PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Journal Name

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

Multiple coupled landscapes and non-adiabatic dynamics with applications to self activating genes

Cong Chen[a], Haidong Feng[b], Kun Zhang[c], Jin Wang[bac‡] and Masaki Sasai[d]

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

Many physical, chemical and biochemical systems (e.g. electronic dynamics and gene regulatory networks) are governed by the continuous stochastic processes (e.g. electron dynamics on a particular electronic energy surface, protein (gene product) synthesis) coupled with the discrete processes (e.g hopping among different electronic energy surfaces and on and off switching of genes). One can also think of the underlying dynamics as the continuous motion on a particular landscape and discrete hoppings among different landscapes. The main difference of such systems from the intra-landscape dynamics alone is the emergence of the timescale involved in transitions among different landscapes in addition to the timescale involved in a particular landscape. Adiabatic limit when inter-landscape hoppings are fast compared to continuous intra-landscape dynamics has been studied both analytically and numerically, but the analytical treatment of non-adiabatic regime where the inter-landscape hoppings are slow or comparable to continuous intra-landscape dynamics remains challenging. In this study, we show that there exists mathematical mapping of the dynamics on 2^N discretely coupled N continuous dimensional landscapes onto one single landscape in 2N dimensional extended continuous space. On this 2N dimensional landscape, eddy current emerges as a sign of non-equilibrium non-adiabatic dynamics and plays an important role in system evolution. Many interesting physical effects such as enhancement of fluctuations, irreversibility, dissipation and optimal kinetics emerge due to non-adiabaticity manifested by the eddy current illustrated for an N=1 self activator. We further generalize our theory to N-gene network with multiple binding sites and multiple synthesis rates for discretely coupled non-equilibrium stochastic physical and biological systems.

Introduction

In physical and biological systems that have frequent energy and matter exchange with environments (e.g. gene regulatory networks), stochasticity is often unavoidable with intrinsic and extrinsic fluctuations[1,2]. These stochastic processes have been modeled mathematically as drifted Brownian motions, where the system evolves in a way similar to a point particle moving diffusively along an underlying landscape. This similarity has led to a landscape picture where steady states are identified as local minimums and non-equilibrium dynamics as flows on the landscape determined by the landscape gradient and a curl flux measuring the degree of deviation from detailed balance[3–7].

A single landscape, however, is not sufficient for describing the process which is dynamically modulated by the discrete change in the system state as in gene regulatory networks[8–10] and molecular motors[11–13]. By regarding discrete change as hopping between different landscapes, these processes are described by coupled multiple landscapes, which are characterized with different timescales, timescale within each landscape and timescale of transitions among landscapes.

In order to quantify this hierarchy in timescales, we introduce an adiabaticity parameter ω. Large ω corresponds to the adiabatic limit where the inter-landscape hoping is more frequent than the intra-landscape motion. Due to the rapid transitions, a single effective landscape emerges from the 'average' of multiple landscapes. In the small ω non-adiabatic limit, the description can be simplified as the system traps in one of the discrete landscapes for a long time[7–10,14,15]. In the moderate non-adiabatic region, however, a proper physical understanding and analytical treatment is still challenging as numerical simulation quickly becomes inefficient as system size grows.

In this study, we use a N-gene regulatory network as an example to show that by developing a continuous spinor repre-

[a] Physics Department, Stony Brook University, NY 11794
[b] Chemistry Department, Stony Brook University, NY 11794
[c] State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, Jilin 130022
[d] Department of Computational Science and Engineering, Nagoya University, Nagoya 464-8603, Japan. Email: sasai@cse.nagoya-u.ac.jp
‡ Email: jin.wang.1@stonybrook.edu

sentation and the associated path integral method, the underlying 2^N discretely (e.g. each gene has 2 discrete states: on and off) coupled continuous stochastic processes can be mapped onto stochastic continuous process in a $2N$ dimensional extended space. In other words, there exists mathematical mapping of the dynamics on 2^N discretely coupled landscapes onto the one single landscape in $2N$ dimensional extended space. Our work, for the first time, provides a general analytic treatment of non-adiabatic dynamics for the discretely coupled continuous non-equilibrium stochastic systems. It gives a physical and quantitative picture to understand dynamics in non-adiabatic regime. On this $2N$ dimensional landscape, eddy current emerges as a sign of non-equilibrium non-adiabatic dynamics and plays an important role in system evolution. We also generalize our result to the general multi-gene network with each gene of having multiple binding sites or multiple activated states, so that it can be applied to variety of gene networks and discretely coupled continuous non-equilibrium stochastic physical and biological systems.

Many interesting physical effects emerge due to non-adiabaticity manifested by the eddy current in extended space. We do a detailed numerical study of $N=1$ self-activator as a special case of our general treatment. Our theory is able to explain the enhancement of fluctuations in non-adiabatic region. Most importantly by extending to 2-dimensional space, the transition between 'on' and 'off' states becomes irreversible due to flux emergence in 2-dimensional space. There is energy or heat dissipation measured by the entropy production associated with this irreversibility.[3] An optimal transition rate can appear at weak non-adiabatic region. We can see different behavior for 'on' to 'off' and 'off' to 'on' transitions due to irreversibility. The optimal rate can play an important role in gene network evolution and similar biological systems.

Model

To demonstrate our point, we study an N-gene regulatory network. Other physical and biological systems work in a similar way (e.g. N-electronic surfaces). For this specific network, each gene has one binding site on which one of the transcription factors can be bound. As a result, each gene has discrete on/off states that have different transcription expression levels[16]. The whole N-gene network has 2^N discrete states. We use s to denote a specific discrete (gene on and off) state of the network, and $P_s(n,t)$ to denote the probability of the system at state s with protein numbers $n=(n_1,n_2,\ldots n_N)$. The master equation is 2^N dimensional. Especially for $N=1$ self activator motif where the transcription factor is a dimer, the master equation is

$$\begin{aligned}
\partial_t P_1(n) &= \frac{h_0}{2}(n+2)(n+1)P_0(n+2) - fP_1(n) \\
&+ g_1(P_1(n-1)-P_1(n)) + k((n+1)P_1(n+1) - nP_1(n)) \\
\partial_t P_0(n) &= -\frac{h_0}{2}n(n-1)P_0(n) + fP_1(n-2) \qquad (1) \\
&+ g_0(P_0(n-1)-P_0(n)) + k((n+1)P_0(n+1) - nP_0(n))
\end{aligned}$$

Here g/k is the protein synthesis/degradation rate, h/f is the binding/unbinding rate. At large volume limit, protein concentration $x=\frac{n}{V_0}$ becomes continuous. After dropping out higher order terms of $\frac{1}{V_0}$ master equation becomes:

$$\begin{aligned}
\partial_t \begin{pmatrix} P_1(n) \\ P_0(n) \end{pmatrix} &= \begin{pmatrix} -\partial_x F_1 + \partial_x^2 D_1 & \\ & -\partial_x F_0 + \partial_x^2 D_0 \end{pmatrix} \begin{pmatrix} P_1(n) \\ P_0(n) \end{pmatrix} \\
&+ \begin{pmatrix} -f & h \\ f & -h \end{pmatrix} \begin{pmatrix} P_1(n) \\ P_0(n) \end{pmatrix} \qquad (2)
\end{aligned}$$

Where we define $h=\frac{h_0}{2}n^2$, redefine $k=k\cdot V_0$. The driving forces and diffusion coefficients are $F_1=\frac{1}{V_0}(g_1-kx)$, $F_0=\frac{1}{V_0}(g_0-kx)$ and $D_1=\frac{1}{2V_0^2}(g_1+kx)$, $D_0=\frac{1}{2V_0^2}(g_0+kx)$.

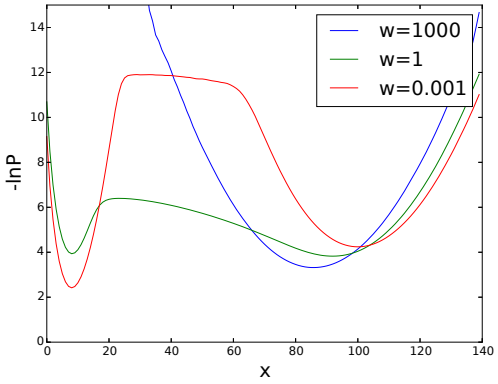


Fig. 1 One-dimensional landscape of self activator at different adiabaticity

This equation has the form of coupled Fokker-Planck equation. The first operator matrix $\mathscr{H}_0$ defines 2 discrete Fokker-Planck landscapes. The second operator matrix $\mathscr{H}_b$ describes the hopping processes that couple the discrete landscapes. As we can see from Fig. 1, at $\omega=0.001$ when the coupling is weak, steady state shows two Fokker-Planck basins. At $\omega=1000$ when the coupling is strong, the two basins merge into one. In a more general N-gene network, the dynamic master equation is 2^N dimensional. For each of the s-components which corresponds to a specific spin

state we have:

$$\begin{aligned}\frac{\partial}{\partial t}P_s(n,t) &= \sum_{i \text{ for } s_i=1} h_{ij}P_{s'}(n_1,\dots n_{i-1},n_i-2,n_{i+1},\dots n_N,t) \\ &+ \sum_{i \text{ for } s_i=0} f_{ij}P_{s''}(n_1,\dots n_{i-1},n_i+2,n_{i+1},\dots n_N,t) \\ &- \sum_{i \text{ for } s_i=0} h_{ij}P_s(n_1,\dots n_{i-1},n_i,n_{i+1},\dots n_N,t) \\ &+ \sum_{i \text{ for } s_i=1} f_{ij}P_s(n_1,\dots n_{i-1},n_i,n_{i+1},\dots n_N,t) \\ &+ \sum_{i \text{ for } s_i=1} g_1^i(P_s(n_1,\dots n_{i-1},n_i-1,n_{i+1},\dots n_N,t)-P_s(n,t)) \\ &+ \sum_{i \text{ for } s_i=0} g_0^i(P_s(n_1,\dots n_{i-1},n_i-1,n_{i+1},\dots n_N,t)-P_s(n,t)) \\ &+ k^i((n_i+1)P_s(n_1,\dots n_{i-1},n_i+1,n_{i+1},\dots n_N,t)-n_iP_s(n,t))\end{aligned} \quad (3)$$

Here s' corresponds to the spin configuration where $s'_i = 0$ while $s_i = 1$ and the rest components of s and s' are the same. s'' corresponds to the spin configuration where $s''_i = 1$ while $s_i = 0$ and the rest equal. In the case the transcription factor is dimer h_{ij} has the form of $\frac{\tilde{h}_{ij}}{2}n_j(n_j-1)$. In the large volume limit, protein concentration $x_i = \frac{n_i}{V_0}$ becomes continuous. The master equation above has the form of coupled Fokker-Planck equation:

$$\begin{aligned}\frac{\partial}{\partial_t}P_s(x,t) &= \sum_{i \text{ for } s_i=1} h_{ij}P_{s'}(x,t) + \sum_{i \text{ for } s_i=0} f_{ij}P_{s''}(x,t) \\ &- \sum_{i \text{ for } s_i=0} h_{ij}P_s(x,t) - \sum_{i \text{ for } s_i=1} f_{ij}P_s(x,t) \\ &- \frac{1}{V_0}\sum_{i \text{ for } s_i=1}\partial_{x_i}(g_1^iP_s(x,t)) + \sum_{i \text{ for } s_i=1}\frac{1}{2V_0^2}\partial^2_{x_i}(g_1^iP_s(x,t)) \\ &- \frac{1}{V_0}\sum_{i \text{ for } s_i=0}\partial_{x_i}(g_0^iP_s(x,t)) + \sum_{i \text{ for } s_i=0}\frac{1}{2V_0^2}\partial^2_{x_i}(g_0^iP_s(x,t)) \\ &+ \sum_i\frac{1}{V_0}\partial_{x_i}(k^ix_iP_s(x,t)) + \frac{1}{2V_0^2}\sum_i\partial^2_{x_i}(k^ix_iP_s(x,t))\end{aligned} \quad (4)$$

Again it has the form of coupled Fokker-Planck equation.

$$\frac{d}{dt}P_s(n,t) = \sum_{s'}\mathscr{H}_{ss'}P_{s'}(n,t) = \sum_{s'}(\mathscr{H}_{0ss'}+\mathscr{H}_{bss'})P_{s'}(n,t) \quad (5)$$

$\mathscr{H}$ in the above equation can be decomposed as $\mathscr{H} = \mathscr{H}_0 + \mathscr{H}_b$, where $\mathscr{H}_0$ is diagonal and describes protein synthesis/degradation, and $\mathscr{H}_b$ represents the binding/unbinding processes with non-diagonal elements describing hopping between discrete on and off gene states. $\mathscr{H}_0$ has terms of g_0^i, g_1^i and k_i, where g_0^i and g_1^i are the protein synthesis rates of i-protein when i-gene is at unbound and bound states, respectively. k_i is the degradation rate of i-protein. In $\mathscr{H}_b$, we use h_{ij} and f_{ij}, where h_{ij} is the binding rate of j-protein with i-gene, f_{ij} is the corresponding unbinding rate. In the case where the transcription factor is dimer, $h_{ij} = \frac{1}{2}\tilde{h}_{ij}n_j(n_j-1)$. f_{ij} eaquals $k \times \omega$ is constant. As mentioned, the key character of such system is the hierarchy of two timescales, one for protein synthesis/degradation and the other for gene activation/repression. To quantify this timescale hierarchy, we introduce adiabaticity parameter as $\omega = f/k$. Other parameters we use to set up the networks are $X_{eq} = f/h$ (equilibrium constant of binding), $X_{ad} = (g_0+g_1)/2k$ (synthesis relative to degradation) and $\delta X = (g_1-g_0)/2k$ (difference in on and off synthesis). We can make discrete protein copy numbers n_i continuous by considering large volume limit, and replace n_i with concentration $x_i = n_i/V_0$. In this way, we always work in the continuous x-space.

In the self activator example we considered, $\mathscr{H}_0$ defines 2 discrete Fokker-Planck states and $\mathscr{H}_b$ describies the coupling between these states. Adiabaticity ω describes the relative strength of $\mathscr{H}_b$ and $\mathscr{H}_0$. When ω is large, we arrive at the adiabatic limit. From the master equation, $\mathscr{H}_b$ dominates over $\mathscr{H}_0$, so that we are always close to the steady state $\mathbf{r}_0$ of $\mathscr{H}_b$: $\mathscr{H}_b\mathbf{r}_0 = 0$. For self-activator, $\mathbf{r}_0 = (\frac{h}{\sqrt{h^2+f^2}}, \frac{f}{\sqrt{h^2+f^2}})^T$. The master equation becomes $\partial_t P(n,t) = 1^T H_0(P(n,t)\mathbf{r}_0)$, which has the form of one-dimensional FPE[9,17]. The other limit, when ω is small, corresponds to the weakly coupling case where hopping between different states is rare. $\mathscr{H}_0$ dominates over $\mathscr{H}_b$. Because $\mathscr{H}_0$ is 2 dimensional and diagonal, if we prepare the initial state at one of the 2 discrete states, the system tends to be trapped there for a long time. This is exactly what we see from Fig. 1.

The moderate ω non-adiabatic regime is of particular interest. As has been pointed out, dynamics in this region can be crucial to physical and biological system dynamics[7,10,15]. However, in this regime, terms in $\mathscr{H}_0$ and $\mathscr{H}_b$ are comparable, the adiabatic approximation no longer works. For general N-gene network, although the master equation is exact, its size grows as 2^N, numerical simulation quickly becomes non-efficient. A clear physical and analytical picture is needed to find a way around this.

Path Integral and Effective Lagrangian

The master equation has the form of Schrödinger equation. The $2^N \times 1$ state vector $\mathbf{P}$ plays the role of wavefunction. $\mathscr{H} = \mathscr{H}_0 + \mathscr{H}_b$ becomes the hamiltonian operator when acting on $\mathbf{P}$. Though $\mathscr{H}$ is non-hermitian, reflecting the non-equilibrium features of the system, the change of gene on/off states is similar to the change of spin-up/down states of electron[18]. To quantify the 'spin state' of the ith gene, we use spinor $|s_i\rangle$ parameterised by θ_i and ϕ_i.

$$|s_i\rangle = \begin{pmatrix}\cos^2\frac{\theta_i}{2}e^{i\phi_i/2} \\ \sin^2\frac{\theta_i}{2}e^{-i\phi_i/2}\end{pmatrix} \qquad \langle s_i| = \begin{pmatrix}e^{-i\phi_i/2}, & e^{i\phi_i/2}\end{pmatrix} \quad (6)$$

With spin up representing the bound state and spin down the unbound state. This spinor is properly normalized: $\langle s^i|s^i\rangle = 1$, the identity operator in spin space can be written as:

$$I_s = \prod_i \frac{1}{4\pi}\int d\cos\theta_i d\phi_i\,|s_i\rangle\langle s_i| \quad (7)$$

The amplitude in the two components of $|s_i\rangle$ has the physical meaning of the probability of the corresponding binding site being at bound/unbound state. A general state of the system $|x,s,t\rangle$can be described with protein copy number $x = (x_1,x_2,\ldots x_N)$ and 'spin' state $s = s_1 \otimes s_2 \ldots s_N$.

Path integral method tells us that the transition probability from an initial point with protein concentration $x_i = (x_i^1,x_i^2,\ldots x_i^N)$(with $1,2,\ldots,N$ labeling genes or proteins) and spin states s_i at time t_i to the final point with (x_f,s_f,t_f) can be calculated by sum over all paths connecting these two points with exponential weight proportional to the integral of effective lagrangian along the path[19]. That is:

$$
\begin{aligned}
P(x_f,s_f,t_f|x_i,s_i,t_i) &= \langle x_f,s_f|\, e^{\int_{t_i}^{t_f}\mathscr{H}d\tau}\,|x_i,s_i\rangle \\
&= \langle x_f,s_f|\left[\lim_{\Delta t\to 0}\prod_{\Delta t} I_x\otimes I_s(1+\mathscr{H}\Delta t)I_x\otimes I_s\right]|x_i,s_i\rangle \\
&= \langle x_f,s_f|\lim_{\Delta t\to 0}\left(\prod_{\tau=t_i+\Delta t}^{\tau=t_f}|p,s,\tau\rangle\langle p,s,\tau|\,(1+\mathscr{H}\Delta t)\,|p,s,\tau-\Delta t\rangle\langle p,s,\tau-\Delta t|\right)|x_i,s_i\rangle \\
&= \text{const}\prod_i \mathscr{D}x\mathscr{D}p\mathscr{D}\cos\theta_i\mathscr{D}\phi_i e^{-\mathscr{L}dt}
\end{aligned}
\tag{8}
$$

Here, $\mathscr{L}$ is the effective lagrangian, which contains $2N$ coordinate q-like variables x_i and c_i with c_i being the probability of the $s_i = 0$ state, and $2N$ momentum p-like variables p_i and ϕ_i.

To simplify our notation we will note $\sin^2(\theta_i/2)$ as c_i, thus $\cos^2(\theta_i/2) = 1 - c_i$ The matrix element $\langle p,s,\tau|\,(1+\mathscr{H}\Delta t)\,|p,s,\tau-\Delta t\rangle$ receives contribution from 3 parts:

$$
\langle p,s,\tau|\,1\,|p,s,\tau-\Delta t\rangle = 1 + i\sum_i p_i\dot{x}_i\Delta t - i\sum_i \phi_i\dot{c}_i\Delta t \tag{9}
$$

$$
\begin{aligned}
\langle p,s,\tau|\,\mathscr{H}_0\,|p,s,\tau-\Delta t\rangle = \quad & \sum_i \frac{1}{V_0}(-ip_i)(g_0^i - k_i x_i)c_i + \sum_i \frac{1}{V_0}(-ip_i)(g_1^i - k_i x_i)(1-c_i) \\
& -\sum_i \frac{1}{2V_0^2}p_i^2(g_0^i + k_i x_i)c_i - \sum_i \frac{1}{2V_0^2}p_i^2(g_1^i + k_i x_i)(1-c_i)
\end{aligned}
\tag{10}
$$

$$
\langle p,s,\tau|\,\mathscr{H}_i\,|p,s,\tau-\Delta t\rangle = \sum_i\sum_j\left(-h_{ij}c_i + h_{ij}c_i e^{-i\phi_i} + f_{ij}(1-c_i)e^{i\phi_i} - f_{ij}(1-c_i)\right) \tag{11}
$$

The effective lagrangian $\mathscr{L}$, turns out to be:

$$
\begin{aligned}
-\mathscr{L} \quad = & \sum_i ip_i\frac{d}{dt}x_i - \sum_i i\phi_i\frac{d}{dt}c_i \\
& \sum_i \frac{1}{V_0}(-ip_i)(g_0^i - k_i x_i)c_i + \sum_i \frac{1}{V_0}(-ip_i)(g_1^i - k_i x_i)(1-c_i) \\
& \sum_i \frac{1}{2V_0^2}p_i^2(g_0^i + k_i x_i)c_i - \sum_i \frac{1}{2V_0^2}p_i^2(g_1^i + k_i x_i)(1-c_i) \\
& \sum_i\sum_j\left(-h_{ij}c_i + h_{ij}c_i e^{-i\phi_i} + f_{ij}(1-c_i)e^{i\phi_i} - f_{ij}(1-c_i)\right)
\end{aligned}
\tag{12}
$$

Effective langrangian is equivalent to Hamiltonian in providing dynamics information. Under spinor representation it is possible to map the original dynamics in continuous x space with discrete gene states into continuous dynamics in x space and spin space.

Deterministic Dynamics and Intrinsic Noise

Effective lagrangian provides information of dynamics in continuous x space as well as in spin space. To reduce the dimensionality and concentrate on the dynamics in the observable space, we can integrate out the conjugate variables. We do so by expanding over conjugate variables p_i and ϕ_i. The first order terms correspond to

classic dynamics, or deterministic dynamics:

$$-\mathscr{L}_{cl} = \sum_i i p_i \frac{d}{dt}x_i - \sum_i i\phi_i \frac{d}{dt}c_i + \sum_i \frac{1}{V_0}(-ip_i)(g_0^i - k_i x_i)c_i + \sum_i \frac{1}{V_0}(-ip_i)(g_1^i - k_i x_i)(1-c_i) + \sum_i\sum_j i\phi_i\left(f_{ij}(1-c_i) - h_{ij}c_i\right) \quad (13)$$

When integrating over ϕ_i and p_i, $\mathscr{L}_{cl}$ contributes $2N$ delta functions, providing $2N$ deterministic equations for x_i and c_i.

$$\dot{x}_i = \frac{1}{V_0}\left(g_0^i c_i + g_1^i(1-c_i) - k_i x_i\right) \quad \dot{c}_i = \sum_j \left(f_{ij}(1-c_i) - h_{ij}c_i\right) \quad (14)$$

We notice that though we didn't use self-consistent assumption, at deterministic level we obtain the same result as the self-consistent approach with zero-fluctuation hamiltonian equations[20]: $\dot{x} = \frac{\partial \mathscr{H}}{\partial p}|_{p=0}$ and $\dot{c} = \frac{\partial \mathscr{H}}{\partial \phi}|_{\phi=0}$.

As we go to the next leading order, we obtain quadratic terms of conjugate variables p_i and ϕ_i. This corresponds to semi-classical dynamics.

$$-\mathscr{L}_{sc} = -\sum_i \frac{1}{2V_0^2}p_i^2(g_0^i + k_i x_i)c_i - \sum_i \frac{1}{2V_0^2}p_i^2(g_1^i + k_i x_i)(1-c_i) - \sum_i \frac{1}{2}\phi_i^2\left(\sum_j\left(f_{ij}(1-c_i) + h_{ij}c_i\right)\right) \quad (15)$$

After the Hubbard-Stratonovich transformation, quadratic terms provide gaussian intrinsic fluctuations[21]. This suggests the effective dynamics is governed from the original $2^N \times N$ dynamics to now by a $2N$ dimensional coupled Langevin system with coordinate dependent diffusion coefficients:

$$\dot{x}_i = \frac{1}{V_0}(g_0^i c_i + g_1^i(1-c_i) - k_i x_i) + \eta_{xi}, \quad (16)$$

$$\dot{c}_i = \sum_{j=1}^N f_{ij}(1-c_i) - \sum_{j=1}^N h_{ij}c_i + \eta_{ci}, \quad (17)$$

with η_{xi} and η_{ci} being intrinsic gaussian noises;

$$\langle \eta_{x_i}(t)\eta_{x_j}(t')\rangle = \frac{1}{V_0^2}\left(g_0^i c_i + g_1^i(1-c_i) + k_i x_i\right)\delta_{ij}\delta(t-t'),$$

$$\langle \eta_{c_i}(t)\eta_{c_j}(t')\rangle = \left(\sum_j\left(f_{ij}(1-c_i) + h_{ij}c_i\right)\right)\delta_{ij}\delta(t-t').$$

If we view the problem in terms of probability distribution, the coupled Langevin equations are equivalent to a $2N$ dimensional FPE governed system:

$$\partial_t P(n,s,t) = -\sum_{x_i}\partial_{x_i}(F_{x_i}P) - \sum_{c_i}\partial_{c_i}(F_{c_i}P) + \sum_{x_i}\partial_{x_i}^2(D_{x_i x_i}P) + \sum_{c_i}\partial_{c_i}^2(D_{c_i c_i}P). \quad (18)$$

Mapping the system stochastic dynamics to the probability evolution via FPE provides a landscape picture for dynamics especially in the non-adiabatic region. Define potential U from the steady state probability distribution ρ^{SS} as $U = -\log\rho^{SS}$. The stable steady states correspond to local minimums of U. The stability of the steady state is determined by the geometry or depth of the basin. Probability conservation tells us the steady state flux is: $j_i^{SS} = -F_i\rho^{ss} + \partial_j(D_{ij}\rho^{ss})$ and $\nabla\cdot\mathbf{j} = 0$. The steady state flux is divergent free and therefore a rotational curl. If we define $\tilde{F}_i = F_i - (\partial_j D_{ij})$, the steady state flux can be written in a way similar to what we have in constant diffusion coefficient case: $\mathbf{j} = -\mathbf{F}\rho^{SS} + \partial\cdot(\mathbf{D}\rho^{SS}) = -\tilde{\mathbf{F}}\rho^{SS} + \mathbf{D}\cdot\partial\rho^{SS}$. The driving force can be decomposed into a gradient part plus a curl part: $F_i = -D_{ij}\partial_j U + j_i^{ss}/\rho^{ss} + \partial_j D_{ij}$ or $\tilde{F}_i = -D_{ij}\partial_j U + j_i^{ss}/\rho^{ss}$.[3]. The system undergoes a drifted Brownian motion on this $2N$-dimensional extended space landscape. If there's only conservative gradient force, the system would be equilibrated at a steady state and there is no flux. The non-zero steady state flux, or eddy current gives a quantitative measure of the detailed balance breaking. As we sill see in the self activator example, it plays an important role in entropy production and non reversible system dynamics.

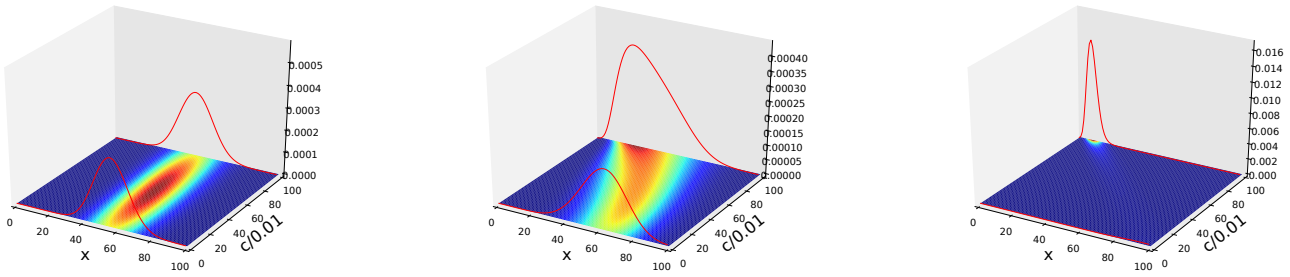
Generalization to N-gene network with multiple binding sites

The example we considered with each gene has only one binding site is rather simplified. If we consider each gene can have up to M binding sites, then each gene has 2^M discrete spin states. For instance, when $M = 2$, the i-gene state has $2^M = 4$ discrete states: $(P_{s_i=00}(n,t), P_{s_i=10}(n,t), P_{s_i=01}(n,t), P_{s_i=11}(n,t))^T$. These 2^M discrete states can have different protein synthesis rate; instead of g_0^i and g_1^i now we have 2^M different $g_{s_i}^i$. We introduce an extra index $1 \le \mu \le M$ to distinguish different binding sites on the same gene. The "spin" state of ith gene can be represented as $|s_i\rangle = \prod_{\mu=1}^M |s_i^\mu\rangle$. The identity operator in spin space is:

$$I_s = \prod_{i=1}^N\left(\frac{1}{(4\pi)^M}\prod_{\mu=1}^M\left(\int_0^1 dc_i^\mu\int_{-2\pi}^{2\pi}d\phi_i^\mu\right)|s_i\rangle\langle s_i|\right)$$

The probability of propagation can still be calculated using path integral when we pay special attention to the index μ. Follow a similar routine we get effective Lagrangian:

$$-\mathscr{L} = \sum_i i p_i\dot{x}_i - \sum_i i\phi_i^\mu\dot{c}_i^\mu + \sum_i\frac{1}{V_0}\left[\sum_{s_i}g_{s_i}^i\left(\prod_{s_i^\mu=1}(1-c_i^\mu)\prod_{s_i^\nu=0}c_i^\nu\right) - k_i x_i\right] - \frac{1}{2V_0^2}\sum_i p_i^2\left[\sum_{s_i}g_{s_i}^i\left(\prod_{s_i^\mu=1}(1-c_i^\mu)\prod_{s_i^\nu=0}c_i^\nu\right) + k_i x_i\right] + \sum_{i,j}\sum_\mu[-h_{ij}^\mu c_i^\mu + h_{ij}^\mu c_i^\mu e^{-i\phi_i^\mu} + f_{ij}^\mu(1-c_i^\mu)e^{i\phi_i^\mu} - f_{ij}(1-c_i)]$$



(a) $\omega = 1000$ **(b)** $\omega = 1$ **(c)** $\omega = 0.01$

Fig. 2 Steady state distribution of self-activator in extended space at (a): $\omega = 1000$, (b):$\omega = 1$, (c):$\omega = 0.01$

When expand the effective lagrangian over p_i and ϕ_i^μ up to the second order, we see the coupled FPE system is equivalent to $N(M+1)$ dimensional coupled Langevin equations:

$$\begin{aligned} \dot{x}_i &= \frac{1}{V_0}\left[\sum_{s_i} g^i_{s_i}\left(\prod_{s_i^\mu=1}(1-c_i^\mu)\prod_{s_i^\nu=0} c_i^\nu\right) - k_i x_i\right] + \eta_{x_i}, \\ \dot{c}_i^\mu &= \sum_j f_{ij}^\mu (1-c_i^\mu) - \sum_j h_{ij}^\mu c_i^\mu + \eta_{c_i^\mu} \end{aligned} \tag{19}$$

$$\langle \eta_{x_i}(t)\eta_{x_j}(t')\rangle = \frac{1}{V_0^2}\left[\sum_{s_i} g^i_{s_i}\left(\prod_{s_i^\mu=1}(1-c_i^\mu)\prod_{s_i^\nu=0} c_i^\nu\right) - k_i x_i\right]\delta_{ij}\delta(t-t')$$

$$\langle \eta_{c_i^\mu}(t)\eta_{c_j^\nu}(t')\rangle = \left(\sum_j \left(f_{ij}^\mu(1-c_i^\mu) + h_{ij}^\mu c_i^\mu\right)\right)\delta_{\mu\nu}\delta(t-t') \tag{20}$$

Generalization to gene with multiple transcription level

However, if the genes are regulated also by mechanism such as chromatin structure change[10,15,22,23] or histone modification[24–26], it should be reasonable to assume that a single gene has more than two states. Then, it is proper to introduce a multiple component spinor:

$$|s_i\rangle = \begin{pmatrix} c_i^1 e^{i\phi_i^1/2} \\ \vdots \\ c_i^M e^{i\phi_i^M/2} \end{pmatrix} \qquad \langle s_i| = \left(e^{-i\phi_i^1/2}\ldots, e^{-i\phi_i^M/2}\right) \tag{21}$$

The normalization condition $\sum_\mu c_i^\mu = 1$ is preserved by the form of $\mathscr{H}_i$. In the example of $N=1$ network where there is one x variable and $(M-1)$ independent c^μ variables. Assume we can still write down the master equation in the coupled FPE form:

$$\partial_t P_s(n,t) = \sum_{s'}(\mathscr{H}_0 + \mathscr{H}_i)_{ss'} P_{s'}(n,t) \tag{22}$$

where $\mathscr{H}_0$ is diagonal in s space and $\mathscr{H}_b$ describes the jumping processes. The identity operator in s space would look like:

$$I_s = \prod_{\mu=1}^{M}\left(\frac{1}{4\pi}\int_0^1 dc^\mu \int_{-2\pi}^{2\pi} d\phi^\mu \,|s\rangle\langle s|\right) \tag{23}$$

Start with coupled FPE that has Schrodinger Equation like form and follow path integral formula we get the transition probability from initial state $|x_i,s_i,t_i\rangle$ to final state $|x_f,s_f,t_f\rangle$ to be:

$$\begin{aligned} P(x_f,s_f,t_f|x_i,s_i,t_i) &= \langle x_f,s_f|\, e^{\int \mathscr{H} dt}\,|x_i,s_i\rangle \\ &= \text{const}\prod_{i,\mu}\int \mathscr{D}x_i\mathscr{D}p_i\mathscr{D}c_i^\mu\mathscr{D}\phi_i^\mu e^{-\int \mathscr{L} dt} \end{aligned} \tag{24}$$

Plug in the identity operator in x space and new one in s spcae. Similarly we have:

$$\langle p,s,t|1|p,s,t-\Delta t\rangle = 1 + i\sum_i p_i\dot{x}_i\Delta t + i\frac{1}{2}\sum_{i,\mu}\phi_i^\mu\dot{c}_i^\mu\Delta t \tag{25}$$

$$\langle p,s,t|\mathscr{H}_0|p,x,t-\Delta t\rangle = \sum_s{}'\left[(\mathscr{H}_0)_{ss}\prod_i c_i(s)\right] \tag{26}$$

$$\langle p,s,t|\mathscr{H}_i|p,x,t-\Delta t\rangle = \sum_{s,s'}{}''\left[(\mathscr{H}_i)_{ss'}\prod_i c_i(s')e^{i(\phi^\nu-\phi^\mu)/2}\right] \tag{27}$$

effective Lagrangian turns out to be:

$$\begin{aligned} -\mathscr{L} = i\sum_i p_i\dot{x}_i + \frac{i}{2}\sum_{i,\mu}\phi_i^\mu\dot{c}_i^\mu + \sum_s{}'\left[(\mathscr{H}_0)_{ss}\prod_i c_i(s)\right] \\ + \sum_{s,s'}{}''\left[(\mathscr{H}_i)_{ss'}\prod_i\left(c_i(s')e^{i(\phi^\nu-\phi^\mu)/2}\right)\right] \end{aligned} \tag{28}$$

Here $\sum_s'$ is a sum over all possible s with s_i at the μ state. $\sum_{s,s'}''$ is a sum over all possible s and s' with s_i at the μ state and s_i' at the ν state. As in the $L=2$ case, $\mathscr{H}_0$ alone defines the L^N N-dimensional x space FP landscapes with proper driving force and intrinsic fluctuation. $\mathscr{H}_b$ is non-diagonal . The adiabatic parameter ω can be defined as $\omega = \frac{|\text{typical element of } \mathscr{H}_b|}{|\text{typical element of } \mathscr{H}_0|}$ to reflect the time scale hierarchy of the two sets of processes. We notice that out of the L c^μ elements only $L-1$ are independent due to the

probability conservation.

The first order terms, corresponding to deterministic classical dynamics, are:

$$-\mathscr{L}_{cl} = i\sum_i p_i\dot{x}_i + \frac{i}{2}\sum_{i,\mu}\phi_i^\mu \dot{c_i^\mu} + \sum_s{}'\sum_i p_i \frac{\partial(\mathscr{H}_0)_{ss}}{\partial p_i}\prod_j c_j(s) + \sum_{\substack{s,s'\\ \mu\neq\nu}}\left[(\mathscr{H}_i)_{ss'}\frac{i(\phi^\nu-\phi^\mu)}{2}\prod_i c_i(s')\right] \quad (29)$$

After integration over p_i and ϕ_i^μ, this contributes delta functions which provide deterministic equations:

$$\dot{x}_i = i\sum_s{}'\left[\left(\frac{\partial(\mathscr{H}_0)_{ss}}{\partial p_i}\right)\Big|_{p=0}\cdot\prod_j c_j(s)\right] \quad (30)$$

$$\dot{c_i^\mu} = \sum_{\substack{s,s'\\ s_i'=\nu\neq\mu=s_i}}\left[(\mathscr{H}_i)_{s's}\prod_j c_j(s) - (\mathscr{H}_i)_{ss'}\prod_j c_j(s')\right] \quad (31)$$

Especially the normalization is conserved at classical level:

$$\sum_\mu \dot{c_i^\mu} = \sum_{\substack{s,s'\\ s_i\neq s_i'}}\left[(\mathscr{H}_i)_{s's}\prod_i c_i(s) - (\mathscr{H}_i)_{ss}\prod_i c_i(s')\right] = 0 \quad (32)$$

The quadratic terms provide information of intrinsic fluctuation. We are extremely interested in c_i^μ part. Notice that cross terms like $\phi^\mu\phi^\nu$ don't contribute after integration. We are left with quadratic terms of ϕ_i^μ:

$$\mathscr{L}_{sc} = \frac{1}{2}\sum_s\sum_i p_i^2\left(\frac{\partial^2(\mathscr{H}_0)_{ss}}{\partial p_i^2}\right) - \frac{1}{8}\sum_{\substack{s,s'\\ s_i=\mu\neq\nu=s_i'}}(\mathscr{H}_i)_{ss'}(\phi^{\mu 2}+\phi^{\nu 2})\prod_i c_i(s')(\phi^{\mu 2}+\phi^{\nu 2}) \quad (33)$$

Just like in $M=2$ spinor case we have worked out, the quadratic terms, after HubbardStratonovich transformation, provide non-constant gaussian fluctuation:

$$\dot{x}_i = i\sum_s{}'\left[\left(\frac{\partial(\mathscr{H}_0)_{ss}}{\partial p_i}\right)\Big|_{p=0}\cdot\prod_j c_j(s)\right] + \eta_{x_I} \quad (34)$$

$$\dot{c_i^\mu} = \sum_{\substack{s,s'\\ s_i'=\nu\neq\mu=s_i}}\left[(\mathscr{H}_j)_{s's}\prod_i c_j(s) - (\mathscr{H}_i)_{ss'}\prod_j c_j(s')\right] + \eta_{c_i^\mu} \quad (35)$$

with diffusion coefficients:

$$\langle\eta_{x_i}(t)\eta_{x_j}(t')\rangle = -\sum_s\left[\frac{\partial^2}{\partial p_i^2}(\mathscr{H}_0)_{ss}\left(\prod_j c_j(s)\right)\right]\delta_{ij}\delta(t-t') \quad (36)$$

$$\langle\eta_{c_i^\mu}(t)\eta_{c_i^\xi}(t')\rangle = \sum_{\substack{s,s'\\ s_i=\mu\neq\nu=s_i'}}\left[(\mathscr{H}_j)_{s's}\prod_i c_j(s) + (\mathscr{H}_i)_{ss'}\prod_j c_j(s')\right]\delta_{\mu\xi}\delta(t-t') \quad (37)$$

Because only $L-1$ of the L c_i^μ variables are independent, those are $N\times L$ dimensional coupled Langevin equations. From the view of probability distribution, this is equivalent to a $N\times L$ dimensional FP landscape system with N-dimensional x space and $N\times(L-1)$ c-space. The evolution in the continuous spinor coordinate c-space plays an important role in the network evolution, especially in the non-adiabatic region. At adiabatic limit, because the dynamics of c variables are totally determined by x-variables, we go back to the N-dimensional x-space FP landscape picture. When $L=2$, we get back to the on/off N-gene network in the last section.

Application to self activator

We will apply our theory to a self activating motif with transcription factor being dimer. This is one of the basic building bricks of larger networks.The master equation, in large volume limit, has the form of coupled Fokker-Planck equation as in Eq.(2). Introducing two components spinor to quantify the binding site, the equivalent two dimensional Fokker-Planck equation in extended space is:

$$\frac{\partial}{\partial t}P = -\partial_x(F_xP) - \partial_c(F_cP) + \partial_x^2(D_xP) + \partial_x^2(D_cP) \quad (38)$$

with driving force and diffusion:

$$F_x = \frac{1}{V_0}(cg_0 + (1-c)g_1 - kx)$$

$$F_c = f(1-c) - \frac{1}{2}h_0x^2c$$

$$D_x = \frac{1}{2V_0^2}[cg_0 + (1-c)g_1 + kx]$$

$$D_c = \frac{1}{2}[f(1-c) + \frac{1}{2}h_0x^2c] \quad (39)$$

As Fig. 1 suggests, when ω is small, the coupling between on/off state is weak. The landscape shows two separate basins. At adiabatic limit($\omega = 1000$), the two basins merge into one as

adiabatic approximation suggests. As Fig. 2 shows, in the new formalism of Fokker-Planck system in extended space, the two separate basins at small ω show up at $c = 0$ 'on' state and $c = 1$ 'off' state. As ω increases, the two basins move against each other and at adiabatic limit coincide at almost the same location, correspond to single peak at adiabatic limit.

The adiabatic approximation that maps the system to a 1-dimensional Fokker-Planck system, though numerically a good approximation, lacks the ability to demonstrate non-equilibriumness of the system as there is no curl flux in one-dimensional systems. One-dimensional system is governed by equilibrium gradient driving force and always leads to detailed balance. In our new formalism, as showed in Fig. 3 that the non-trivial flux (white arrows) originated from the spin coupling emerges in the expanded x-c space as a sign of system being at non-equilibrium state.

The Fano Factor of steady states, defined as variance/mean, quantifies the global stability. For a single Poisson peak, Fano Factor equals one. Large fano factor indicates large deviation from single Poisson peak and large fluctuation. Fig. 4(a) shows how Fano factor changes with adiabaticity. In the adiabatic limit, as adiabatic approximation suggests, due to frequent binding/unbinding processes, system is always close to the steady state of $\mathbf{H}_b$. It is a weighted average of 'on' and 'off' peak and is itself Poisson like. The fano factor is close to one. In the moderate non-adiabatic region, Fano factor increase as adiabaticity decreases. It is because when the coupling between 'on' and 'off' states is weak, the adiabatic peak splits into 'on' and 'off' peaks that contribute to Fano factors.

While for 1-dimensional landscape of self activator, the driving force is purely gradient, the transitions from both off to on and on to off are reversible. When extended to 2-dimensional landscape the transition paths are clearly non-reversible with emergence of eddy-current in addition to gradient component. The 2-dimensional landscape makes it possible to study irreversible optimal transitions between off and on states. We focus on the mean first passage time (MFPT) for both on to off and off to on states. From Fig. 4(b) we see a kinetic turnover behavior where an optimal time or transition rate exists with respect to adiabaticity. From mean first passage time, we see that in the adiabatic regime, the discrete landscapes are strongly coupled. The rate limiting step is determined by the adiabatic barrier between the two states on a single effective landscape (intra-landscape dynamics). Decreasing the ω decreases the coupling between discrete landscapes. The 'averaged' single basin at adiabatic limit splits into 'on' and 'off' states as ω decreases. The effective barrier becomes less than the one in the adiabatic case and the transition rate become higher. On the other hand, in the small ω regime, the rate limiting step is determined by the non-adiabatic jumping between the landscapes. The discrete landscapes are weakly coupled, system tends to stay in either 'on' or 'off' state for a long time. Decreasing the ω further decreases the frequency or probability of jumping and therefore the transition rate decreases. This explains the turnover behavior of rate or kinetic time with respect to adiabaticity ω.[14,27]

Difference in forward and backward three-point correlation functions, as Fig. 4(c) shows, indicates the system evolution is irreversible in time due to the presence of the flux that breaks detailed balance. As ω decreases (non-adiabaticity increases), the irreversibility becomes more pronounced. We should note that these non-equilibrium features become more evident as ω decreases from the adiabatic regime to the non-adiabatic regime.

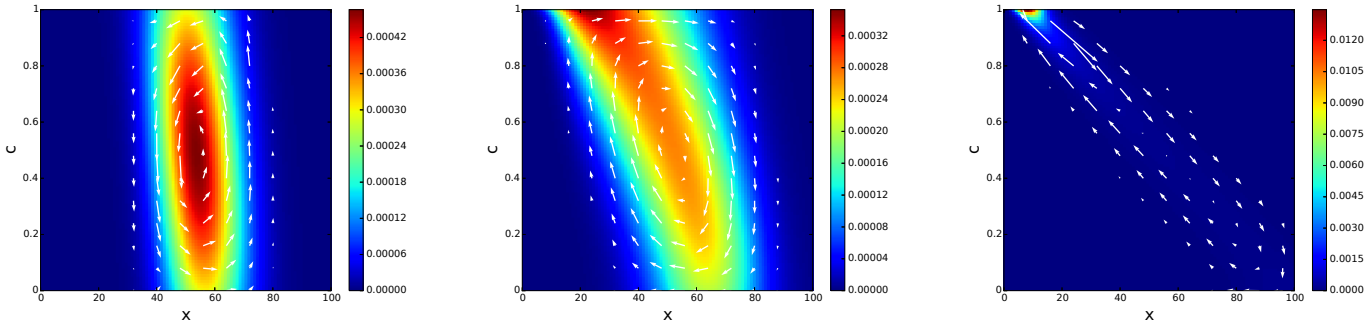
The entropy production rate representing the dissipation directly associated with the degree of deviation from equilibrium, is closely related to the new flux in extended space. Fig. 4(d) shows EPR decreases as ω increases. The non-zero flux promotes the non-adiabatic fluctuations and dissipation in the smaller ω regime. Three-point correlation and entropy production rate both show that non-adiabatic regime where the discrete landscapes are weakly coupled shows higher deviation from the equilibrium reflected by the higher heat dissipation and higher irreversibility. Our theory provides a unified landscape theory in this regime with 2-dimensional eddy current that measures the degree of the system being deviated from equilibrium.

Above all, our theory is capable to explain the non-equilibrium and irreversible properties of network dynamics with eddy current emerging in extended space, and provides a theoretical framework to study dynamics (transition rate, path) within a single unified landscape for both adiabatic and non-adiabatic dynamics.

Conclusions

In this work, we systematically studied discretely coupled stochastic processes and associated non-adiabatic dynamics that are common in physics and biology. We started with exact master equations, and showed by using the continuous spinor representation we are able to map the dynamics in 2^N discretely coupled landscapes into a single landscape in the extended space. Our result, at deterministic level, agrees with self-consistent approximation. It also unveils the role intrinsic fluctuation played in system evolution. Intrinsic fluctuation lies not only in synthesis and degradation processes, but also in the binding/unbinding processes. In other words, for a coupled Fokker-Planck system (e.g. N-gene network), we have intrinsic fluctuation not only within each of the discrete Fokker-Planck landscapes, but also in the hopping processes. This is clearly demonstrated in the unified landscape formalism with fluctuation in both protein concentration x-variables and discrete state specifying c-variables.

The unified landscape picture provides a theoretical and analytical framework for non-equilibrium dynamics at non-adiabatic region. The mapping from 2^N discretely coupled N-dimensional stochastic system into a $2N$-dimensional Fokker-Planck system significantly improves numerical calculation efficiency in non-adiabatic region where adiabatic approximation is not valid. As we can see from the self activator example, it captures thermal dynamical properties like fluctuation enhancement, irreversibility, transition rate turn over. The steady state flux introduced in the extended space is the key signature of the non-equilibrium dynamics. It gives a quantitative measure of the detailed balance breaking. It also enables us to study optimal transition path and rate under familiar Fokker-Planck framework.



(a) $\omega = 1000$ (b) $\omega = 1$ (c) $\omega = 0.01$

Fig. 3 Steady state distribution of self-activator in extended space with white arrow being non-gradient flux at (a): $\omega = 1000$, (b):$\omega = 1$, (c):$\omega = 0.01$

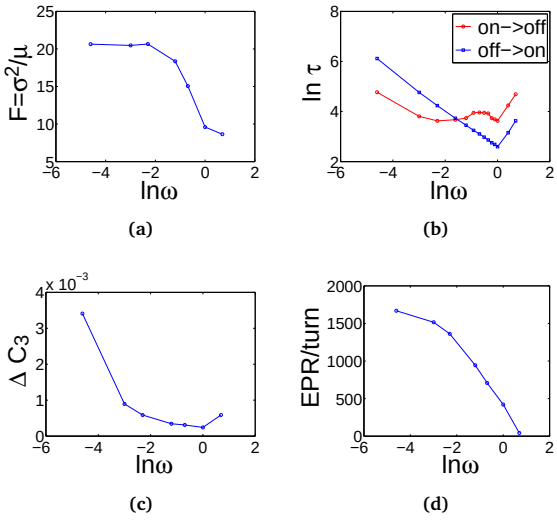


Fig. 4 Thermodynamic quantities of self activator calculated in extended space under unified landscape framework (a):Fano factor. (b):Mean first passage time(MFPT). (c):Difference between forward and backward three-point correlation functions. (d):Entropy production rate per turn of the gene on-off switch.

acknowledgements

The research studies presented in this study were supported in part by National Science Foundation (USA: NSF-0926287, NSF-0947767) and National Natural Science Foundation of China (NSFC-21190040, NSFC-11174105, NSFC-91225114, NSFC-91430217). C.C. would like to thank Dr. Chunhe Li for fruitful discussions.

References

1 M. Thattai and A. van Oudenaarden, *Proceedings of the National Academy of Sciences*, 2001, **98**, 8614–8619.

2 P. S. Swain, M. B. Elowitz and E. D. Siggia, *Proceedings of the National Academy of Sciences*, 2002, **99**, 12795–12800.

3 J. Wang, L. Xu and E. Wang, *Proceedings of the National Academy of Sciences*, 2008, **105**, 12271–12276.

4 J. Wang, L. Xu, E. Wang and S. Huang, *Biophysical Journal*, 2008, **99**, 29–39.

5 J. Wang, K. Zhang and E. Wang, *The Journal of Chemical Physics*, 2010, **133**, –.

6 J. Wang, K. Zhang, L. Xu and E. Wang, *Proceedings of the National Academy of Sciences*, 2011, **108**, 8257–8262.

7 H. Feng and J. Wang, *Scientific Reports*, 2012, **2**, year.

8 J. E. M. Hornos, D. Schultz, G. C. P. Innocentini, J. Wang, A. M. Walczak, J. N. Onuchic and P. G. Wolynes, *Phys. Rev. E*, 2005, **72**, 051907.

9 D. Schultz, J. N. Onuchic and P. G. Wolynes, *The Journal of Chemical Physics*, 2007, **126**, –.

10 M. Sasai, Y. Kawabata, K. Makishi, K. Itoh and T. P. Terada, *PLoS Comput. Biol.*, 2013, **9**, e1003380.

11 F. Jülicher, A. Ajdari and J. Prost, *Rev. Mod. Phys.*, 1997, **69**, 1269–1282.

12 H. Qian, *Journal of Mathematical Chemistry*, 2000, **27**, 219–234.

13 Q.-M. Nie, A. Togashi, T. N. Sasaki, M. Takano, M. Sasai and T. P. Terada, *PLoS Comput. Biol.*, 2014, **10**, e1003552.

14 H. Feng, B. Han and J. Wang, *The journal of physical chemistry. B*, 2011, **115**, 1254.

15 C. Li and J. Wang, *Journal of The Royal Society Interface*, 2013, **10**, year.

16 G. Tkaik and A. M. Walczak, *Journal of Physics: Condensed Matter*, 2011, **23**, 153102.

17 T. B. Kepler and T. C. Elston, *Biophysical journal*, 2001, **81**, 3116–3136.

18 M. Sasai and P. G. Wolynes, *Proceedings of the National Academy of Sciences*, 2003, **100**, 2374–2379.

19 F. Richard and H. Albert Roach, *Quantum mechanics and path integrals*, McGraw-Hill, New York, 1965.

20 B. Zhang and P. G. Wolynes, *Proceedings of the National Academy of Sciences*, 2014.

21 K. Zhang, M. Sasai and J. Wang, *Proceedings of the National Academy of Sciences*, 2013, **110**, 14930–14935.

22 H. Feng and J. Wang, *Scientific reports*, 2012, **2**, year.

23 E. de Wit, B. A. M. Bouwman, Y. Zhu, P. Klous, E. Splinter, M. J. A. M. Verstegen, P. H. L. Krijger, N. Festuccia, E. P. Nora,

M. Welling, E. Heard, N. Geijsen, R. A. Poot, I. Chambers and W. de Laat, *Nature*, 2012, **501**, 227.

24 H. Zhang, X.-J. Tian, A. Mukhopadhyay, K. S. Kim and J. Xing, *Phys. Rev. Lett.*, 2014, **112**, 068101.

25 S. S. Ashwin and M. Sasai, *http://arxiv.org/abs/1410.2337*, 2014.

26 K. Sneppen and I. B. Dodd, *PLoS Comput. Biol.*, 2012, **8**, e1002643.

27 A. M. Walczak, J. N. Onuchic and P. G. Wolynes, *Proceedings of the National Academy of Sciences of the United States of America*, 2005, **102**, 18926–18931.