



# Potential landscape of high dimensional nonlinear stochastic dynamics with large noise

The Harvard community has made this article openly available. [Please share](#) how this access benefits you. Your story matters

Citation	Tang, Ying, Ruoshi Yuan, Gaowei Wang, Xiaomei Zhu, and Ping Ao. 2017. "Potential landscape of high dimensional nonlinear stochastic dynamics with large noise." Scientific Reports 7 (1): 15762. doi:10.1038/s41598-017-15889-2. <a href="http://dx.doi.org/10.1038/s41598-017-15889-2">http://dx.doi.org/10.1038/s41598-017-15889-2</a> .
Published Version	<a href="https://doi.org/10.1038/s41598-017-15889-2">doi:10.1038/s41598-017-15889-2</a>
Citable link	<a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:34493224">http://nrs.harvard.edu/urn-3:HUL.InstRepos:34493224</a>
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a>

# SCIENTIFIC REPORTS



OPEN

## Potential landscape of high dimensional nonlinear stochastic dynamics with large noise

Ying Tang<sup>1</sup>, Ruoshi Yuan<sup>2,4</sup>, Gaowei Wang<sup>3</sup>, Xiaomei Zhu<sup>3</sup> & Ping Ao<sup>3</sup>

Quantifying stochastic processes is essential to understand many natural phenomena, particularly in biology, including the cell-fate decision in developmental processes as well as the genesis and progression of cancers. While various attempts have been made to construct potential landscape in high dimensional systems and to estimate transition rates, they are practically limited to the cases where either noise is small or detailed balance condition holds. A general and practical approach to investigate real-world nonequilibrium systems, which are typically high-dimensional and subject to large multiplicative noise and the breakdown of detailed balance, remains elusive. Here, we formulate a computational framework that can directly compute the relative probabilities between locally stable states of such systems based on a least action method, without the necessity of simulating the steady-state distribution. The method can be applied to systems with arbitrary noise intensities through A-type stochastic integration, which preserves the dynamical structure of the deterministic counterpart dynamics. We demonstrate our approach in a numerically accurate manner through solvable examples. We further apply the method to investigate the role of noise on tumor heterogeneity in a 38-dimensional network model for prostate cancer, and provide a new strategy on controlling cell populations by manipulating noise strength.

Studying stochastic dynamics is a central task to understand various natural and experimental phenomena in physics<sup>1,2</sup>, chemistry<sup>3</sup>, and biology<sup>4–6</sup>. Specifically, stochastic transitions induce current switching in a semiconductor<sup>7</sup>, reveal population stabilization<sup>8</sup> or extinction<sup>9</sup>, and provide a mechanistic understanding for the genesis and progression of complex diseases such as cancers<sup>10,11</sup>. Potential landscape<sup>12–14</sup>, emergent from the underlying dynamical system, serves as a powerful tool to quantify multi-stability and estimate transition rates. However, a general approach to obtain this potential landscape in practice remains elusive. A major challenge is that real-world systems are intrinsically high dimensional, e.g. the gene regulatory network<sup>4,10</sup>, which makes simulation based approaches<sup>15</sup> computationally unfeasible. In addition, systems may also subject to significant random fluctuations<sup>16–18</sup> that have functional roles such as driving cell fate decisions<sup>3,19–21</sup>. To properly integrate such noise effects can be challenging as well.

To be specific, previous attempts, which compute the potential landscape as the logarithm of a simulated steady-state distribution, suffer from the exponentially increasing computational cost of stochastic simulations, and thus encounter the curse of dimensionality. More efficient simulation approaches are developed when the detailed balance condition is satisfied<sup>22</sup>, however, this condition breaks down for nonequilibrium systems<sup>23</sup>. Except simulations, methods on the basis of the Wentzel-Kramers-Brillouin (WKB) approximation<sup>24</sup> or Freidlin-Wentzell theory<sup>25–28</sup> are proposed, but their applications are restricted to systems with small noise. Therefore, on one hand, the high dimensionality defies the use of the expensive stochastic simulation; on the other hand, stochastic simulation seems inevitable except when the noise of the system is small. This conundrum avoids quantifying stability and stochastic transitions in real-world systems.

Towards resolving this problem, we notice that information critical for many applications can be extracted from the stable states, their relative steady-state probabilities and stability. These stable states have direct

<sup>1</sup>Department of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai, 200240, China. <sup>2</sup>School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China. <sup>3</sup>Key Laboratory of Systems Biomedicine Ministry of Education, Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai, 200240, China. <sup>4</sup>Present address: Department of Systems Biology, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts, 02115, USA. Ying Tang and Ruoshi Yuan contributed equally to this work. Correspondence and requests for materials should be addressed to P.A. (email: [aoping@sjtu.edu.cn](mailto:aoping@sjtu.edu.cn))

correspondence to experimental observables such as the biological phenotypes<sup>21,29,30</sup>. To obtain such information, we develop a computational framework based on the path integral, through a least action principle and the A-type interpretation<sup>31,32</sup> of the stochastic differential equation (SDE). A dual role potential function lies at the core of the A-type interpretation, which exactly corresponds to 1) the steady-state distribution for the SDE and 2) the Lyapunov function<sup>33</sup> for the SDE's deterministic counterpart, an ordinary differential equation (ODE). With this correspondence, the relative probability between stable states can be computed directly, without the necessity of simulating the whole steady-state distribution. As a result, the present method has two major advantages: 1) it greatly reduces the computational cost as we focus on calculating the potential difference between stable states; 2) it is robust under arbitrary noise strength through the general consistency between ODE and SDE, and thus break the small noise restriction.

Our approach can be applied to a wide range of high dimensional stochastic dynamics, and enables us to investigate the role of large noise. It provides an efficient way to calculate probability ratios and transition rates between stable states. We demonstrate the method in a numerically accurate manner through an example with various noise intensities, and then apply it to a 38-dimensional network model for prostate cancer<sup>11</sup>. In particular, the tumor heterogeneity<sup>6,29,34</sup> is observed controllable by the noise intensity. The result may uncover a mechanistic basis for hyperthermia.

## Results

**Formulation.** *Deviation between ODE and SDE.* ODE and SDE can model a wide range of dynamics<sup>33,35</sup>, for example, chemical reactions<sup>2</sup>, population stabilization<sup>8</sup>, and carcinogenesis<sup>11,30</sup>. ODEs have been successfully used to quantitatively model average behaviors of a stochastic process<sup>10,36–38</sup>, for example, stable states correspond to biological phenotypes<sup>39</sup>; SDEs, by further adding a vanishing-mean noise term, recapitulate the stochastic process<sup>5,19,40,41</sup>. A major advantage of SDE modeling is the extractable dynamical information from its ODE counterpart, which may greatly reduce the computational cost of stochastic modeling. However, positions of the fixed points from the ODE can be altered after adding noise, under standard Ito's or Stratonovich's<sup>35</sup> interpretation, known as the "noise effects"<sup>42–45</sup>. We also demonstrate here that even for a two dimensional example with simple additive white noise, the dynamical structure can be greatly altered by noise, e.g. from multi-stable to uni-stable, when applying prevailing simulation methods like Ito's or Stratonovich's<sup>35</sup>, as shown in Fig. 1(a–c). Such deviation is expected to be dramatic when modeling real-world systems with significant noises<sup>16–18</sup>. The methods based on WKB approximation<sup>24</sup> or Freidlin-Wentzell theory<sup>25–28</sup> encounter this problem, and thus they are mainly applied to the cases with small noise.

We identify two independent causes for the deviation between ODE and SDE when using the prevailing stochastic integrations: 1) the existence of a non-detailed balance part; 2) a variable-dependent diffusion matrix (multiplicative noise). Both causes are related to the freedom of choosing a stochastic interpretation for SDE. With the presence of multiplicative noise, it is known that conventional stochastic integrations like Ito's and Stratonovich's lead to distinct modeling results<sup>35</sup>. However, the deviation by non-detailed balance has not been noticed before, and it can occur even for systems with additive noise, where Ito's and Stratonovich's show no difference. Therefore, the prevailing stochastic simulations could not achieve the consistency for systems without detailed balance. More details are given in Sect. I of Supplemental Information.

Regarding the issues raised above, one may ask the following question: Is there a possibility to eliminate such unexpected "noise effects" and establish a general consistency between ODE and SDE modeling even under large fluctuations? If possible, dynamical information from the ODE counterpart can be inherited by the SDE modeling, such that the valleys' positions in the landscape of SDE is obtainable by calculating stable states of ODE. The consistency is particularly necessary in a scenario where an ODE model is properly constructed and quantitatively correspond to experimental data on average, but are invalidated by the usual way of simulating SDE in reconstituting the original stochastic process and vice versa.

*Bridging ODE and SDE.* We provide a background on the framework that bridges ODE and SDE:

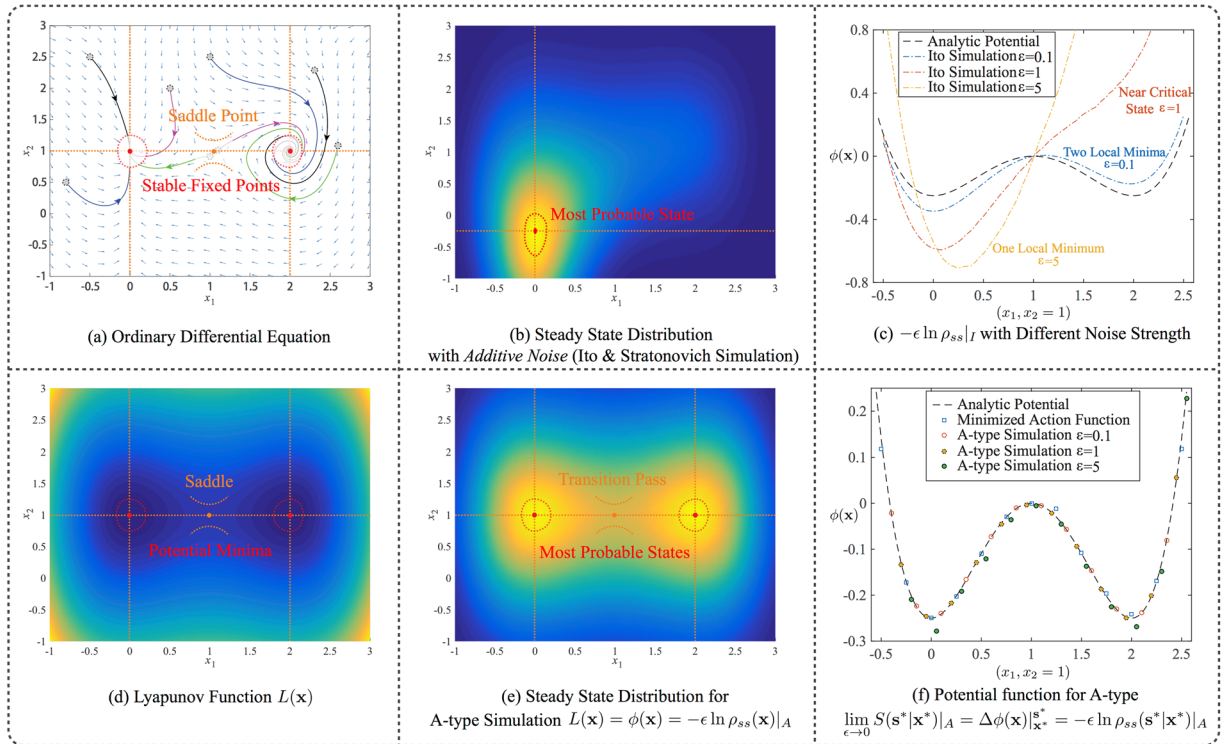
$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \quad (1)$$

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + G(\mathbf{x})\zeta(t), \quad (2)$$

where the  $N$ -dimensional vector  $\mathbf{x}$  denotes state variables, and  $\dot{\mathbf{x}}$  represents its time evolution. The deterministic part is  $\mathbf{f}(\mathbf{x})$ , and the  $M$ -dimensional Gaussian white noise  $\zeta(t)$  has  $\langle \zeta(t) \rangle = 0$ ,  $\langle \zeta(t)\zeta^\tau(t') \rangle = 2\varepsilon\delta(t-t')$ , where  $\varepsilon$  is the noise strength playing the role of temperature, the superscript  $\tau$  denotes transpose,  $\delta(t-t')$  is the Dirac delta function, and  $\langle \dots \rangle$  represents noise average. Here,  $G(\mathbf{x})G^\tau(\mathbf{x}) = D(\mathbf{x})$  defines the symmetric positive definite diffusion matrix  $D(\mathbf{x})$ . The multiplicative noise  $G(\mathbf{x})\zeta(t)$  models that system state can in turn regulate noise by feedback, or inhomogeneity of the noisy environment. The results below are valid for multiplicative noise. For Eq. (2) with multiplicative noise, a freedom in choosing the integration method leads to different stochastic interpretations<sup>35</sup>, and we will specify a choice below.

It is challenging to generally construct Lyapunov function<sup>33</sup> for ODEs, because  $\mathbf{f}$  is typically nonlinear and cannot be written directly as the gradient of a potential function  $U(\mathbf{x})$ :  $\mathbf{f} \neq -\nabla_{\mathbf{x}}U$ . Even so, a decomposed dynamics equivalent to Eq. (2) was discovered<sup>14</sup>, and Eqs (1) and (2) can be coherently decomposed as:

$$\mathbf{f}(\mathbf{x}) = -[D(\mathbf{x}) + Q(\mathbf{x})]\nabla L(\mathbf{x}), \quad (3)$$



Method	Stochastic simulation	Gradient expansion	Least “action” method (present)
Computation cost	$\sim \mathcal{O}(n^N)$	$\sim \mathcal{O}(N^2)$	$\sim \mathcal{O}(N)$

**Figure 1.** The modeling results on the potential (Lyapunov) function and steady-state distribution for an illustrative example with additive noise. Top panel: the prevailing stochastic simulation leads to a deviated stability structure. Bottom panel: The Lyapunov function, A-type simulation and potential function by least action method consistently reveal stability structure. (a) Vector field showing two stable fixed points and one saddle point. (b) Steady-state distribution by Ito’s (Stratonovich’s) simulation of SDE with  $\epsilon = 1$ , where positions of fixed points from ODE and bistable topology are altered by large noise. (c) Deviations on potential values along the line  $x_2 = 1$  between the analytical construction and Ito’s simulation of SDE with noise strengths  $\epsilon = 0.1, 1, 5$ . (d) Lyapunov function for ODE. (e) Steady-state distribution by A-type simulation of SDE with  $\epsilon = 1$ , where positions of fixed points and bistable topology are consistent with those from ODE and Lyapunov function. (f) Potential values along the line  $x_2 = 1$  calculated by least action method, and A-type simulation of SDE with noise strengths  $\epsilon = 0.1, 1, 5$ . Parameters are:  $d = a = H = 1$ . Table: computational costs of the methods on calculating potential barrier between fixed points with respect to system’s dimension  $N$ , where  $n$  is the number of mesh points in each dimension for stochastic simulation.

$$\dot{\mathbf{x}} = -[D(\mathbf{x}) + Q(\mathbf{x})]\nabla\phi(\mathbf{x}) + G(\mathbf{x}) * \zeta(t), \tag{4}$$

where the matrix  $Q(\mathbf{x})$  is anti-symmetric with  $\nabla\phi^T Q \nabla\phi = 0$ , and the asterisk means A-type stochastic integration<sup>31</sup>. A Lyapunov function  $L(\mathbf{x})$  and a potential function  $\phi(\mathbf{x})$  are constructed, and  $L(\mathbf{x})$  satisfies  $dL/dt \leq 0$  for any trajectory of Eq. (1)<sup>46–48</sup>. For Eq. (4), by solving the corresponding Fokker-Planck equation (FPE):

$$\partial_t \rho(\mathbf{x}, t) = \nabla_{\mathbf{x}}^T [D(\mathbf{x}) + Q(\mathbf{x})][\nabla_{\mathbf{x}}\phi(\mathbf{x}) + \epsilon \nabla_{\mathbf{x}}]\rho(\mathbf{x}, t), \tag{5}$$

which is obtained from the zero-mass limit on a  $2N$ -dimensional Klein-Kramers equation<sup>32</sup>, the steady state obeys Boltzmann-Gibbs distribution  $\rho_s(\mathbf{x}) = \exp[-\phi(\mathbf{x})/\epsilon]$ . As the steady state is invariant under transformation  $\phi \rightarrow \phi + C$  for any constant  $C$ , we have chosen  $C$  such that the distribution is normalized.

*A-type integration* is defined as the connection between the SDE (2) and the FPE (5), and realized by two explicit limiting procedures: first the usual integration limit and then the zero mass limit<sup>32</sup>. Even for systems with additive noise, it is different from the conventional  $\alpha$ -type stochastic integration<sup>31</sup> ( $\alpha = 0$  is Ito’s,  $\alpha = 1/2$  is Stratonovich’s), except that when detailed balance condition holds ( $Q = 0$ ) it corresponds to  $\alpha = 1$ . An exact transformation from A-type integration to Ito’s has been achieved<sup>31</sup>. *A-type simulation* can thus be implemented as follows: Eq. (4) is transformed to be an equivalent SDE under Ito’s interpretation:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \epsilon \Delta \mathbf{f}(\mathbf{x}) + G(\mathbf{x}) \cdot \zeta(t), \tag{6}$$

where  $\Delta \mathbf{f}_i(\mathbf{x}) = \sum_j \partial_{x_j} [D_{ij}(\mathbf{x}) + Q_{ij}(\mathbf{x})]$ , and the dot denotes Ito’s integration. Thus, one can simulate Eq. (6) with Ito’s scheme to realize A-type simulation for Eq. (2).

An advantage led by the A-type simulation is that for arbitrary noise strength the sampled steady-state distribution of Eq. (2) corresponds to the Lyapunov function and potential function:

$$L(\mathbf{x}) = \phi(\mathbf{x}) = -\varepsilon \ln \rho_{ss}(\mathbf{x})|_A, \quad (7)$$

where the subscript A denotes A-type simulation.

Then, the deterministic part of Eq. (2) has the same decomposition as Eq. (1), and positions of fixed points from Eq. (1) are not changed after added noise. For Ito's or Stratonovich's integrations one cannot recognize fixed points of ODEs from the simulated distribution even for additive noise, as shown in Fig. 1(a–c). A-type simulation also reserves topology of the landscape for arbitrary noise strength<sup>31,47</sup>, as exemplified in Fig. 1(d–f). Note that two numerical methods have been used to estimate the steady-state distribution or potential function in Eq. 7: the A-type simulation<sup>31</sup> and the gradient expansion method<sup>14</sup>. To implement these two methods, the matrix  $Q(\mathbf{x})$  needs to be solved.

**Efficient computation of potential difference.** The essential information for multi-stable systems is the relative stability between stable states, which can be extracted from the potential difference<sup>1,2</sup>. Based on the property that fixed points for ODE and locally most probable states for SDE are identical in our framework, we have the following protocol for a typical situation, where two stable fixed points  $\mathbf{x}_1^*$  and  $\mathbf{x}_2^*$  are connected by a saddle point  $\mathbf{s}^*$  (Protocol I):

1. Calculate the potential difference between each stable fixed point and the saddle point  $\Delta\phi(\mathbf{x})|_{\mathbf{x}_1^*}^{\mathbf{s}^*}$ ,  $\Delta\phi(\mathbf{x})|_{\mathbf{x}_2^*}^{\mathbf{s}^*}$  by the least action method Eq. (14) below. The detail of finding the least action path is given in Methods. The potential difference between  $\mathbf{x}_1^*$  and  $\mathbf{x}_2^*$  is  $\Delta\phi(\mathbf{x})|_{\mathbf{x}_1^*}^{\mathbf{x}_2^*} = \Delta\phi(\mathbf{x})|_{\mathbf{x}_1^*}^{\mathbf{s}^*} - \Delta\phi(\mathbf{x})|_{\mathbf{x}_2^*}^{\mathbf{s}^*}$ .

The path integral formulation for Eq. (2) is applied here to calculate the potential difference. The formulation needs to be consistent with the stochastic integration used<sup>49</sup>. For A-type integration,  $P(\mathbf{s}^*, T_2|\mathbf{x}^*, T_1) = \int_{\mathbf{x}^*}^{\mathbf{s}^*} \mathcal{D}_A \mathbf{x} \exp\{-S_T[\mathbf{x}]|_A/\varepsilon\}$ , where the action  $S_T[\mathbf{x}]|_A$  is a function of paths  $[\mathbf{x}(t)]$  with  $\mathbf{x}^*$  and  $\mathbf{s}^*$  as start point at time  $T_1$  and end point at  $T_2$  separately. As we have transformed Eqs (2) to (6), it is more convenient to use the equivalent path integral formulation:

$$P(\mathbf{s}^*, T_2|\mathbf{x}^*, T_1) = \int_{\mathbf{x}^*}^{\mathbf{s}^*} \mathcal{D}_I \mathbf{x} \exp\{-S_T[\mathbf{x}]|_I/\varepsilon\}, \quad (8)$$

where the subscript I means Ito's integration. The measure on paths is defined as:  $\int_{\mathbf{x}^*}^{\mathbf{s}^*} \mathcal{D}_I \mathbf{x} \equiv \lim_{K \rightarrow \infty} \prod_{k=1}^{K-1} \int d\mathbf{x}^k / \sqrt{\det[2\pi\varepsilon dt D(\mathbf{x}^{k-1})]}$ , where functions of  $\mathbf{x}$  take pre-points in each interval. The time is discretized into  $K$  segments with  $T_1 = t_1 < \dots < t_k < \dots < t_K = T_2$ , each interval being  $dt$  and  $\mathbf{x}^k = \mathbf{x}(t_k)$ . The action function:

$$S_T[\mathbf{x}]|_A = \frac{1}{4} \int_{T_1}^{T_2} dt \left[ \dot{\mathbf{x}} - \mathbf{f}(\mathbf{x}) - \varepsilon \Delta \mathbf{f}(\mathbf{x}) \right]^T D^{-1}(\mathbf{x}) \left[ \dot{\mathbf{x}} - \mathbf{f}(\mathbf{x}) - \varepsilon \Delta \mathbf{f}(\mathbf{x}) \right], \quad (9)$$

where the integration obeys Ito's rule<sup>35</sup>. Note that the present action function is different from that of Freidlin-Wentzell's framework<sup>25</sup> except when  $\varepsilon \rightarrow 0$ . Even for SDE with additive noise, the difference still exists, because  $\Delta \mathbf{f}(\mathbf{x}) = \sum_j \partial_{x_j} Q_{ij}(\mathbf{x})$  can be nonzero for systems without detailed balance.

By using the decomposition in Eq. (4), we have:

$$\begin{aligned} S_T[\mathbf{x}]|_A &= \frac{1}{4} \int_{T_1}^{T_2} dt \left[ \dot{\mathbf{x}} - D\nabla\phi + Q\nabla\phi - \varepsilon\Delta\mathbf{f} \right]^T D^{-1}(\dot{\mathbf{x}} - D\nabla\phi + Q\nabla\phi - \varepsilon\Delta\mathbf{f}) \\ &+ \int_{T_1}^{T_2} dt \dot{\mathbf{x}}^T \nabla\phi + \int_{T_1}^{T_2} dt (\nabla\phi)^T (Q\nabla\phi - \varepsilon\Delta\mathbf{f}) \\ &\geq \Delta\phi|_{\mathbf{x}^*}^{\mathbf{s}^*} - \varepsilon \int_{T_1}^{T_2} dt D\nabla^2\phi - \varepsilon \int_{T_1}^{T_2} dt (\nabla\phi)^T \Delta\mathbf{f}, \end{aligned} \quad (10)$$

where we have used  $\nabla\phi^T Q\nabla\phi = 0$ , and Ito's formula<sup>35</sup>:  $\int_{T_1}^{T_2} dt \dot{\mathbf{x}}^T \nabla\phi = \Delta\phi|_{\mathbf{x}^*}^{\mathbf{s}^*} - \varepsilon \int_{T_1}^{T_2} dt D\nabla^2\phi$  with  $D\nabla^2$  denoting  $D_{ij} \partial_{x_i} \partial_{x_j}$ . For clarity, we ignore the symbol  $(\mathbf{x})$  for functions of  $\mathbf{x}$  in the derivation. The inequality in Eq. (10) becomes equality for the least action path:

$$\dot{\mathbf{x}} = D(\mathbf{x})\nabla\phi(\mathbf{x}) - Q(\mathbf{x})\nabla\phi(\mathbf{x}) + \varepsilon\Delta\mathbf{f}(\mathbf{x}), \quad (11)$$

where the trajectory tends to go reversely from  $\mathbf{s}^*$  to  $\mathbf{x}^*$ . Thus, Eq. (9) counts the accumulation of noise, whose minimization equals to the ‘‘uphill’’ energy  $\Delta\phi(\mathbf{x})|_{\mathbf{x}^*}^{\mathbf{s}^*}$ . When the trajectory passes the saddle point, it goes ‘‘downhill’’ obeying  $\dot{\mathbf{x}} = -D(\mathbf{x})\nabla\phi(\mathbf{x}) - Q(\mathbf{x})\nabla\phi(\mathbf{x}) + \varepsilon\Delta\mathbf{f}(\mathbf{x})$ , where minimization of Eq. (9) is zero.

In the limit of  $\varepsilon \rightarrow 0$ ,

$$S_T[\mathbf{x}]|_A \geq \Delta\phi(\mathbf{x})|_{\mathbf{x}^*}^{\mathbf{s}^*}, \quad (12)$$

and the least action path follows the time-reversal adjoint dynamics<sup>50</sup> of Eq. (1) with decomposition Eq. (3):  $\dot{\mathbf{x}} = D(\mathbf{x})\nabla\phi(\mathbf{x}) - Q(\mathbf{x})\nabla\phi(\mathbf{x})$ . Since the dynamics is deterministic, the least action path connecting  $\mathbf{s}^*$  to  $\mathbf{x}^*$  is specified. Then, minimization of the action function

$$S(\mathbf{s}^*|\mathbf{x}^*)|_A \doteq \inf_{\{T>0\}} \inf_{\{x(T_1)=\mathbf{s}^*, x(T_2)=\mathbf{x}^*\}} S_T[\mathbf{x}]|_A \quad (13)$$

equals to the potential difference between  $\mathbf{s}^*$  and  $\mathbf{x}^*$ , which is unique if the ideal least action path following  $\dot{\mathbf{x}} = D(\mathbf{x})\nabla\phi(\mathbf{x}) - Q(\mathbf{x})\nabla\phi(\mathbf{x})$  is found.

With Eq. (7), we reach a formula to calculate potential barrier:

$$\lim_{\varepsilon \rightarrow 0} S(\mathbf{s}^*|\mathbf{x}^*)|_A = \Delta\phi(\mathbf{x})|_{\mathbf{x}^*}^{\mathbf{s}^*} = -\varepsilon \ln \rho_{ss}(\mathbf{s}^*|\mathbf{x}^*)|_A. \quad (14)$$

We emphasize that the first equality holds only when  $\varepsilon \rightarrow 0$ , but the second equality is for arbitrary noise strength, which is different from that of Freidlin-Wentzell's framework valid for  $\varepsilon \rightarrow 0$ <sup>25</sup>. The significance of Eq. (14) is that it enables to obtain potential difference by minimizing the action when  $\varepsilon \rightarrow 0$ , and the result exactly corresponds to that of Lyapunov function for ODE and by A-type simulation of SDE with arbitrary noise strength.

**Probability ratio and transition rate.** As the steady state obeys the Boltzmann-Gibbs distribution, the probability ratio between stable states is:

$$\frac{\rho(\mathbf{x}_2^*)}{\rho(\mathbf{x}_1^*)} = \exp\left[-\frac{1}{\varepsilon} \Delta\phi(\mathbf{x})\right]_{\mathbf{x}_1^*}^{\mathbf{x}_2^*}, \quad (15)$$

cell state where  $\Delta\phi(\mathbf{x})|_{\mathbf{x}_1^*}^{\mathbf{x}_2^*} = \Delta\phi(\mathbf{x})|_{\mathbf{x}_1^*}^{\mathbf{s}^*} - \Delta\phi(\mathbf{x})|_{\mathbf{x}_2^*}^{\mathbf{s}^*}$ . Equation (15) is valid under arbitrary noise strength, and thus can show variation of the ratio with different noise intensities. It provides the probability ratios of quantities such as the number of different cell states. Specifically, we apply it to analyze tumor heterogeneity by large noise in the section of Application.

When  $\varepsilon$  is small compared to the height of the potential barrier  $\Delta\phi(\mathbf{x})|_{\mathbf{x}_1^*}^{\mathbf{s}^*}$ , the asymptotic transition rate formula<sup>1,2</sup> from the stable fixed point  $\mathbf{x}_1^*$  to  $\mathbf{x}_2^*$  is:

$$R(\mathbf{x}_2^*|\mathbf{x}_1^*) \propto \exp\left[-\frac{1}{\varepsilon} \Delta\phi(\mathbf{x})\right]_{\mathbf{x}_1^*}^{\mathbf{s}^*}. \quad (16)$$

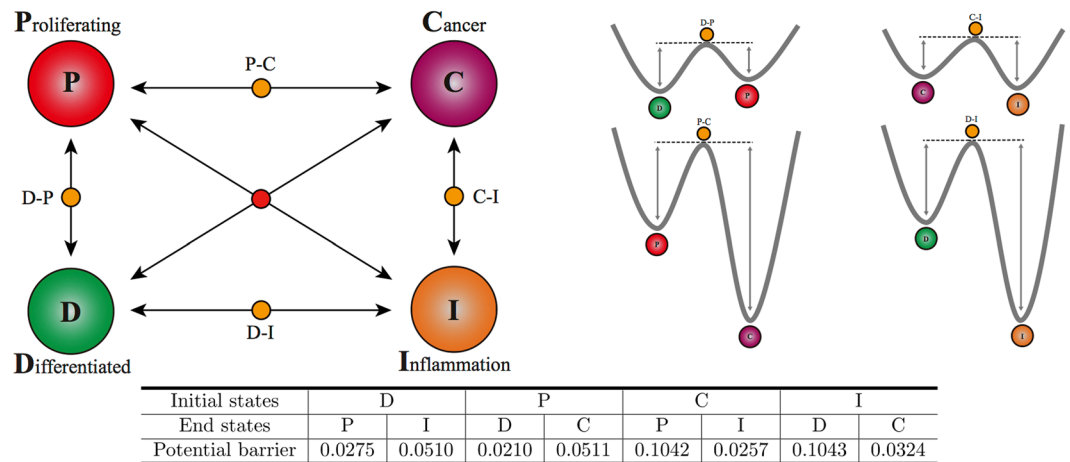
Different from<sup>51</sup>, in our framework the non-detailed balance part does not provide correction terms that explicitly appear in the pre-factor of the rate formula as analyzed in Sect. IV of Supplemental Information.

**Comparison on computational cost.** Computational costs of methods mentioned above for systems with respect to dimension  $N$  are analyzed here. For stochastic simulation of Eq. (4), we mesh each dimension into  $n$  points, and the computational cost is exponentially proportional to dimension,  $cost \sim \mathcal{O}(n^N)$ . This method also has the problem of slow convergence when noise strength is small. For the gradient expansion<sup>14</sup>, as in each step a matrix needs to be evaluated, the computational cost is approximately  $cost \sim \mathcal{O}(N^2)$ . Both of the two methods need to solve the matrix  $Q$  for A-type interpretation.

However, information critical for many stochastic models is mainly obtained from estimating the relative probabilities between stable states. Under such circumstances, it is sufficient to find the potential differences between fixed points, without the necessity of obtaining the whole steady-state distribution. Thus, we only need to calculate the one dimensional least action path connecting fixed points, and the computational cost should be linearly proportional to the dimension,  $cost \sim \mathcal{O}(N)$ . As a result, the least action method is efficient in high dimensional systems. We list the results in the table of Fig. 1.

**Protocol to obtain a global landscape.** Protocol I can be extended to obtain a global landscape for systems with multiple stable states (Protocol II):

1. Find the fixed points under consideration from solving the ODE counterpart, such as by Newton iteration method<sup>52</sup>.
2. Classify all the fixed points into two groups by calculating the eigenvalues of the linearized Jacobian matrix in their neighborhood: stable fixed points (no eigenvalue with positive real part) and unstable points (at least one eigenvalue with positive real part).
3. Choose a saddle point as reference. Start from the points in a small neighborhood of the saddle point, and simulate the ODE to find all the stable fixed points reached. Calculate potential difference between the saddle point and the stable fixed points by the least action method in Eq. (14). The detail of the minimization procedure in Eq. (14) is given in Methods.
4. Repeat step 3 for all saddle points. Assign relative potential difference between the saddle points if they reach a common stable fixed point.
5. For any other points in state space, simulate the ODE to find the fixed point it reaches. Obtain their potential difference by the least action method. The total computational cost depends on the potential value of how many points are calculated.
6. With the calculated potential values, extract the relative probabilities between the states by Eq. (15).



**Figure 2.** Potential barriers calculated by Protocol II in the prostate cancer model<sup>11</sup>. Left panel: The chosen four cell states: differentiated (D), proliferating (P), cancer (C), inflammation (I). They are stable fixed points obtained from ODE of the 38-dimensional system. The states D-P, D-I, C-I, P-C are saddle points, and the red fixed point in the middle is unstable. Right panel: the heights of potential barriers between stable fixed points connected by saddles. The lengths of arrows are proportional to barrier heights listed in the table below. Table: potential barriers between stable fixed points are calculated by the least action method. We set  $K = 100$ ,  $T = 20$ , and have checked that larger  $K$  and  $T$  values lead to convergent results. The parameters of the system chosen here are for typical cancer patients<sup>11</sup>, where cancer and inflammation states are more stable.

The consistency between ODE and SDE enables to utilize information of fixed points and basins of attraction from ODE, which greatly improves the efficiency of our algorithm. As each point in state space reaches a single fixed point, its potential value is uniquely determined if the ideal least action method is found numerically during minimization procedure, which leads to a global landscape without ambiguity. Specifically, the probability ratio between a fixed and a point within the potential well is calculated by Eq. (15), which corresponds to cell-to-cell variability inside the attractor. For dynamical systems with complex attractors such as limit cycle<sup>33,47,53</sup>, the point on the stable limit cycle can be treated similarly as the stable fixed point, and thus our method can be generalized.

**Application.** Heterogeneity of cell populations is widely observed in biological systems such as cancer<sup>54</sup>, where different cell phenotypes emerge in tumor tissues<sup>34,55</sup>. It is proposed that an underlying regulation network and quantification on the network dynamics by SDE models can describe various cell states and transitions between them<sup>6,15</sup>. Starting from the SDE model, the calculated potential landscape provides an integrated picture to study heterogeneity. Specifically, valleys in the landscape correspond to different cell states, and the potential barrier separating them quantifies the transition rates. This approach of landscape is helpful to understand systematically the effect of perturbations on cell state interconversions.

Here, we investigate whether the variation of noise strength leads to changes of cell states, as noise plays a crucial role in biological processes, for example, it drives the cell fate decision<sup>20,21</sup>. The previous methods<sup>24–28</sup> can not be applied to study the function of large noise, because identification on valleys of landscape and calculation on potential barrier by these methods are restricted to the zero noise limit. Now, we are able to quantify the role of large noise on heterogeneity in high dimensional network dynamics, because the present calculation on landscape is robust under arbitrary noise strength. From our method, the ratios of cell states can be controlled by manipulating noise strength, which allows the cell-to-cell variability under the same gene regulation network.

**Prostate cancer model.** As an illustrative example, we demonstrate the effectiveness of our approach by applying it to a network model for the prostate cancer<sup>11</sup>. The network dynamics is modeled by a 38-dimensional SDE. Each dynamical variable represents the expression or activity level of a gene. The dynamics of each component is written as a sum of the generation and degradation, including the activation or inhibition by the other genes. We use the standard Hill equation to model such interactions<sup>56,57</sup>.

The ODE counterpart and the parameters are given in Sect. VII of the Supplemental Information. From the analysis on the ODE, we know the system has 10 stable fixed points and 16 saddle points. We find that four stable fixed points correspond to various cell states shown in Fig. 2: differentiated (D), proliferating (P), cancer (C), inflammation (I) with their positions given in<sup>11</sup>. For clarity, we consider additive noise case with diffusion matrix  $D(\mathbf{x})$  as an identity matrix in this example. It should be emphasized that the deviation between ODE and SDE appears even with additive noise, because this system does not obey the detailed balance condition. As a result, considering additive noise is sufficient to demonstrate the advantage of our method based on A-type integration compared with the prevailing methods based on other stochastic integrations.

We use the least action method to calculate heights of the potential barriers between stable states, as shown in Fig. 2. We note that in order to have the global landscape, we use the continuous condition to set the potential value of  $P$  to  $D$  specifically. According to Eq. (15), the probability ratios of different cell states can be calculated under various noise strengths. We list the ratios between the states of D, P, C, I in Fig. 3, which shows qualitative

Cell types Noise ↓ More	Stable states	D	P	C	I
	Ratios ( $\epsilon = 0.01$ )		0.0048	0.0025	0.5117
Ratios ( $\epsilon = 0.02$ )		0.0696	0.0501	0.7154	1
Ratios ( $\epsilon = 0.05$ )		0.3448	0.3034	0.8759	1
Ratios ( $\epsilon = 0.1$ )		0.5882	0.5471	0.9353	1
Ratios ( $\epsilon = 1$ )		0.9524	0.9429	0.9905	1

**Figure 3.** The probability ratios between D, P, C, I states of the prostate cancer model. The ratios are calculated by Eq. (15) with heights of potential barriers listed in the table of Fig. 2, where the values are normalized by the I state. We consider noise intensities:  $\epsilon = 0.01, 0.02, 0.05, 0.1, 1$ . The tumor heterogeneity emerges when noise strength becomes large where the four types of cells are almost equally distributed.

different results with changing noise. When noise is small, e.g.  $\epsilon = 0.01$ , most of the cells belong to the cancer state and the inflammation state, whereas little are in the differentiated state and the proliferating state. When noise becomes large such as  $\epsilon = 1$ , various cell states are almost equal in number. This demonstrates the emergence of tumor heterogeneity with respect to increasing the noise strength.

We elucidate more on the application of controlling the ratios of cell states through varying the noise intensity, which can be implemented by tuning temperature. First, it was demonstrated that an therapy with combination of hyperthermia and other treatments, such as immunotherapy and radiotherapy, can improve the efficiency of cancer cure<sup>58</sup>. In those cases, temperature plays the role of enhancer to switch on and off the effectiveness of other therapies. Specifically, drug cytotoxicity triggered by temperature variation leads to the death of tumor cells, and therefore combination of hyperthermia and chemotherapy is regarded as an effective treatment of cancer<sup>59</sup>. Second, as different levels of heating were found to bring distinct modulatory effect on tumor targets, our method valid for arbitrary noise strength may be applied to study sensitivity of thermal treatment regulated by temperature, which will provide new designs on clinical trials. Third, the regional hyperthermia to radiotherapy<sup>60</sup> shows an improvement on survival rates of cancer patients, because hyperthermia can guide the action of chemotherapy to specific heated tumor region. This can be modeled by multiplicative noise, as exemplified in Sect. I of the Supplemental Information. Therefore, the present approach could form the theoretical basis for hyperthermia that employs effect of temperature in tumor treatment.

With the obtained potential barrier, transition rates of cell state interconversions are given by Eq. (16), which provides a quantitative understanding on the cancer genesis. Under the given parameters and noise strength, the result provides a set of predictions: 1) the cancer and the inflammation states are more stable than the proliferating and the differentiated states; 2) transitions from the cancer state to the proliferating state and from the inflammation state to the differentiated state are difficult than the other way around; 3) transition to the cancer state from the inflammation state is more frequent than from the proliferating state. These suggest that the model may describe a cancer patient, and new strategies for medical treatments should be designed to raise the potential energy of the cancer and the inflammation states.

## Discussion

The above results depend on a specific choice on the stochastic integration. From the mathematical aspect, different stochastic integrations are equivalent, and can be transformed to each other by modifying the drift term  $\mathbf{f}(\mathbf{x})$  correspondingly. From the physical aspect, for a system with clearly separated sources for the deterministic force and the stochastic force, i.e. the two forces can be measured independently in an experiment, we might say that a stochastic interpretation for the SDE is chosen by Nature. There are particular scenarios in which such measurements are possible, for example, the experiments<sup>61,62</sup> have shown that A-type is chosen for a class of systems, which demonstrates that the A-type integration is not only a theoretical treatment.

However, in many extended applications such as phenomenological models for biological systems, separating deterministic and stochastic forces would not even be meaningful, and we would rather consider stochastic integrations as mathematically equivalent tools that is chosen by modeling. Under the situation, SDE model set by a combination of the drift term  $\mathbf{f}(\mathbf{x})$  and stochastic integration can be non-unique, and each is an effective description for the system. If the deterministic rate equations are properly constructed and quantitatively correspond to the experimental measurements on average, e.g. the stable fixed points from the ODE part are consistent with the distribution peaks from data, then A-type integration should be a proper choice for calculating the steady-state distribution. Because the correspondence between ODE and SDE modeling under arbitrary noise strength is a property for the A-type integration. The cancer model discussed above belongs to this category, as its ODE part was demonstrated to match the experimental observation on average<sup>11</sup>.

In biology, noise has a variety of sources<sup>63</sup>, such as locations of molecules, micro-environmental fluctuations, gene expression noise, and cellular processes like cell growth. For complex systems like cancer, noise may come from different sources. SDE model reconstitutes the random fluctuations into a single noise term, which reflects the various sources of noise<sup>19,64</sup>. Therefore, several experimental operations can implement the change of the noise strength discussed here in real biological systems.

Our method can be applied to systems that are modeled by master equation (CME) with discrete dynamical variable<sup>35</sup>. First, CME can be transformed to be the chemical Langevin equation with continuous variable<sup>40</sup>, which can be cast into the form of Eq. (2). Then, our method is applicable to improve efficiency. The approximation is tolerably accurate when the copy number of variables are large, and it also requires that the dynamical process



has a time scale during which multiple reactions occur and the reaction rate does not change dramatically<sup>40</sup>. These conditions are expected to hold for the present high dimensional cancer dynamics<sup>10,11</sup>, where the proteins usually has high copy numbers. Second, CME may also be expanded to a FPE and further corresponds to Eq. (4) with consistent modeling predictions, as demonstrated through an explicit procedure in Sect. II of Supplemental Information. Third, for systems with low copy numbers, SDE can still provide an appropriate description on the effect of noises<sup>19</sup>. Fourth, for stochastic processes on the level of single molecules, such as gene burst process<sup>65</sup>, CME is a more proper approximation to capture the discrete nature of species<sup>9</sup>. Nevertheless, this kind of noise will diminish by accumulation of proteins with long lifetime<sup>66</sup>.

Mathematically, SDE and CME are two independent modeling methodologies, and are on an equal footing to describe the stochastic dynamics. Both CME and SDE are models with intrinsic discrepancy to the real dynamical process. From computational side, a whole set of CME to describe the stochastic dynamics in detail is typically high dimensional, and the Gillespie algorithm<sup>67</sup> to simulate CME is time consuming. Thus, the present method handling SDE valid for arbitrary noise strength is practically useful to investigate high dimensional systems with large fluctuations, particularly when the ODE counterpart is properly constructed and quantitatively correspond to the average experimental data.

Several other remarks are in order. First, our calculation on the potential difference is applicable to systems both with and without detailed balance condition<sup>14</sup>, i.e.  $Q = 0$  or not. Breakdown of detailed balance inducing a curl flux in the state space affects the least action path, and generally leads to  $S(\mathbf{x}_1^*|\mathbf{x}_2^*)|_A \neq S(\mathbf{x}_2^*|\mathbf{x}_1^*)|_A$ . For such cases, the least action path also differ from the deterministic saddle-node trajectories<sup>25</sup>. Second, there are many efficient numerical methods to calculate fixed points of ODEs in high dimension<sup>52</sup>, such as Newton iteration method. Third, the present action function has the dimension of energy, and the conventional action in classical physics has the dimension of energy multiplied by time. Fourth, positions for the locally most probable states in Eq. (3) is a subset of fixed points' positions for Eq. (1), because  $\nabla\phi(\mathbf{x}) = 0$  is sufficient but not necessary to  $\mathbf{f}(\mathbf{x}) = -[D(\mathbf{x}) + Q(\mathbf{x})]\nabla\phi(\mathbf{x}) = 0$ . Fifth, the constructed potential function is also useful to extract thermodynamical free energy for non-equilibrium systems<sup>68,69</sup>.

We next compare our framework with the previous works. First, authors in<sup>24</sup> used a WKB method and reached a Hamiltonian-Jacobi equation  $\nabla\phi(\mathbf{x})^T D(\mathbf{x})\nabla\phi(\mathbf{x}) + \nabla\phi(\mathbf{x})^T \mathbf{f}(\mathbf{x}) = 0$  in the lowest order of  $\varepsilon \rightarrow 0$ . However, computational cost of solving this partial differential equation increases exponentially. This Hamiltonian-Jacobi equation is valid for arbitrary orders of  $\varepsilon$  in our framework<sup>32</sup>. Second, our method is different from the previous path integral approach<sup>15</sup>, where their action function gives an effective potential rather than the exact potential  $\phi(\mathbf{x})$  constructed consistently in Eqs (3) and (4). Third, the unexpected "noise effects" in using SDEs have been widely reported<sup>42-44</sup>, and whether the phenomena are produced by the physical effect of the zero-mean noise or by the intricacy of using various stochastic integrations<sup>45,70</sup> (also see the example in Sect. II of Supplemental Information) is a question without a definite answer. These effects in general defy the use of dynamical information from the ODE counterpart, and our method provide a possibility to reserve useful results by ODE analysis for SDE with arbitrary noise strength.

We analyze the difference between our framework and Freidlin-Wentzell's<sup>25</sup>, based on which the quasi-potential has been calculated in many systems recently<sup>26-28</sup>. First, the consistency between action function's form and stochastic integration (classified in Table II of Supplemental Information) has not been considered in Freidlin-Wentzell's framework. Only the present action function of Ito's form is the same as the usual Freidlin-Wentzell's action<sup>25</sup>. Second, our decomposition  $\mathbf{f}(\mathbf{x}) = -[D(\mathbf{x}) + Q(\mathbf{x})]\nabla\phi(\mathbf{x})$  with  $\nabla\phi(\mathbf{x})Q(\mathbf{x})\nabla\phi(\mathbf{x}) = 0$  is generally different from the usual Freidlin-Wentzell form  $\mathbf{f}(\mathbf{x}) = -\nabla U(\mathbf{x}) + l(\mathbf{x})$  with  $\nabla U(\mathbf{x}) \cdot l(\mathbf{x}) = 0$ , except when diffusion matrix  $D(\mathbf{x})$  is proportional to identity. For general  $D(\mathbf{x})$ , the minimization of action function Eq. (17) does not directly equal to the function  $U(\mathbf{x})$  even in the limit of  $\varepsilon \rightarrow 0$ . For example, when  $D(\mathbf{x})$  is a diagonal matrix with distinct constant elements, the action  $S_T[\mathbf{x}]|_I \geq \sum_i \int_{T_1}^{T_2} dt (\dot{x}_i \partial_{x_i} U/D_{ii} - l_i \partial_{x_i} U/D_{ii}) \neq \Delta U(\mathbf{x})|_{x^*}$ . On the other hand, if we apply the action without the diffusion matrix as in<sup>26</sup>,  $\hat{S}_T(\mathbf{x}) = \sum_i \int_{T_1}^{T_2} dt (\dot{x}_i - f_i)^2/2$ , it does not include the effect of  $D(\mathbf{x})$ . A more detailed comparison is given in Sect. V of Supplemental Information.

Recently, the existence of decomposition Eq. (4) and its equivalence to Freidlin-Wentzell's quasi-potential in the zero noise limit is discussed<sup>71</sup>. Here, we provide a framework to directly extend the potential function to the cases with finite noise. Besides, the classical Freidlin-Wentzell's quasi-potential is locally constructed around each stable attractors, and complicated methods to glue the locally constructed quasi-potentials are required<sup>25</sup>. In this paper, we find that a unique potential function can be obtained by finding the ideal least action path for the minimization in Eq. (13). Even  $Q$  in Eq. (4) generally may not be uniquely determined for a given Eq. (2)<sup>71</sup>, certain boundary conditions or physical constrains can be added to specify  $Q$  in order to guarantee a unique least action path for Eq. (13).

In order to be more efficient in finding the least action path as given by the time-reversal adjoint dynamics, here we provide a strategy of minimization called ODE-based-adaptive-time method. First, we simulate ODE, Eq. (1), to get a path from  $s^*$  (considering  $s^*$  is a saddle point, add a small perturbation to it as the starting point) to a corresponding stable fixed point  $x^*$ . We record the total duration for the trajectory as  $T_{s^* \rightarrow x^*}^{ODE}$ . Since the least action path is time-reversal of the adjoint dynamics, its duration and length should be the same as that of the corresponding trajectory for the original ODE. Therefore, we choose  $T_2 - T_1 = T_{s^* \rightarrow x^*}^{ODE}$ , and put the constrain  $\inf_{\{T_2 - T_1 = T_{s^* \rightarrow x^*}^{ODE}\}}$  to the minimization in Eq. (13). Note that the constrain of trajectory time here could not guarantee to find the ideal least action path, however, it can reduce the sampling space of paths connecting  $x^*$  to  $s^*$ . To ensure that the ideal least action path is picked, new types of minimization algorithm considering dynamics by  $-Q(\mathbf{x})\nabla\phi(\mathbf{x})$  need to be developed.

To conclude, we have identified two independent causes for the deviation between the SDE and its ODE counterpart when using the prevailing stochastic integrations: the existence of a non-detailed balance component or a variable-dependent diffusion matrix (multiplicative noise). We have developed a new numerical approach to study multi-stability and stochastic transitions between stable states for the SDE modeling. We are able to compute efficiently probability ratios between stable states in high dimensional nonequilibrium systems subject to large noise, through a least action method under the A-type stochastic interpretation. The modeling results of a prostate cancer network reveal a new mechanism to control the ratios of cell states by manipulating the noise intensity. Our approach should also be practically useful to study the role of the noise in the dynamical modeling of many other real-world high dimensional stochastic processes.

## Methods

We have demonstrated in Sect. III of Supplemental Information that differences of action functions for Eq. (2) with various stochastic integrations can be neglected when  $\varepsilon \rightarrow 0$ . Therefore, we can choose action with specific stochastic integration for the convenience of numerical calculations. Here, we adopt the action with Itô's integration<sup>49</sup>:

$$S_T[\mathbf{x}]|_I = \frac{1}{4} \int_{T_1}^{T_2} dt [\dot{\mathbf{x}} - \mathbf{f}(\mathbf{x})]^T D^{-1}(\mathbf{x}) [\dot{\mathbf{x}} - \mathbf{f}(\mathbf{x})]. \quad (17)$$

When  $\varepsilon \rightarrow 0$ , its minimization equals to the potential difference  $\lim_{\varepsilon \rightarrow 0} S(\mathbf{s}^*|\mathbf{x}^*)|_I = \lim_{\varepsilon \rightarrow 0} S(\mathbf{s}^*|\mathbf{x}^*)|_A = \Delta\phi(\mathbf{x})|_{\mathbf{x}^*}$  with the least action path satisfying  $\dot{\mathbf{x}} = D\nabla\phi - Q\nabla\phi$ , and deviation between the least action path and that by  $\dot{\mathbf{x}} = -D\nabla\phi - Q\nabla\phi + \varepsilon\Delta\mathbf{f}$  for Eq. (9) disappears as well.

Numerically, we use discretized form of Eq. (17) as the object function to minimize. The discretized scheme adopted corresponds to the stochastic integration<sup>49</sup>, such as pre-point scheme is needed for Itô's interpretation. Thus, we minimize the discretized action:

$$S_T[\mathbf{x}]|_I = \frac{1}{4} \sum_{k=1}^K \Delta t_k \left[ \frac{\mathbf{x}^k - \mathbf{x}^{k-1}}{\Delta t_k} - \mathbf{f}(\mathbf{x}^{k-1}) \right]^T D^{-1}(\mathbf{x}^{k-1}) \left[ \frac{\mathbf{x}^k - \mathbf{x}^{k-1}}{\Delta t_k} - \mathbf{f}(\mathbf{x}^{k-1}) \right], \quad (18)$$

where we divide the time intervals to be  $T_1 = t_0 < \dots < t_k < \dots < t_K = T_2$  with  $\Delta t_k = t_k - t_{k-1}$ . The trajectory  $\mathbf{x}(t)$  is divided into  $K$  pieces with  $\mathbf{x}^k = \mathbf{x}(t_k)$ . Besides, Eq. (17) is explicitly independent of  $\varepsilon$ , and thus its numerical minimization has implicitly taken the limit  $\varepsilon \rightarrow 0$ . We choose a straight line connecting the two points as initial path, and set  $T = -T_1 = T_2$ . We use the `fmincon` function with the interior-point algorithm in the toolbox of MATLAB to do minimization. A package of MATLAB code to implement our method is provided in supplementary material. The minimization procedure requires to calculate gradient of the action function for the optimization step, and each step has a cost linearly proportional to dimension. Thus, in numerical realization the computational cost may become superlinear to dimension, and is at most quadratic to dimension. Such increasing demand on computational power is also encountered for the gradient expansion method. Therefore, both the stochastic simulation and the gradient expansion method are more computationally expensive than the present method.

For Eq. (2), when  $Q(\mathbf{x})\nabla\phi(\mathbf{x})$  is dominate compared to  $D(\mathbf{x})\nabla\phi(\mathbf{x})$ , the least action path rotates and is relatively long in length. Then, the number of points  $K$  is required to be large enough such that the minimization procedure can find out the long least action path. Besides, we find in the numerical experiment that if  $T$  is too large, the least action path may pass an additional saddle point (limit cycle) with higher potential energy before going through the expected saddle point. Special care is needed to choose suitable  $K$  and  $T$  in these cases, and the ODE-based-adaptive-time method proposed in Discussion is a candidate to improve the efficiency. There are also numerical methods to adjust the grid points on the least action path, such as the adaptive minimum action method<sup>72</sup>. They can be applied to optimize our current numerical code.

## References

- Kramers, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, 284–304 (1940).
- Hänggi, P., Talkner, P. & Borkovec, M. Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* **62**, 251–341 (1990).
- Eyring, H. The activated complex in chemical reactions. *J. Chem. Phys.* **3**, 107–115 (1935).
- McAdams, H. H. & Arkin, A. Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA* **94**, 814–819 (1997).
- Zhu, X., Yin, L., Hood, L. & Ao, P. Calculating biological behaviors of epigenetic states in the phage  $\lambda$  life cycle. *Funct. Integr. Genomics* **4**, 188–195 (2004).
- Wilkinson, D. J. Stochastic modelling for quantitative description of heterogeneous biological systems. *Nat. Rev. Genet.* **10**, 122–133 (2009).
- Bomze, Y., Hey, R., Grahm, H. T. & Teitworth, S. W. Noise-induced current switching in semiconductor superlattices: Observation of nonexponential kinetics in a high-dimensional system. *Phys. Rev. Lett.* **109**, 026801 (2012).
- Parker, M., Kamenev, A. & Meerson, B. Noise-induced stabilization in population dynamics. *Phys. Rev. Lett.* **107**, 180603 (2011).
- Khasin, M. & Dykman, M. I. Extinction rate fragility in population dynamics. *Phys. Rev. Lett.* **103**, 068101 (2009).
- Wang, G., Zhu, X., Gu, J. & Ao, P. Quantitative implementation of the endogenous molecular-cellular network hypothesis in hepatocellular carcinoma. *Interface Focus* **4**, 20130064 (2014).
- Zhu, X., Yuan, R., Hood, L. & Ao, P. Endogenous molecular-cellular hierarchical modeling of prostate carcinogenesis uncovers robust structure. *Prog. Biophys. Mol. Biol.* **117**, 30–42 (2015).
- Wright, S. *The Roles of Mutation, Inbreeding, Crossbreeding, and Selection in Evolution*, vol. 1 (Proceedings of the Sixth International Congress of Genetics, 1932).
- Waddington, C. H. *The Strategy of the Genes*, vol. 20 (Allen and Unwin, London, 1957).
- Ao, P. Potential in stochastic differential equations: novel construction. *J. Phys. A* **37**, L25 (2004).

15. Wang, J., Zhang, K., Xu, L. & Wang, E. Quantifying the Waddington landscape and biological paths for development and differentiation. *Proc. Natl. Acad. Sci. USA* **108**, 8257–8262 (2011).
16. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).
17. Paulsson, J. Summing up the noise in gene networks. *Nature* **427**, 415–418 (2004).
18. Choi, P. J., Cai, L., Frieda, K. & Xie, X. S. A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science* **322**, 442–446 (2008).
19. Lei, X., Tian, W., Zhu, H., Chen, T. & Ao, P. Biological sources of intrinsic and extrinsic noise in cl expression of lysogenic phage lambda. *Sci. Rep.* **5** (2015).
20. Losick, R. & Desplan, C. Stochasticity and cell fate. *Science* **320**, 65–68 (2008).
21. Dar, R. D., Hosmane, N. N., Arkin, M. R., Siliciano, R. F. & Weinberger, L. S. Screening for noise in gene expression identifies drug synergies. *Science* **344**, 1392–1396 (2014).
22. Trendelkamp-Schroer, B. & Noé, F. Efficient estimation of rare-event kinetics. *Phys. Rev. X* **6**, 011009 (2016).
23. Tang, Y., Yuan, R., Chen, J. & Ao, P. Work relations connecting nonequilibrium steady states without detailed balance. *Phys. Rev. E* **91**, 042108 (2015).
24. Lu, M., Onuchic, J. & Ben-Jacob, E. Construction of an effective landscape for multistate genetic switches. *Phys. Rev. Lett.* **113**, 078102 (2014).
25. Freidlin, M. I. & Wentzell, A. D. *Random Perturbations of Dynamical Systems*, 3rd edn. (Springer-Verlag, Berlin, 2012).
26. Zhou, J. X., Aliyu, M., Aurell, E. & Huang, S. Quasi-potential landscape in complex multi-stable systems. *J. R. Soc. Interface* **9**, 3539–3553 (2012).
27. Lv, C., Li, X., Li, F. & Li, T. Constructing the energy landscape for genetic switching system driven by intrinsic noise. *PLoS one* **9**, e88167 (2014).
28. Wells, D. K., Kath, W. L. & Motter, A. E. Control of stochastic and induced switching in biophysical networks. *Phys. Rev. X* **5**, 031036 (2015).
29. Huang, S. Non-genetic heterogeneity of cells in development: more than just noise. *Development* **136**, 3853–3862 (2009).
30. Wang, G., Zhu, X., Hood, L. & Ao, P. From phage lambda to human cancer: endogenous molecular-cellular network hypothesis. *Quant. Biol.* **1**, 32–49 (2013).
31. Shi, J., Chen, T., Yuan, R., Yuan, B. & Ao, P. Relation of a new interpretation of stochastic differential equations to Itô process. *J. Stat. Phys.* **148**, 579–590 (2012).
32. Yuan, R. & Ao, P. Beyond Itô versus Stratonovich. *J. Stat. Mech.* **2012**, P07010 (2012).
33. Strogatz, S. H. *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering* (Westview Press, 2014).
34. Li, S., Zhu, X., Liu, B., Wang, G. & Ao, P. Endogenous molecular network reveals two mechanisms of heterogeneity within gastric cancer. *Oncotarget* (2015).
35. Gardiner, C. W. *Handbook of Stochastic Methods*, 3rd edn (Springer-Verlag, Berlin, 2004).
36. Barkal, N. & Leibler, S. Robustness in simple biochemical networks. *Nature* **387**, 913–917 (1997).
37. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
38. Tyson, J. J., Chen, K. C. & Novak, B. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell Biol.* **15**, 221–231 (2003).
39. Wang, G. *et al.* Endogenous network states predict gain or loss of functions for genetic mutations in hepatocellular carcinoma. *J. R. Soc. Interface* **13**, 20151115 (2016).
40. Gillespie, D. T. The chemical Langevin equation. *J. Chem. Phys.* **113**, 297–306 (2000).
41. Das, J. *et al.* Digital signaling and hysteresis characterize ras activation in lymphoid cells. *Cell* **136**, 337–351 (2009).
42. Horsthemke, W. & Lefever, R. *Noise-Induced Transitions: Theory and Applications in Physics, Chemistry, and Biology*, 2nd edn. (Springer-Verlag, Berlin, 2006).
43. Sagués, F., Sancho, J. M. & Garca-Ojalvo, J. Spatiotemporal order out of noise. *Rev. Mod. Phys.* **79**, 829–882 (2007).
44. Assaf, M., Roberts, E., Luthy-Schulten, Z. & Goldenfeld, N. Extrinsic noise driven phenotype switching in a self-regulating gene. *Phys. Rev. Lett.* **111**, 058102 (2013).
45. Tang, Y., Yuan, R., Chen, J. & Ao, P. Controlling symmetry-breaking states by a hidden quantity in multiplicative noise. *Phys. Rev. E* **90**, 052121 (2014).
46. Tang, Y., Yuan, R. & Ma, Y. Dynamical behaviors determined by the Lyapunov function in competitive Lotka-Volterra systems. *Phys. Rev. E* **87**, 012708 (2013).
47. Yuan, R., Wang, X., Ma, Y., Yuan, B. & Ao, P. Exploring a noisy van der Pol type oscillator with a stochastic approach. *Phys. Rev. E* **87**, 062109 (2013).
48. Ma, Y., Tan, Q., Yuan, R., Yuan, B. & Ao, P. Potential function in a continuous dissipative chaotic system: Decomposition scheme and role of strange attractor. *Int. J. Bifurcat. Chaos* **24**, 1450015 (2014).
49. Tang, Y., Yuan, R. & Ao, P. Summing over trajectories of stochastic dynamics with multiplicative noise. *J. Chem. Phys.* **141**, 044125 (2014).
50. Qian, H. The zeroth law of thermodynamics and volume-preserving conservative system in equilibrium with stochastic damping. *Phys. Lett. A* **378**, 609–616 (2014).
51. Maier, R. S. & Stein, D. L. Escape problem for irreversible systems. *Phys. Rev. E* **48**, 931–938 (1993).
52. Ascher, U. M. & Petzold, L. R. *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*, vol. 61 (SIAM, Philadelphia, 1998).
53. Ge, H. & Qian, H. Landscapes of non-gradient dynamics without detailed balance: Stable limit cycles and multiple attractors. *Chaos* **22**, 023140 (2012).
54. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *cell* **144**, 646–674 (2011).
55. Junttila, M. R. & de Sauvage, F. J. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature* **501**, 346–354 (2013).
56. Yuan, R., Zhu, X., Wang, G., Li, S. & Ao, P. Cancer as robust intrinsic state shaped by evolution: a key issues review. *Rep. Prog. Phys.* **80**, 042701 (2017).
57. MacArthur, B. D., Ma'ayan, A. & Lemischka, I. R. Systems biology of stem cell fate and cellular reprogramming. *Nat. Rev. Mol. Cell Biol.* **10**, 672–681 (2009).
58. Wust, P. *et al.* Hyperthermia in combined treatment of cancer. *Lancet Oncol.* **3**, 487–497 (2002).
59. Issels, R. D. Hyperthermia adds to chemotherapy. *Eur. J. Cancer* **44**, 2546–2554 (2008).
60. Hildebrandt, B. *et al.* The cellular and molecular basis of hyperthermia. *Crit. Rev. Oncol. Hematol.* **43**, 33–56 (2002).
61. Lançon, P., Batrouni, G., Lobry, L. & Ostrowsky, N. Drift without flux: Brownian walker with a space-dependent diffusion coefficient. *Europhys. Lett.* **54**, 28 (2001).
62. Volpe, G., Helden, L., Brettschneider, T., Wehr, J. & Bechinger, C. Influence of noise on force measurements. *Phys. Rev. Lett.* **104**, 170602 (2010).
63. Raser, J. M. & O'Shea, E. K. Noise in gene expression: origins, consequences, and control. *Science* **309**, 2010–2013 (2005).
64. Süel, G. M., Garcia-Ojalvo, J., Liberman, L. M. & Elowitz, M. B. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* **440**, 545–550 (2006).

65. Pedraza, J. M. & Paulsson, J. Effects of molecular memory and bursting on fluctuations in gene expression. *Science* **319**, 339–343 (2008).
66. Eldar, A. & Elowitz, M. B. Functional roles for noise in genetic circuits. *Nature* **467**, 167–173 (2010).
67. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361 (1977).
68. Tang, Y., Yuan, R. & Ao, P. Anomalous free energy changes induced by topology. *Phys. Rev. E* **92**, 062129 (2015).
69. Tang, Y., Yuan, R. & Ao, P. Nonequilibrium work relation beyond the Boltzmann–Gibbs distribution. *Phys. Rev. E* **89**, 062112 (2014).
70. Vaccario, G., Antoine, C. & Talbot, J. First-passage times in  $d$ -dimensional heterogeneous media. *Phys. Rev. Lett.* **115**, 240601 (2015).
71. Zhou, P. & Li, T. Construction of the landscape for multi-stable systems: Potential landscape, quasi-potential, A-type integral and beyond. *J. Chem. Phys.* **144**, 094109 (2016).
72. Zhou, X., Ren, W. & E, W. Adaptive minimum action method for the study of rare events. *J. Chem. Phys.* **128**, 104111 (2008).

## Acknowledgements

We thank Hao Ge, Hong Qian, Nigel Goldenfeld, Xue Lei and Xiaojie Qiu for the critical comments. Discussions with Lei Zhang, Tiejun Li, Jianhua Xing, Joseph Zhou, and Site Li are greatly acknowledged. This work was supported in part by the Natural Science Foundation of China No. NSFC91329301 and No. NSFC91529306; and by the grants from the State Key Laboratory of Oncogenes and Related Genes (No. 90-10-11).

## Author Contributions

Y.T., R.Y., G.W. and P.A. had the original idea for this work. Y.T. and R.Y. did the theoretical analysis. R.Y. and Y.T. conducted the numerical implementation, and wrote the MATLAB code. Y.T., R.Y., G.W. and X.Z. analyzed the biological examples. All authors discussed the results and contributed to writing the article.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-017-15889-2>.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017