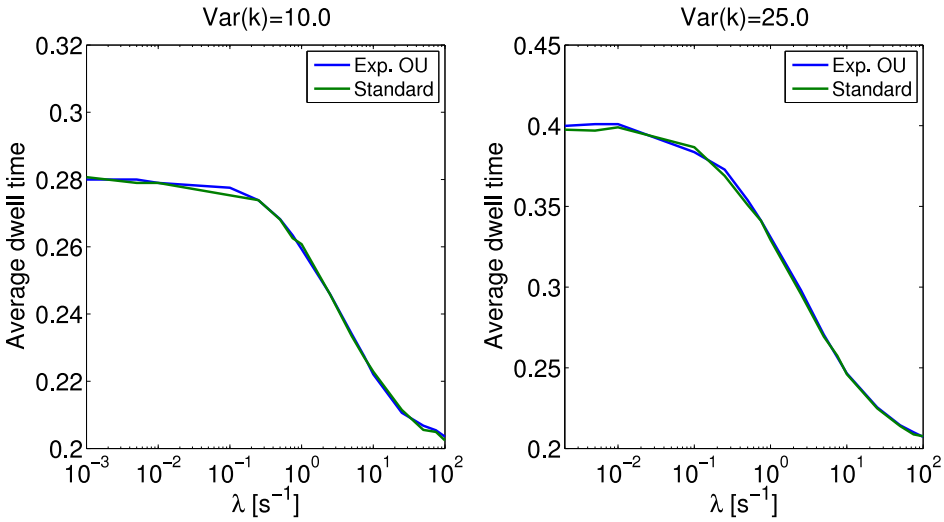

FIG. 8. We compute the expected dwell time of a polymer A on the operator for the continuous (labeled Exp. OU) and discrete model (labeled Standard) by Monte Carlo sampling. We find that the expected dwell times are comparable in both processes, which explains why the mean first passage times computed in the main text are similar for both models.

For the jump Markov process, we simply simulate the continuous time Markov process. At any instant in time t when the parameter is k_t, there are two competing exponentials: unbinding which occurs with rate k_t, and parameter switching which occurs with rate λ. We use the Gillespie method, and redraw the value of k_t when the reaction event "rate switching" occurs. We see that the expected dwell time is essentially the same for both models in the parameter regime that is shown to give bistability (Figure 8).

[1]W. E. Moerner and M. Orrit, Science **283**, 1670-1676 (1999).
[2]S. Weiss, Science **283**, 1676-1683 (1999).
[3]X. S. Xie and J. K. Trautman, Annu. Rev. Phys. Chem. **49**, 441-480 (1998).
[4]J. Cao, Chem. Phys. Lett. **327**, 38-44 (2000).
[5]S. Yang and J. Cao, J. Phys. Chem. B **105**, 6536-6549 (2001).
[6]A. M. van Oijen, P. C. Blainey, D. J. Crampton, C. C. Richardson, T. Ellenberger, and X. S. Xie, Science **301**, 1235-1238 (2003).
[7]O. Flomenbom, K. Velonia, D. Loos, S. Masuo, M. Cotlet, Y. Engelborghs, J. Hofkens, A. E. Rowan, R. J. M. Nolte, M. V. der Auweraer, F. C. de Schryver, and J. Klafter, Proc. Natl. Acad. Sci. U. S. A. **102**, 2368-2372 (2005).
[8]L. Iversen, H. L. Tu, W. C. Lin, S. Christensen, S. M. Abel, J. Iwig, H. J. Wu, J. Gureasko, C. Rhodes, R. Petit, S. Hansen, P. Thill, C. H. Yu, D. Stamou, A. K. Chakraborty, J. Kuriyan, and J. T. Groves, Science. **120**, 50-54 (2014).
[9]G. K. Schenter, H. P. Lu, and X. S. Xie, J. Phys. Chem. A **103**, 10477-10488 (1999).
[10]R. Zwanzig, Acc. Chem. Res. **23**, 148-152 (1990).
[11]M. Ptashne, *A Genetic Switch: Phage lambda and Higher Organisms*, 2nd ed. (Blackwell, Oxford, 1992).
[12]T. S. Gardner, C. R. Cantor, and J. J. Collins, Nature **403**, 339-342 (2000).
[13]A. Arkin, J. Ross, and H. H. McAdams, Genetics **149**, 1633-1648 (1998).
[14]P. B. Warren and P. Rein ten Wolde, J. Phys. Chem. B **109**, 6812-6823 (2005).
[15]D. Schultz, A. Walczak, J. Onuchic, and P. G. Wolynes, Proc. Natl. Acad. Sci. U. S. A. **105**, 19165-19170 (2008).
[16]W. E, W. Ren and E. Vanden-Eijnden, Comm. Pure Appl. Math. **57**, 637-656 (2004).
[17]M. Heymann, Ph.D. thesis, New York University, 2007.
[18]M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, 2nd ed. (Springer, 1998).
[19]H. Touchette, Phys. Rep. **478**, 1 (2009).
[20]C. C. Govern, M. Yang, and A. K. Chakraborty, Phys. Rev. Lett. **108**, 058102 (2012).
[21]D. T. Gillespie, J. Comput. Phys. **22**, 403 (1976).
[22]D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[23]D. T. Gillespie, Physica A **188**, 404-425 (1992).
[24]P. Thill, J. Chem. Phys. **141**, 015102 (2014).