Towards a critical transition theory under different temporal scales and noise strengths

Jifan Shi and Tiejun Li*
LMAM and School of Mathematical Sciences, Peking University, Beijing 100871, China

Luonan Chen†
Key Laboratory of Systems Biology, Innovation Center for Cell Signaling Network, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, and Chinese Academy of Sciences, Shanghai 200031, China; Institute of Industrial Science, University of Tokyo, Tokyo 153-8505, Japan

(Received 22 October 2015; revised manuscript received 3 January 2016; published 21 March 2016)

The mechanism of critical phenomena or critical transitions has been recently studied from various aspects, in particular considering slow parameter change and small noise. In this article, we systematically classify critical transitions into three types based on temporal scales and noise strengths of dynamical systems. Specifically, the classification is made by comparing three important time scales τ_λ, τ_{tran}, and τ_{ergo}, where τ_λ is the time scale of parameter change (e.g., the change of environment), τ_{tran} is the time scale when a particle or state transits from a metastable state into another, and τ_{ergo} is the time scale when the system becomes ergodic. According to the time scales, we classify the critical transition behaviors as three types, i.e., state transition, basin transition, and distribution transition. Moreover, for each type of transition, there are two cases, i.e., single-trajectory transition and multitrajectory ensemble transition, which correspond to the transition of individual behavior and population behavior, respectively. We also define the critical point for each type of critical transition, derive several properties, and further propose the indicators for predicting critical transitions with numerical simulations. In addition, we show that the noise-to-signal ratio is effective to make the classification of critical transitions for real systems.

DOI: 10.1103/PhysRevE.93.032137

I. INTRODUCTION

Critical transitions of nonlinear systems long have been studied in various fields, including earthquake, disease outbreak, stock-market crash, and so on [1–16]. In particular, the topic to predict the critical point has attracted scientists for hundreds of years. In the ancient time, people used the life experiences to make predictions but with limited accuracy due to lack of theoretical support. Even up to recent times, few precise indicators for real problems have been found, although there are many mathematical models established to characterize the critical phenomena.

In the 2000s, one simple but effective prediction index, i.e., "critical slowing down" phenomenon, was proposed by Scheffer *et al.* based on the bifurcation theory [2,3]. They found that the system would exhibit a drastically slowing down of relaxation time and a significant increasing of variance when the system is close to the bifurcation point from a steady state. Since then, much research has been devoted to the development of similar models and many statistical indicators have been proposed [4,5]. These theories and models have been widely applied to ecology [6–8], meteorology [9–12,15], economics [13,14], and so on. In addition, Cavalcante *et al.* named the extreme events in chaotic systems "Dragon kings" and considered their predictability and controllability [17]. A rigorous and detailed mathematical description was given by Kuehn in 2011 [18]. Recently, an extension of Scheffer's model to multivariable systems or complex networks was established by Chen *et al.* in 2012 [19–24]. Their dynamical network biomarker (DNB) showed early-warning signals effectively by exploring high-throughput data of gene sequences for biological and medical systems, which opens a door to the prediction of complex disease even in the predisease stage. Further extension to single sample statistics was also presented [25].

However, most of those works mentioned above assume sufficiently small noise perturbations, and thus the critical transition always happens near the bifurcation point. On the other hand, it is still an open question how to effectively characterize the critical transition to detect its early warning signals for a system with moderate or large noise [16]. Thompson *et al.* found the early and delayed escape under different drift and noise with first order autoregression [26]. Dakos *et al.* made a step forward in this direction by discussing the "flickering" phenomenon [27]. They compared its statistical behaviors with the small noise case and found that almost all indicators could hardly be used anymore for large noise cases.

There are a number of other studies related to the stochastic systems with nonsmall noise transitions. Horsthemke and Lefever established a noise-induced transition theory in the one-dimensional case [28]. They focused on the maximal points of the probability density function (pdf) and set the transition point as a state where a sudden change of their number (for maximal points) occurred. Arnold defined a P-bifurcation point where the distribution changed from unimodal to bimodal [29]. The transition path theory (TPT) was developed to study the statistical properties of transition paths as an ensemble [30]. Burglund *et al.* discussed noise-induced phenomena under different scales of Kramers time [31]. In addition, there are also theoretical works on nonequilibrium phase transition, which were used to analyze the qualitative change of density function [32–34]. These works motivate us to consider the critical transitions in moderate and large noise

*tieli@math.pku.edu.cn
†lnchen@sibs.ac.cn

cases with a general mathematical framework and study their dynamical properties before the transitions.

In this paper, we aim to build a general mathematical framework for understanding critical transitions with the consideration of both small noise and large noise. To characterize the critical transitions, we will show that there are three key time scales, i.e., (a) the time scale of parameter change denoted as τ_λ, (b) the time scale when a particle (or state of the system) transits out of a metastable basin denoted as τ_{tran}, and (c) the time scale when the system becomes ergodic denoted as τ_{ergo}. Clearly, the traditional theory based on small noise, e.g., critical slowing down, is just a special case when $\tau_\lambda \ll \tau_{\text{tran}} \lesssim \tau_{\text{ergo}}$, which is called the *state transition* in this paper. For such a case, the critical transition happens at or near the bifurcation point, i.e., the attraction basin of the stable equilibrium becomes sufficiently small. When the noise is not sufficiently small, generally there are at least two other important types of transition in mathematics. One type is the *basin transition*, which is a particle transiting out of an attraction basin, and it generally happens not at but *before* the bifurcation point, i.e., the attraction basin of the stable equilibrium is not sufficiently small. This corresponds to the regime $\tau_{\text{tran}} < \tau_\lambda < \tau_{\text{ergo}}$ of the system. Another type corresponds to the regime of the system $\tau_{\text{tran}} \lesssim \tau_{\text{ergo}} \ll \tau_\lambda$. We call it the *distribution transition*, which may happen far before the bifurcation point due to the ergodic state of the system. For such dynamics, what can be observed is just the point cloud of trajectories or probability distribution of the system. Based on such time scales, we identify the features of critical transitions and further propose different indicators to predict the critical point for each case. In terms of dynamical behaviors, there are an additional two different ways of transitions, i.e., single-trajectory transition and multitrajectory ensemble transition, which correspond to the transition of individual behavior and population behavior, respectively. We note that a related idea on discussing the time-scale issues in mesoscopic dynamics was also proposed in Ref. [35].

We should also remark that the purpose of this paper does not intend to solve the problem of the prediction for critical transitions but aims to present the theoretical framework with a minimal model to characterize the critical transitions for both small and large noises. However, as we can see, the considered minimal model shows essential features of general stochastic dynamics for critical transitions. Note that before the critical transition for case (a) or case (b) from one stable equilibrium, there is no information on the dynamics of the transited stable equilibrium, but for case (c), the information on the dynamics of the transited stable equilibrium is considered to be available due to the ergodic condition. In such a sense, the critical transition for case (b) can be considered as a conditional distribution transition, in contrast to the distribution transition for case (c), while the state transition for case (a) can be also considered as one extreme case of the conditional distribution transition.

The structure of this paper is as follows. In Sec. II, we present a minimal model and its generalization. Then we define the three time scales associated with the model. In Sec. III, we classify the systems into three regimes. In each regime, we investigate the critical transition behavior and study the properties of the critical point. We also discuss the differences between the single-trajectory and multitrajectory samplings. In Sec. IV, we list some indicators or properties useful in predicting the critical points. Simulation results are also presented. We show that the noise-to-signal ratio (NSR) can discriminate the regimes of the stochastic systems for classifying the transitions. Finally, we make the conclusion with the discussion on some omitted but interesting topics in Sec. V.

II. MODEL

Critical transition is a complicated behavior, governed by stochastic dynamics. In this section, we first propose a minimal model, which is a one-dimensional stochastic differential equation (SDE) with a gradient force and additive noise, to describe such behaviors. We then discuss its extension to general cases with multiple variables. In the end, we take a toy model having an analytical form for further discussions without loss of generality.

A. A one-dimensional model with additive noise

We start from a gradient SDE with changing parameters to model the critical phenomena in one dimension:

$$\begin{cases} dx_\tau = -\partial_x U(x_\tau,\lambda)d\tau + \sqrt{2\eta}\mathrm{d}w_\tau, \\ \frac{d\lambda}{d\tau} = k. \end{cases} \tag{1}$$

In Eq. (1), x_τ is the position of a particle or a state of the system at time τ. $U(x,\lambda)$ is a potential function changing with the parameter λ. With the change of the parameter, there is a bifurcation which qualitatively changes the dynamics. Actually, the parameter k adjusts the difference of time scales between x and λ. The smaller k is, the slower λ varies than x. w_τ is a standard Brownian motion with mean zero and covariance $\langle w_\tau w_{\tau'}\rangle = \min(\tau,\tau')$. And η denotes the noise strength. Here we employ the convention in probability theory to use dw_τ instead of the white noise $dw_\tau/d\tau$ because of the irregularity of Brownian paths. The system (1) is not an almighty model for describing the critical transitions, but it can exhibit some essential features for general dynamical systems.

Equation (1) is a system with one variable x, two parameters τ and λ with three degrees of freedom, U, k, and η. If we set our observation time, i.e., the whole changing period of λ as $O(1)$, then we claim that there are three typical time scales in the dynamics expressed by Eq. (1):

(i) τ_λ, the time scale of parameter λ.

(ii) τ_{tran}, the waiting time when a particle (or state of the system) transits across the boundary of a stable metastable well (or the boundary of the attracting basin for a stable equilibrium) into another.

(iii) τ_{ergo}, the time scale when the system becomes ergodic.

It is easy to see that the order of τ_λ is $O(1/k)$ in Eq. (1). For τ_{tran}, we have an approximate expression in special situations. When $k = 0$ and η is sufficiently small, we have $\log(\tau_{\text{tran}}) \sim O(h_U/\eta)$, where h_U is the energy barrier height of the potential well $U(x,\lambda)$ at a fixed λ. τ_{ergo} is one order higher than the maximum τ_{tran} between all metastable states. If we can get the largest negative eigenvalue λ_1 (with the least magnitude) of the forward operator induced by Eq. (1), then τ_{ergo} is in the order of $O(1/|\lambda_1|)$ (cf. Chapter 5 of Ref. [36]).

To set our observation time for λ to be $O(1)$, we rescale Eq. (1) as

$$\begin{cases} dx_t = -\partial_x V(x_t,\lambda)dt + \sqrt{2\sigma}dw_t, \\ \frac{d\lambda}{dt} = 1, \end{cases} \tag{2}$$

by letting $t = k\tau$, $V(x,\lambda) = U(x,\lambda)/k$, and $\sigma = \eta/k$. Thus λ will change with unit rate and $\tau_\lambda = O(1)$. With this manipulation, we reduce the number of degrees of freedom of the system (1) from three into two in (2): V and σ. For every V and σ, there is a simple correspondence for U, k, and η.

For the three time scales τ_λ, τ_{tran}, and τ_{ergo}, though in general the last two can hardly have analytic expressions, it is obvious that the time rescaling from Eq. (1) to Eq. (2) will not change their ranking. Moreover, τ_{ergo} is always larger than τ_{tran}. Now with a fixed $V(x,\lambda)$, it is σ that determines the ranking of τ_λ, τ_{tran}, and τ_{ergo}. Concerning the primitive parameters in Eq. (1) corresponding to the same $V(x,\lambda)$, smaller temporal parameter k, which means slower environmental change, or larger noise strength η, will lead to larger effective noise amplitude σ.

In Sec. III, we will classify critical phenomena into three types with respect to different rankings of τ_λ, τ_{tran}, and τ_{ergo}. The ranking is mainly determined by noise amplitude σ, which is a combination of the primitive parameters k and η.

B. A general multidimensional model with multiplicative noise

The model in the previous subsection is a simplest minimal model, which has limitations in some sense. More often, we can consider a general dynamical system as:

$$\begin{cases} d\boldsymbol{X}_t = \boldsymbol{f}(\boldsymbol{X}_t,\boldsymbol{p})dt + \sqrt{2\sigma}\boldsymbol{g}(\boldsymbol{X}_t,\boldsymbol{p})d\boldsymbol{W}_t, \\ \frac{d\boldsymbol{p}}{dt} = \boldsymbol{h}(t), \end{cases} \tag{3}$$

where the coordinates $\boldsymbol{X}_t \in \mathbb{R}^n$, parameter $\boldsymbol{p} \in \mathbb{R}^m$, and $\boldsymbol{W}_t$ is a n-dimensional Brownian motion with independent components. We choose the positive functions $\boldsymbol{g}(\boldsymbol{x},\boldsymbol{p}),\boldsymbol{h}(t) \sim O(1)$. We suppose that dynamics of a nonlinear system are around a stable equilibrium initially with the gradual change of parameters, and we only consider codimension one bifurcations through the interactions of $\boldsymbol{f}(\boldsymbol{x},\boldsymbol{p})$, $\boldsymbol{g}(\boldsymbol{x},\boldsymbol{p})$, and σ without loss of generality.

To discuss the relation with our minimal model, let us give some suitable remarks on Eq. (3). First, with a given initial condition, the parameter $\boldsymbol{p}$ determines a simple curve in $\mathbb{R}^m$. Using the arclength as the natural coordinate of the curve, the parameter equation can be reduced into $d\lambda/dt = s(t)$. Thus Eq. (3) can be transformed into

$$\begin{cases} d\boldsymbol{X}_t = \boldsymbol{f}(\boldsymbol{X}_t,\lambda)dt + \sqrt{2\sigma}\boldsymbol{g}(\boldsymbol{X}_t,\lambda)d\boldsymbol{W}_t, \\ \frac{d\lambda}{dt} = s(t). \end{cases} \tag{4}$$

Second, in a high-dimensional system, $\boldsymbol{f}(\boldsymbol{x},\lambda)$ may not be written in a gradient form (in a one-dimensional system, we can always transform it into a gradient form), and the effect of $\boldsymbol{g}(\boldsymbol{x},\lambda)$ is not trivial. For Eq. (4), we can define the three time scales τ_λ, τ_{tran}, and τ_{ergo} similar to those in Eq. (1) of Sec. II A. In a gradient system, τ_{tran} is the time when a particle transits across the barrier of potential well, but in a general system, τ_{tran} is the time when the particle or the state crosses over the attracting basin's boundary of the stable equilibrium.

Third, we take the viewpoint of "effective potential" in a nonequilibrium phase transition [28,32] or "quasipotential" in large derivations [37,38]. We utilize an "ideal potential," which has been mentioned without a name in Ref. [27], defined as

$$V(\boldsymbol{x},\lambda) = -\sigma \ln p_{\text{st}}(\boldsymbol{x},\lambda). \tag{5}$$

Here $p_{\text{st}}(\boldsymbol{x},\lambda)$ is the steady-state pdf induced by Eq. (4) with a fixed λ. As the system Eq. (4) has only stable equilibrium points and codimension one bifurcations, we claim that the boundary of ideal potential V can grasp the boundary of the attracting basin after sufficiently long time. Though τ_λ, τ_{tran}, and τ_{ergo} can no longer be estimated through V, the ideal potential still gives us an intuition on the scope of attracting basins.

Clearly, although we know Eq. (3) is much more complicated than the minimal model Eq. (2), with the previous discussion, we can do the same classification to Eq. (3) or Eq. (4) according to the different rankings of τ_λ, τ_{tran}, and τ_{ergo}. Thus the later discussions on our minimal model will shed light on the investigation for general models, and our study on the simplest system (2) presents the necessary steps towards the understanding of more complicated critical transitions.

C. Toy model for simulation

In the classification and simulations, we choose the potential function as

$$U(x,\lambda) = \tfrac{1}{4}x^4 - \tfrac{3}{2}x^2 + \lambda x, \tag{6}$$

which is unimodal or bimodal at different λ values. We take $k = 0.01$ and vary η (or σ). Hence, we have the corresponding

$$V(x,\lambda) = 100\left(\tfrac{1}{4}x^4 - \tfrac{3}{2}x^2 + \lambda x\right). \tag{7}$$

To study the critical transitions, we take the dynamics in Eq. (2) with Eq. (7) and choose $\lambda(0) = -3$, $x(0) \approx 2.1038$, which corresponds to the unique stable fixed point of potential $V(x,\lambda(0))$. The terminal time is set at $t = 6$, i.e., $\lambda = 3$. When doing the simulations, we use the Euler-Maruyama scheme

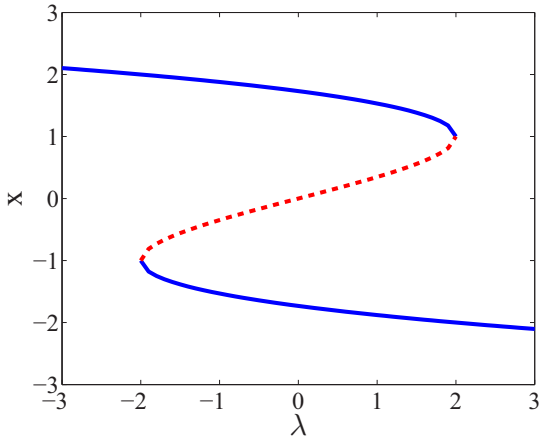

FIG. 1. Bifurcation diagram of the dynamics in Eq. (2) with $V(x,\lambda)$ in Eq. (7). The stable equilibrium points are marked by a solid blue line and the unstable equilibrium points are marked by a dashed red line. At $\lambda = -2$, the bistable state occurs, and at $\lambda = 2$ one metastable well disappears.

and choose a sufficiently small step size $\Delta t = 10^{-6}$ as the default. The bifurcation diagram of $V(x,\lambda)$ is shown in Fig. 1 with its stable (solid blue line) and unstable (dashed red line) equilibrium points marked.

III. CLASSIFICATION OF CRITICAL TRANSITIONS IN DIFFERENT SCALES

A. Classification

According to the ranking and size of the typical time scales induced by the system (2), we classify the critical transitions into three cases.

Case I: state transition, for $\tau_\lambda \ll \tau_{\text{tran}} \lesssim \tau_{\text{ergo}}$, usually under relatively fast temporal parameter k or small parameter η, i.e., small noise amplitude σ.

Case II: basin transition, for $\tau_{\text{tran}} < \tau_\lambda < \tau_{\text{ergo}}$, usually under relatively medium temporal parameter k and parameter η, i.e., medium noise amplitude σ. It can be viewed as a conditional distribution transition.

Case III: distribution transition, for $\tau_{\text{tran}} \lesssim \tau_{\text{ergo}} \ll \tau_\lambda$, usually under relatively slow temporal parameter k or large parameter η, i.e., large noise amplitude σ.

We know that the ranking of τ_λ, τ_{tran}, and τ_{ergo} is only determined by σ with the same V in Eq. (2). We thus also call Case I the small noise case, Case II the medium noise case, and Case III the large noise case. The main results of this paper are concluded in Table I, including the classifications and the features of each case. We will discuss the details in the next subsection.

Horsthemke and Lefever [28] mainly focused on the distribution transition, which is Case III in our classification. And they especially discussed the one-dimensional model with the qualitative change of extrema of the stationary distribution as a signal. We will extend it to a high-dimensional space and use KL divergence (Kullback-Leibler divergence) as a new indicator. Berglund and Gentz [31] used Kramers time to classify the transitions but without τ_{ergo}. They have discussed stochastic resonance where drift term $\boldsymbol{f}$ is periodic in time instead of our medium noise case (Case II), which is much common in transition phenomena. In particular, as far as we know, the difference between a single-trajectory sample and a multitrajectory sample has never discussed before. After all, our classification is clear and general, and we will also derive indicators as early-warning signals shown in Sec. IV.

B. Discussion on models and classification

Now we discuss in detail how the classification is done and what the definition of the critical transition is in each case. As mentioned previously, we will classify the dynamics described by Eq. (2) into three types or cases: state transition, basin transition, and distribution transition. Simulations are based on the dynamics Eqs. (2) and (7).

1. State transition under small noise

Case I occurs when $\tau_\lambda \ll \tau_{\text{tran}} \lesssim \tau_{\text{ergo}}$, which means that σ is much smaller than the barrier height of the potential and the particle or the state of the system $x(t)$ almost always stays in a close neighborhood of the stable equilibrium point. Thus during the observation, we can hardly see the transition until the state becomes unstable.

In this case, we define the critical point as the bifurcation point, where a metastable well disappears and the particle has to transit into another. This definition has no ambiguity between the multitrajectory and single-trajectory statistics, because all paths or trajectories make the sudden change at the same instant. In the sufficiently small noise case, many criteria of early warning signals for predicting the critical transitions are proposed, like the increasing of variance and autocorrelation and so on. "Critical slowing down" is the basis of these statistics [2,4] for one-dimensional systems, and "strong fluctuation and correlation" is the signal for complex networks or multidimensional systems [19].

Transition path of a single particle is illustrated in Fig. 2(a), which is simulated by a Euler-Maruyama scheme for dynamical Eqs. (2) and (7) under small noise $\sigma = 0.125$. The bifurcation point is at $\lambda = 2$, which is also the critical point in this case. We can observe a sudden transition of the position of the state, clearly. In probability space for a multitrajectory ensemble, we can compute the first jump pdf, which is almost

TABLE I. Classification of critical transitions and features of each case.

	Case I	Case II	Case III
Transition type	*State transition*	*Basin transition*	*Distribution transition*
Time scales	$\tau_\lambda \ll \tau_{\text{tran}} \lesssim \tau_{\text{ergo}}$	$\tau_{\text{tran}} < \tau_\lambda < \tau_{\text{ergo}}$	$\tau_{\text{tran}} \lesssim \tau_{\text{ergo}} \ll \tau_\lambda$
Noise amplitude σ	Small	Medium	Big
Temporal parameter k	Fast	Medium	Slow
Primitive noise η	Small	Medium	Large
Single-trajectory critical point (λ^*)	Bifurcation point	Point when a particle crosses the boundary of an attracting basin for the stable equilibrium	Point when sudden or largest change of distribution (computed by time sequence in windows) occurs
Ensemble critical point (λ^*)	Bifurcation point	Point when most trajectories transit	Point when sudden or largest change of distribution (computed by trajectory ensemble at each time) occurs
Properties near critical point	Increase of variance, autocorrelation, correlation and so on	Large deviations from mean; noise and barrier height in the same order	Sudden change of some moments; increasing KL divergence between adjacent distributions

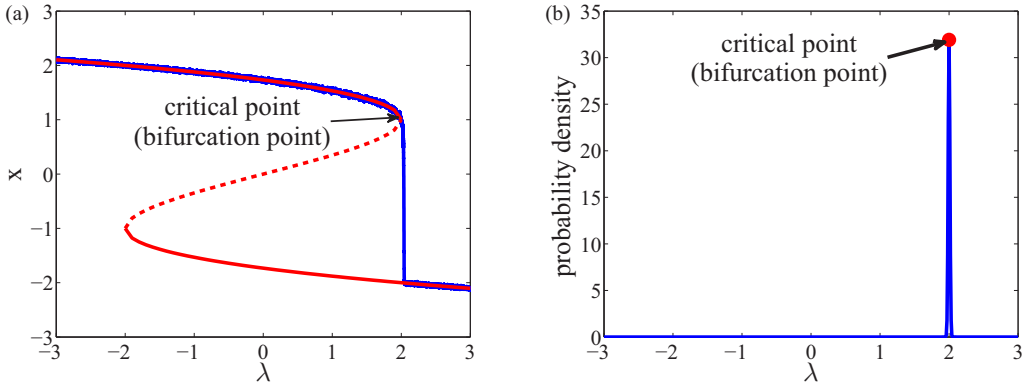

FIG. 2. State transition case under small noise. The dynamics is given by Eq. (2) with potential (7). Noise amplitude $\sigma = 0.125$ is relatively small to the potential barrier. Panel (a) is a path for a single particle under small noise and a transition can be seen. The solid red line is the stable equilibrium points and the dashed red line is the unstable equilibrium points. The trajectory of a state or particle is in blue. The critical point is defined as the bifurcation point $\lambda = 2$. Panel (b) shows the first jump pdf for multiple particles under small noise strength. In Fig. 2(b), the first jump probability density is nearly a δ function at $\lambda = 2$, which is also the bifurcation point. From both Figs. 2(a) and 2(b), we know there is no difference between the multiple- and single-trajectory statistics in this small noise case.

a δ function at $\lambda = 2$ and is shown in Fig. 2(b). The evolution of the probability density of x is also presented in Fig. 3.

2. Basin transition under medium noise

Case II occurs when $\tau_{\text{tran}} < \tau_\lambda < \tau_{\text{ergo}}$ with a relatively medium noise size. In this case we can observe the particle transiting at some place before the system bifurcates, but in general it will never get back. One realization of the path is explicitly shown in Fig. 4(a).

We want to emphasize that every trajectory has a quite different escape time in the medium noise case. Therefore, we distinguish two concepts: the single-trajectory transition and multitrajectory ensemble transition. For the single-trajectory transition, we define the critical point as the time when the particle crosses the boundary of an attraction basin, i.e., the separatrix between two neighboring attraction basins. For the multiple-trajectory transition, we define the critical point as the most probable transition point, at which the particles leave the attraction basin of the metastable state with the maximal escaping rate. We call it the *ensemble transition point*. Though the total transition probability saturates until the bifurcation point, the particles hardly stay in the initial metastable well eventually.

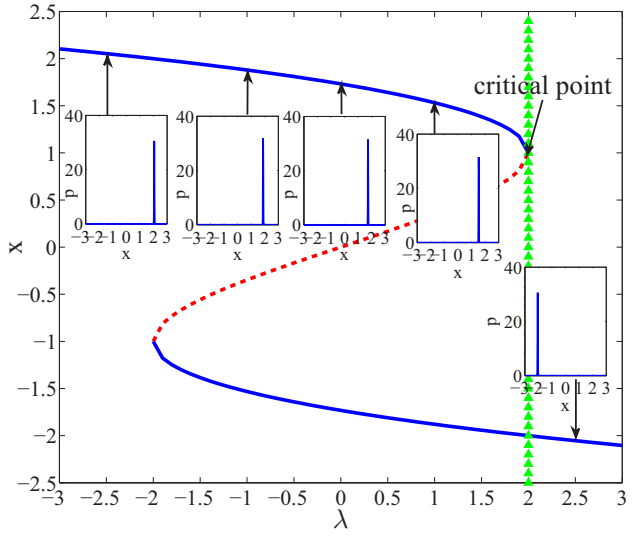

FIG. 3. The evolution of pdf of x for the dynamics (2) and (7) with $\sigma = 0.125$ for the small noise case. The solid blue line is the stable equilibrium points and the dashed red line is the unstable equilibrium points. Insets are pdfs at $\lambda = -2.5, -1, 0, 1, 2.5$. If we plot all the pdfs at every λ, it will exhibit a sudden transition at $\lambda = 2$ (green triangle line).

Now let us list some simple properties of the system in this case.

(1) For a single trajectory, the particle follows the uphill path

$$\dot{x} = \nabla V(x, \lambda) \tag{8}$$

with a high probability before the transition (cf. Refs. [39,40]).

(2) For single-particle trajectories, the closer the particle gets to the boundary, the more seldom it stays there.

(3) For a multitrajectory ensemble, the most probable transition point corresponds to the parameter λ^* at which the height of the basin h is in the same order of the noise amplitude σ.

Simulation of a single-trajectory path is shown in Fig. 4(a) with $\sigma = 50$. We can see that the transition occurs far before the bifurcation point. For multiple trajectories, we simulate 10^5 paths to compute the probability density of the first jump at each parameter. The numerical pdf of the transition point is shown in Fig. 4(b). The most probable transition point is marked as the critical point for ensemble transition. We can observe that the transition time generally differs for the single trajectory and multiple trajectories. The evolution of probability density for x is plotted in Fig. 5.

3. Distribution transition under large noise

The third case is $\tau_{\text{tran}} \lesssim \tau_{\text{ergo}} \ll \tau_\lambda$, which means that during the observation the state of the system has transited over the whole reachable space many times. Here σ is relatively

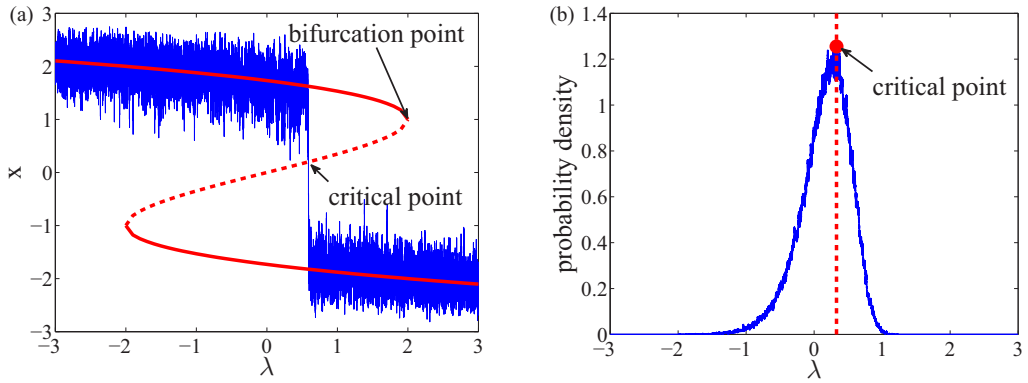

FIG. 4. Basin transition under medium noise. The dynamics is described by Eq. (2) with potential (7). We choose $\sigma = 50$, which is neither too small nor too large relative to potential barrier. A basin transition path for a single particle and the first jump pdf for multiple particles are shown in Figs. 4(a) and 4(b), respectively. In Fig. 4(a), the solid red line is the stable equilibrium points and dashed red line is the unstable equilibrium points. The trajectory of a single particle is in blue. For the single trajectory, the critical point is at which it crosses the boundary of a metastable well and moves to another metastable state. In Fig. 4(b), 10^5 trajectories are simulated and we record the first jump time for each. The maximum of the density function is defined to be the critical point of ensemble transition, which is marked red in the figure ($\lambda \approx 0.33$). The definitions of critical transitions differ for single-trajectory transition and ensemble transition.

larger than the two former cases. As τ_{ergo} is sufficiently small, the system reaches its equilibrium quickly at every λ.

Because the system is ergodic with a fast time scale, the single-path statistics in an $O(1)$ time interval and multitrajectory statistics in space make no difference. To define the critical point, we focus on the sudden qualitative change of the equilibrium distribution.

Some theories have been proposed on studying the qualitative change of distributions. The noise-induced transition theory utilizes the number and position of stable equilibrium points as indicators, which are also the local maxima of pdf [28]. The phase transition theory commonly uses an order parameter, which is obtained by mean-field approximation or other methods in lattice models [33,34].

However, using the number or position of equilibrium points to measure qualitative change has some shortcomings. For example, in Eq. (2) the equilibrium pdf can be given for each fixed λ as

$$p_{\text{st}}(x,\lambda) = \frac{1}{Z_\lambda} e^{-\frac{V(x,\lambda)}{\sigma}}, \tag{9}$$

where Z_λ is the normalization constant. A new stable equilibrium appears at $x = -1$ when $\lambda = -2$, but the probability around $x = -1$ is almost zero. Though the number of stable fixed points has a sudden change, there is no considerable change for the statistical quantities like the mean or variance. In practice, it is difficult to detect.

The order parameters utilized in phase transitions like the mean or variance are usually low-order approximations of moments to the density function. For characterizing the distribution transition, KL divergence (relative entropy) between two density functions is a good candidate of indicators, which extends the approximation of pdf to the full order in terms of moments.

We define the critical point in this case as the sudden change point of distributions, and we take the Fisher information metric (FIM) or KL divergence between two pdfs in consecutive time as indicators. The FIM or its KL divergence approximation as the indicator involves the information of probability density in whole order.

We plot a simulation path with $\sigma = 200$ in Fig. 6(a). In state space, it is difficult to observe any critical phenomenon. The first jump probability density is shown in Fig. 6(b). Since the system is ergodic, the first transition happens whenever the bistable state occurs. The particles or states have been already moving over the whole space, so the peak in Fig. 6(b) is not meaningful. However, in probability space, we can plot the evolution of the whole pdf at each parameter λ (see Fig. 7). Thus we can compute the Fisher information metric in theory

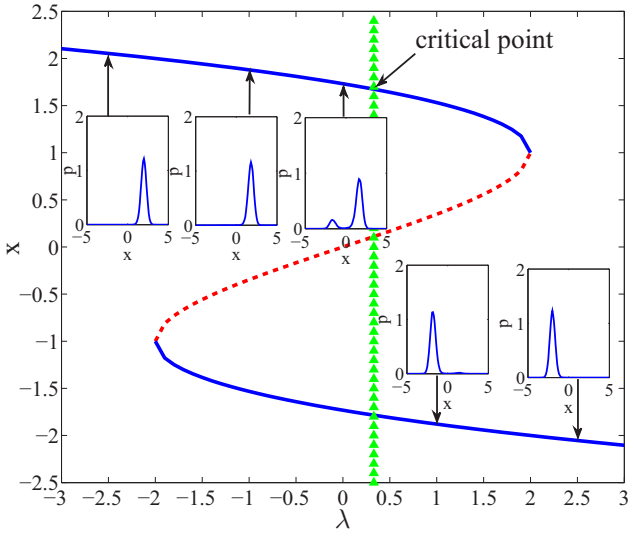

FIG. 5. The evolution of pdf of x with $\sigma = 50$ here for the medium noise case. The solid blue line is the stable equilibrium points and the dashed red line is the unstable equilibrium points. Insets are pdfs of x at $\lambda = -2.5, -1, 0, 1, 2.5$. By comparing the neighboring pdfs at every λ, they exhibit the largest change before $\lambda = 1$ but after $\lambda = 0$ (about 0.33 marked by the green triangle line), which corresponds to the ensemble transition point.

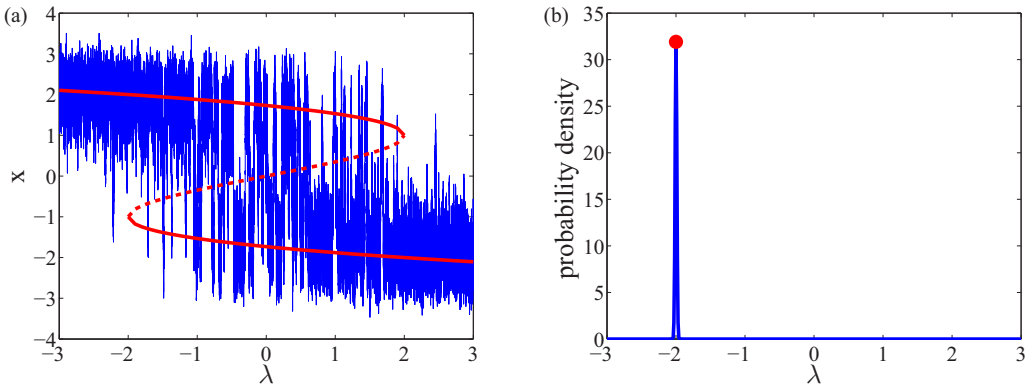

FIG. 6. Distribution transition under large noise. We choose $\sigma = 200$ which is relatively large to potential barrier. A single particle trajectory and the first jump pdf for multiple particles are shown in (a) and (b), respectively. In (a), the solid red line is the stable equilibrium points, the dashed red line is the unstable equilibrium points, and the blue line is a single trajectory of a particle. This trajectory is flickering all the time. In (b), the first jump probability density is nearly a δ function at $\lambda = -2$, where bistability occurs. This peak point marked red is not meaningful since the particle is always moving over the whole space even after that point.

and its KL divergence approximation by simulation. The result is shown in Fig. 8, and the peak at $\lambda = 0$ is set to be the critical point of distribution transition.

We call this regime distribution transition because if we let $\sigma \to 0$ and keep $\tau_{\text{ergo}} \ll \tau_\lambda$ in Eq. (2) [namely $k \to 0$ faster than η in Eq. (1)], we will have a double-peaked δ distribution centered at two metastable states when $\lambda = 0$, but have only one δ distribution for the other values of λ. In this case, the FIM diverges at the critical point $\lambda = 0$. On the other hand, if we let $\eta \to 0$ faster than k, it corresponds to the small noise case. A similar discussion can be referred to Ref. [35].

C. Extension to the general model

We have shown in Sec. II B that the general model Eq. (3) or Eq. (4) also has the time scales τ_λ, τ_{tran}, and τ_{ergo}. The main differences between Eqs. (4) and (2) are $\boldsymbol{x}$, which is a vector in high dimensions, and $s(t)$, which is inhomogeneous in time. In the case of $s(t) \sim O(1)$, we can perform a time rescaling $\tau(t) = \int_0^t s(u)du$ to make λ change with the unit rate during the observation. We then have the similar classification of transitions as the one-dimensional case.

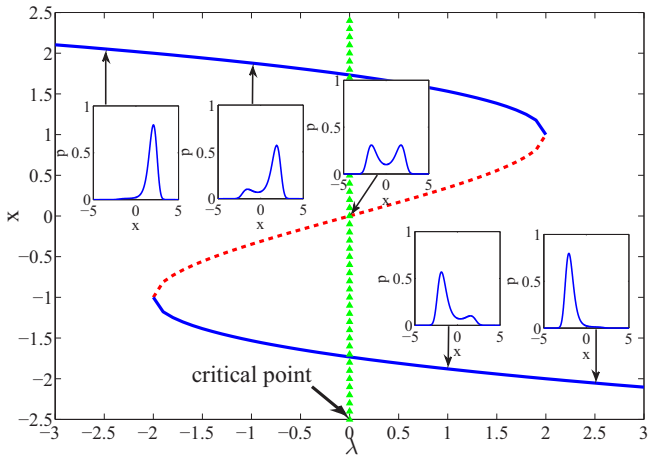

FIG. 7. The evolution of pdf with $\sigma = 200$ for the large noise case. The solid blue line is the stable equilibrium points and the dashed red line is the unstable equilibrium points. Insets are pdfs of x at $\lambda = -2.5, -1, 0, 1, 2.5$. The Fisher information metric for the pdfs in consecutive time gives the critical $\lambda = 0$ with the biggest change, which corresponds to the distribution transition point we defined (green triangle line).

1. State transition under small noise

In this case, σ is relatively small such that $\tau_\lambda \ll \tau_{\text{tran}} \lesssim \tau_{\text{ergo}}$. The critical transition point is defined as the bifurcation point. Differing from the one-dimensional case, $\boldsymbol{X}_t$ is now in $\mathbb{R}^n$. When we process the data, we should do clustering just as the procedure in DNB theory [19,20,25]. The basic assumption in DNB theory is that we can transform the system into eigenmodes, in which the principal eigenmode bifurcates. There are some but not all components related to the principal eigenmode. Their variances and autocorrelations will increase drastically when the system is close to critical transitions. Moreover, the correlations among these components increase fast while the correlations between this group and others decrease to zero.

2. Basin transition under medium noise

In this case, σ is in a medium level such that $\tau_{\text{tran}} < \tau_\lambda < \tau_{\text{ergo}}$. As discussed in the one-dimensional case, we have to distinguish the single-trajectory transition and multi-trajectory ensemble transition. For a single-particle trajectory, we set the critical point at which the particle transits out of the basin of metastable state. For ensemble transition, the transition point is defined as the parameter with which the system achieves the maximal escaping rate. The properties of the system near the critical point are just extensions of the one-dimensional case.

3. Distribution transition under large noise

In this case, σ is relatively large such that $\tau_{\text{tran}} \lesssim \tau_{\text{ergo}} \ll \tau_\lambda$. The system can reach its equilibrium fast and we get the states distributed as p_{st} at each λ. The ergodicity ensures us no statistical difference between a single trajectory and the ensemble of trajectories. As in the one-dimensional case, we

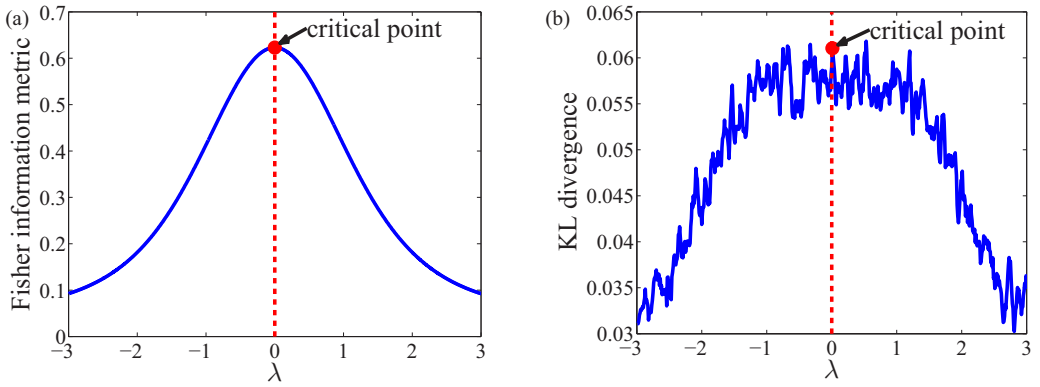

FIG. 8. (a) The theoretical Fisher information metric (FIM) at each λ and (b) KL divergence of adjacent distributions with discretization. The KL divergence is computed from the empirical distributions for each pair λ and $\lambda + \Delta\lambda$. $\Delta\lambda = 10^{-2}$ and 10^4 trajectories are simulated to estimate the distribution. We define the critical point as the peak of the curve, where the largest change of distribution occurs. The curve will get more peaked when σ is smaller, while ergodicity is preserved.

define the critical point at which the equilibrium pdf has a qualitative change or maximal change. We can quantify the change using Fisher information metric and KL divergence as an approximation. In high dimensions, the KL divergence or FIM is difficult to be numerically computed precisely by data. As we know, the KL divergence is affine invariant. We can first use principal component analysis [41] to obtain the principle components and then sum up the KL divergences in each principal component to make the estimation.

Clearly, we can only observe the dynamics around the original stable equilibrium before the state transition or basin transition, but we can observe the dynamics around both the original and the transited stable equilibria before the distribution transition due to the ergodic condition. Thus, the critical transition for the basin transition can be considered as a conditional distribution transition, in contrast to the distribution transition for the distribution transition, while the state transition can be also considered as one extreme case of the conditional distribution transition.

IV. INDICATORS AND PROPERTIES FOR EARLY WARNING SIGNALS

With the previous discussion on our minimal model and related extensions, we now aim to study how to make predictions on the critical transition from the available information in data.

A. Difference between a single trajectory and a multiple-trajectory ensemble

Before discussing the indicators, we first explain the differences between the single trajectory and multiple-trajectory ensemble.

The key difference between the two concepts is whether we focus on individual sudden change or parameter-induced group change. Intuitively, transition for a single trajectory is mainly an individual behavior, e.g., someone catches a cold. While the ensemble transition for multiple trajectories is a group behavior usually caused by the change of parameter or environment, e.g., the breakout of influenza. The smaller the external noise is, the less the difference between the two behaviors is. When the noise tends to zero, the stochastic dynamics degenerates to the deterministic case like Case I. On the other hand, when the noise is sufficiently large such that the system becomes ergodic quickly, the statistical behavior in short time between these two is not remarkable and we are in Case III. Thus the difference matters only in Case II with medium noise.

The difference of the two behaviors is also embodied in processing the data: The statistical quantities can be computed at each instant for multiple trajectories, but the time windows are needed for single-trajectory data.

B. Indicators and properties in different cases

Based on the analysis of our minimal model, we propose several indicators in different cases. Though these indicators have different forms, they are actually based on similar mechanisms. We can divide those indicators into two classes: *signals in state space* and *signals in probability space*. Variance, correlation, and so on are among the former class, and the latter includes KL divergence and FIM, etc. In different cases, the difficulty in computing different indicators from the data varies dramatically.

1. State transition case

The prediction problem of early warning signal has been studied in vast literature in this case. The major mission is to detect the bifurcation point before the system bifurcates. Scheffer *et al.* presented an early work based on the critical slowing down behavior [2]. Chen's group extended it into multivariate network and proposed the "DNB" (dynamical network biomarker) indicator [19] or "DNM" (dynamical network marker) [24] by further considering correlations among the variables among a network. To make a long story short, the most useful indicators are the variance and correlation. Many other proposals can be found in Ref. [4].

Corresponding to the trajectory in Fig. 2, we compute its variance through a moving time window. We can observe in Fig. 9 that the variance of x increases drastically before the

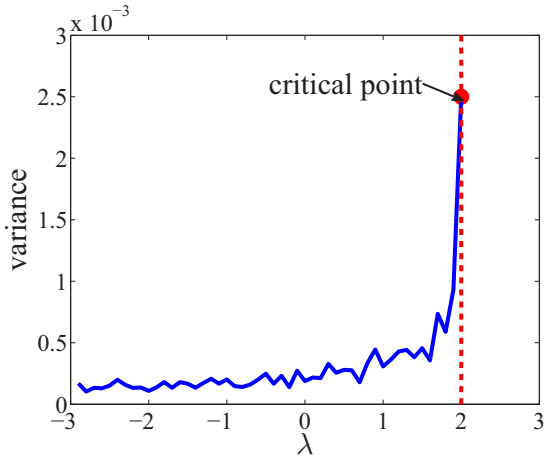

FIG. 9. Critical indicator—variance in Case I. The simulation result is computed by using single trajectory in moving time windows. We plot the curve (solid blue line) before the critical point (dashed red line). The dramatic increase of variance is clearly observed.

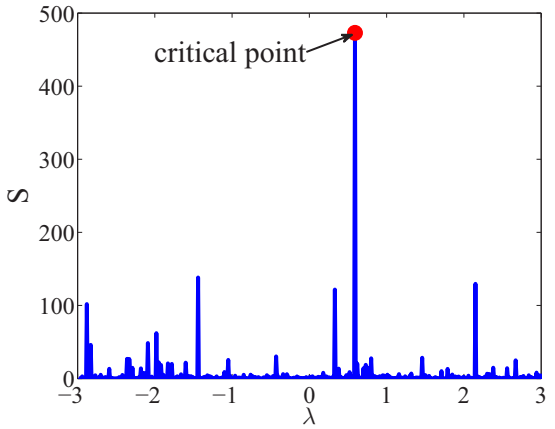

FIG. 10. Critical indicator—signal S [see Eq. (10)] in Case II. The simulation result is computed using the single-trajectory data by moving the time window. We take $N = 10^4$ and $\delta = 0.1$. The highest sharp peak (red point) is the transition point where the particle travels across the boundary of the attraction basin.

critical point, i.e., the bifurcation point. The result will be apparent if we compute via multiple trajectories.

2. Basin transition case

In this case, we distinguish the single trajectory and multiple-trajectory ensemble, in contrast to the state transition. The major mission is to detect when the single particle or many particles get close to the boundary of a metastable basin.

For the prediction of a single-trajectory critical transition, we propose the following indicator:

$$S(x,\delta,N) = \frac{|x - \bar{x}_N|}{p_N(x,\delta)}, \tag{10}$$

where x is the current particle position, $\bar{x}_N$ is the mean position by the previous N points, and $p_N(x,\delta)$ is the percentage of points in δ neighbor of x among the previous N points. The fact that $|x - \bar{x}_N|$ becomes larger and $p_N(x,\delta)$ becomes smaller near the transition point means that the particle is getting into a place seldom visited before. Hence, the larger S is, the closer the particle reaches the boundary of the attraction basin. Here δ and N will be chosen according to the considered problem.

For the trajectory obtained in Fig. 4(a), the indicator S is shown in Fig. 10. S increases indeed when approaching the transition point, and each peak in the figure represents an attempt to jump out.

For the ensemble transition, in which we define the critical point as the instant when the system achieves the maximal escaping rate, there seems no suitable indicators if only the trajectories of particles which do not cross the boundary are available. This is often the real case, especially in medical examples. Although there are no clear rules of dynamics, our simulations support that the critical transition point occurs where σ has the same order as the depth of the basin. We sampled 100 trajectories for each σ varied from 1 to 100 and then plotted h_V/σ for every σ in Fig. 11(a) where h_V is the barrier height when the trajectory transits across the boundary. In Fig. 11(b), we simulate the dynamics Eq. (1) with potential (6), where k varies from 10^{-3} to 10^2 and η varies from 10^{-6} to 10^2. The amount of 10^3 trajectories are simulated at each parameter to compute h_U/η, where h_U is the barrier height at the first transition point. The colored contour means that the first transition is located in $\lambda \in [-2,2]$, and the blank domain means that either the particle walks around the whole space before $\lambda = -2$ or it does not transit until $\lambda = 2$ (see bifurcation diagram in Fig. 1). The Fig. 11(a) is a slice of Fig. 11(b) which are marked by red line with square symbol. Both results in Fig. 11 show that the particles prefer to jump when the barrier height is close to the noise amplitude. These results are helpful for understanding the critical transition pattern and providing heuristics for constructing the indicators in the medium noise case.

3. Distribution transition case

As discussed in the previous section, we define the critical point in Case III as the instant in which the sudden change of distribution occurs. The FIM or KL divergence of adjacent distributions achieves maximum there. Hence, one natural indicator is FIM or KL divergence, or we can take moments as a cheap alternative. In the high-dimensional case, we take the principal component analysis and make the sum of KL divergences for the marginal distribution on each component.

We take a lattice model as an example to show the applicability of the proposed indicators. The considered dynamics has the following form [32,42,43]

$$dx_i = \left[f(x_i) + \frac{1}{2} g(x_i) g'(x_i) - \frac{D}{N} \sum_{j \in \mathcal{N}(i)} (x_i - x_j) \right] dt + \sigma g(x_i) dw_i, \tag{11}$$

where $\mathcal{N}(i)$ is the neighbors of site i, N is the number of nearest neighbors, σ is the amplitude of noise, D controls the strength of spatial interaction, and $w_i(t)$ are independent standard Brownian motions.

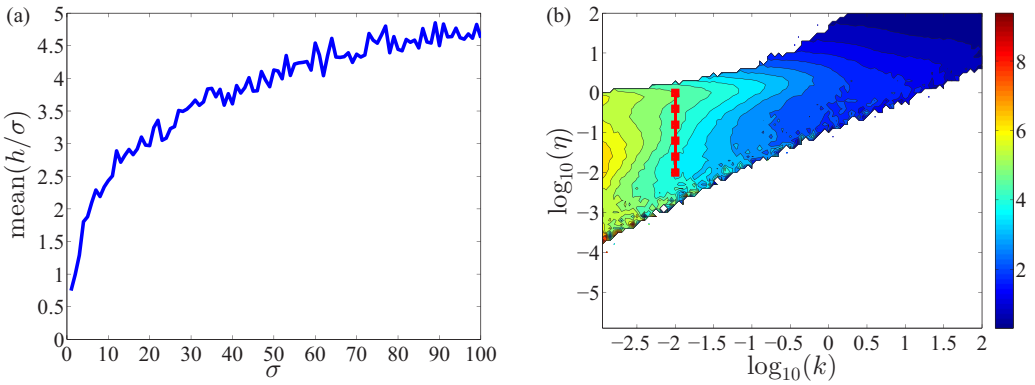

FIG. 11. The ratio between the barrier height and noise amplitude when the first transition occurs. (a) h_V/σ of various σ in dynamics Eq. (2) with potential Eq. (7). (b) h_U/η of various k and η in dynamics Eq. (1) with potential Eq. (6). Average is taken by 100 and 1000 trajectories in Figs. 11(a) and 11(b), respectively. The blank domain in Fig. 11(b) means either the particle walks around the whole space before $\lambda = -2$ (top left blank) or does not transit until the bifurcation occurs at $\lambda = 2$ (bottom blank). Figure 11(a) is the slice marked by red line with square symbols in Fig. 11(b). This shows the first transition usually happens when barrier height is in the same order of the noise amplitude.

We choose 10×10 square lattice in two dimensions, $D = 25, \sigma = 1$, and

$$f(x) = ax + x^3 - x^5, \quad g(x) = 1 + x^2. \tag{12}$$

We vary a from -2 to 0 with step size 0.1. In simulations, $\Delta t = 10^{-4}$ and 10^6 steps are taken in each a. The system achieves ergodicity for each parameter. The results of order parameter $|\langle x_i\rangle|$, average of variances $\mathrm{Var}(x_i)$ and average of KL divergence on each site are shown in Fig. 12. The signals at the critical point are clear.

C. Classification by noise-to-signal ratio

We have discussed about the models and transitions. But for a specific problem, the way to identify the regime of a critical

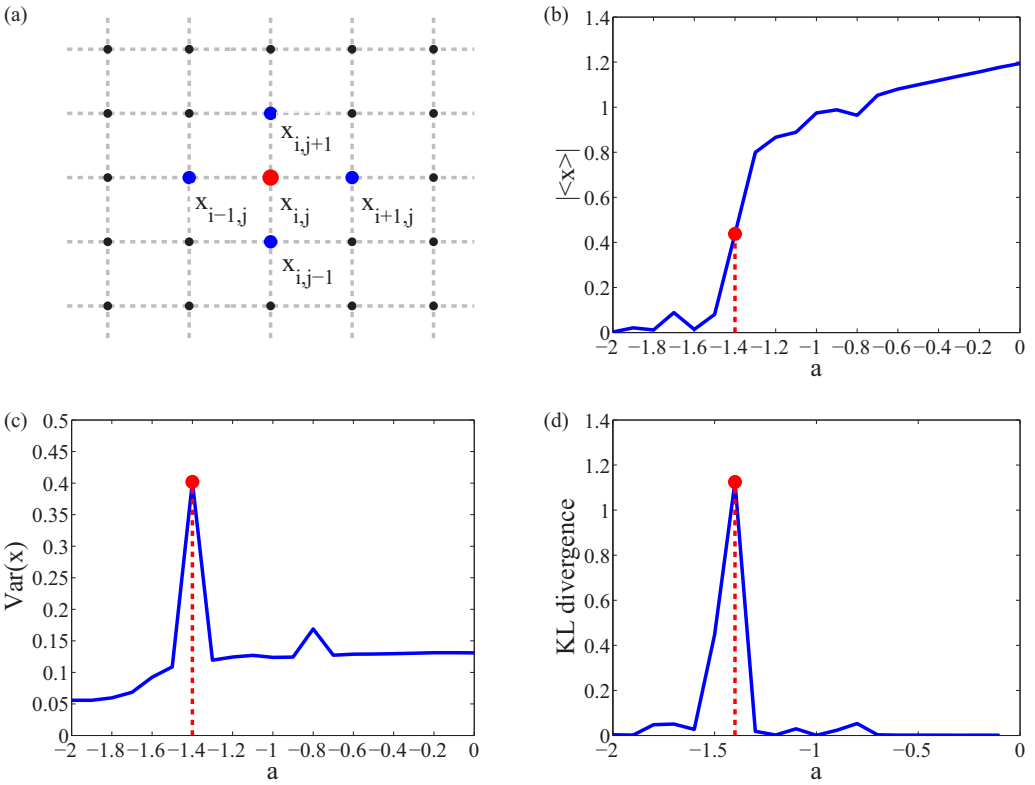

FIG. 12. Critical transition of a lattice model with large noise. The system Eq. (11) is ergodic in each parameter a. Figure 12(a) is a sketch map of the dynamics on a 10×10 lattice. The absolute mean value Fig. 12(b), average variance of sites Fig. 12(c), and KL distance of adjacent equilibrium distributions Fig. 12(d) are computed. The red points in the last three figures correspond to the critical transition point where a sudden change of distribution occurs.

transition is a nontrivial problem. The NSR can give heuristic classification on this problem, though it is difficult to give an exact judgment. In the model Eq. (2) with Eq. (7), we define the NSR as

$$\mathrm{NSR} = \log \frac{\sigma}{\mu}, \tag{13}$$

where σ is the amplitude of noise and μ is the magnitude of the characteristic potential energy barrier. To identify to which regime the considered dynamics with different parameters belongs, we do a large number of simulations and make the grouping from the numerical behaviors of the sampled trajectories. From the discussions in the previous subsection, it is natural to make the following classification according to the characteristics of the simulated path as

Case I: if the trajectory has only one transition point, which is very close to the bifurcation point;

Case II: if the trajectory has only one transition point, but it is far before the bifurcation point;

Case III: if the trajectory flicks over the whole space and crosses the boundary between two attraction basins many times.

We take the dynamics (2) with potential (7). When doing the simulation, $\mu = 100$ is fixed and σ varies. We simulate trajectories for every NSR expressed by Eq. (13) from -4 to 1 with step size 0.1 and observe into which case the trajectories fall. The results are shown in Fig. 13. In Fig. 13, we set each case an number—1 for Case I, 2 for Case II, and 3 for Case III. At each NSR from -4 to 1 with step 0.1, 100 trajectories are simulated and classified. The average case number is

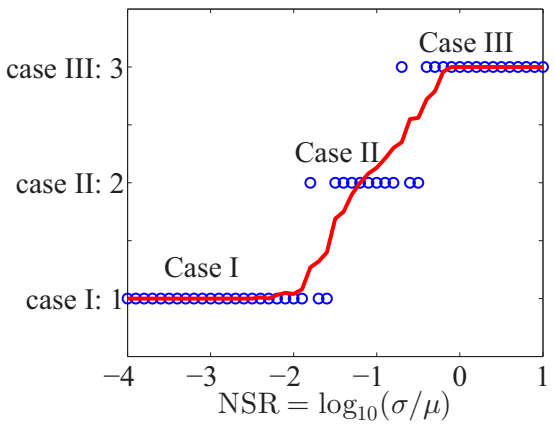

FIG. 13. Simulations to identify into which case the trajectory belongs by the noise-to-signal ratio (NSR). The vertical axis shows the classification according to the characteristics of sample paths with different parameters: Case I is for small noise, in which the only one transition point is near the bifurcation point; Case II is for medium noise, in which the only one transition point is far before the bifurcation point; and Case III is for large noise, in which the trajectory flicks over the space. We set Case I to value 1, Case II to 2, and Case III to 3. The NSR varies from -4 to 1 with step size 0.1, and the blue circles are results of several trajectories corresponding to the related NSR. The red line is the average result of 100 trajectories at each NSR. In the transitional zones with intermediate NSR, we have overlapping between regimes. From the figure, we can know how the system evolves with the changing of NSR.

plotted by the red line. And the blue circles are results of several trajectories. From these results, we can summarize the following heuristic rules:

(i) When NSR $\ll -2$, i.e., $\sigma/\mu \ll O(10^{-2})$, the system falls in Case I;

(ii) When NSR ~ -1, i.e., $\sigma/\mu \sim O(10^{-1})$, the system falls in Case II with high probability;

(iii) When NSR $\gtrsim 0$, i.e., $\sigma/\mu \gtrsim O(1)$, the system falls in Case III.

These observations may provide useful references for the consideration of specific problems with a model as Eq. (1) or Eq. (3).

V. CONCLUSION AND DISCUSSION

We studied critical phenomena or critical transitions with both a minimal model and its generalized model. The analysis covers both conceptual and computational aspects. One major contribution in our paper is that we classify the critical transitions into three types according to their time scales and noise strengths, which are state transition, basin transition, and distribution transition, corresponding to the small, medium, and large noise cases respectively. In each case, we define the critical transition and provide the early-warning indicator or property. For the small noise case, variance and correlation are good signals to predict the critical transition. For a medium noise case with multiple particles, we find that the most probable jumping point is located at which the potential barrier has the same order as the noise amplitude. For large noise case, KL divergence shows significant signals. We also proposed to distinguish two concepts in the medium noise case, i.e., the single-trajectory statistics and multiple-trajectory statistics. These two concepts lead to different locations of transition points in the medium noise, but they almost coincide in the small and large noise cases. The simulations support our theoretical results and the proposed indicators.

There are many related works considering the critical tipping points in other ways. We next briefly discuss several representative models and their relation with our work.

(a) R-tipping theory. Wieczorek *et al.* and Luke *et al.* have studied the excitability and compost bomb instability of dynamical systems [44,45]. From these works, Ashwin *et al.* established a rate-dependent tipping (R-tipping) point theory [46,47] which mainly focused on the changing rate of the parameters. According to different rates of parameters, the traceability of trajectories will suddenly change, and that is where the R-tipping occurs. R-tipping may be found even without bifurcation and noise. In our model [see Eq. (2)], we have transformed the rate changing effect into the noise term. We show that there are noise-induced transitions across the basin of attraction without an R-tipping. Thus, the R-tipping transition is a parallel theory with our framework, which focuses on different aspects of critical phenomena.

(b) Sandpile model. The discrete sandpile model has been long used in the modeling of earthquake and so on [48,49]. We can consider the sandpile model as a discrete situation in our medium noise case where no parameters change and $\tau_\lambda = +\infty$. The noise is set as adding a ball. We know the critical point is the state when the number of balls becomes closing to the threshold (N) of a cell. Just as demonstrated in

Fig. 11, the most dangerous time (the number of balls is $N-1$) is the state when the barrier height ($N-$ the number of balls) gets close to the noise amplitude (counts of ball added each time, which is just 1). We can set the size of the largest cluster of cells which have $N-1$ balls as an order parameter. It will be a good indicator to the critical corruption. In principle, our idea or framework in the medium noise case is also plausible in this discrete model, but more detail analyses are necessary for future studies. A continuous sandpile model was also proposed in 1996 [50], but the start point and equations differ from ours.

(c) Percolation models. Sornette *et al.* used different models, i.e., Bohman-Frieze-Wormald [51] or the hierarchical fiber bundle model [52,53], to consider the transition problem in the percolation. They found several statistics which are powerful in prediction based on the pdfs. Though the model differs, we can consider that the system is in the situation $\tau_{\text{ergo}} \ll \tau_\lambda$. Thus, at any time the system is ergodic and we can obtain the equilibrium pdfs. Many statistics can be computed and used for prediction with these pdfs. Percolation models are not in a SDE form but the classification also makes sense for this case.

In our framework, some important questions remain to be further studied in the future.

(i) Sampling frequency issue. Most of the results are simulated with a fixed small k, which means that the parameter λ changes very slowly. Although our theory still works for large k, which is transformed to a different noise magnitude σ, the large-k case generally corresponds to very low sampling size since the system varies too quickly in real time. We need a high-frequency sampling to obtain sufficient information. This generally imposes difficulty in real problems.

(ii) Time window issue for single-trajectory data. In dealing with the single-trajectory data, we move the time window to get the statistics. In general, the width of the time window influences the final result. To make the result more accurate, the width of the window should tend to zero. In this case, more sample points in unit time are needed. Thus, a balance is required in real situations.

(iii) More general bifurcation patterns. The bifurcation we discussed so far is only for the codimension one case. For codimension two or higher, the situation will be much more complicated. Generally, it is difficult to know external or internal parameters simply based on the observed data. Besides, although the indicators in small noise case can be used in both catastrophic and noncatastrophic cases [54], the signal is not so significant in the latter situation.

(iii) Indicators for transition. So far we have good indicators for Case I with small noise. For Case II, the proposed indicator for the single-trajectory critical transition has a clear theoretical background but needs further tests and improvements. The best way to set the parameters like the number of points N or neighbor size δ is a practical issue. For Case III with large noise, one may encounter the problem whereby the noise dominates the signal if the noise magnitude is too large big.

(iv) More general models. As known to us, R-tipping theory, the sandpile model, the percolation models, and many others cannot be simply written in the form of an SDE with white noise perturbations. However, their critical phenomena and classifications share many similarities with our model. More general models need to be established to consider these models in a unified framework.

These problems challenge future research on the study of critical transition theory.

ACKNOWLEDGMENTS

The authors acknowledge Professor Hong Qian for his valuable discussion. This work (L.C.) was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB13040700) and the NSFC (Grants No. 91529303, No. 91439103, and No. 61134013). T.L. acknowledges the support of the NSFC (Grants No. 11171009, No. 11421101, and No. 91530322) and the National Science Foundation for Excellent Young Scholars (Grant No. 11222114).

[1] D. Sornette, Proc. Natl. Acad. Sci. USA **99**, 2522 (2002).
[2] M. Scheffer, S. Carpenter, J. A. Foley, C. Folke, and B. Walker, Nature **413**, 591 (2001).
[3] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara, Nature **461**, 53 (2009).
[4] V. Dakos, S. R. Carpenter, W. A. Brock, A. M. Ellison, V. Guttal, A. R. Ives, S. Kefi, V. Livina, D. A. Seekell, E. H. van Nes *et al.*, PLoS ONE **7**, e41010 (2012).
[5] S. Kéfi, V. Guttal, W. A. Brock, S. R. Carpenter, A. M. Ellison, V. N. Livina, D. A. Seekell, M. Scheffer, E. H. van Nes, and V. Dakos, PLoS ONE **9**, e92097 (2014).
[6] S. R. Carpenter, Proc. Natl. Acad. Sci. USA **102**, 10002 (2005).
[7] S. Carpenter and W. Brock, Ecol. Lett. **9**, 311 (2006).
[8] S. R. Carpenter, J. J. Cole, M. L. Pace, R. Batt, W. Brock, T. Cline, J. Coloso, J. R. Hodgson, J. F. Kitchell, D. A. Seekell *et al.*, Science **332**, 1079 (2011).
[9] V. Dakos, M. Scheffer, E. H. van Nes, V. Brovkin, V. Petoukhov, and H. Held, Proc. Natl. Acad. Sci. USA **105**, 14308 (2008).
[10] T. M. Lenton, H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber, Proc. Natl. Acad. Sci. USA **105**, 1786 (2008).
[11] J. M. T. Thompson and J. Sieber, Int. J. Bifurcat. Chaos **21**, 399 (2011).
[12] J. M. T. Thompson and J. Sieber, IMA J. Appl. Math. **76**, 27 (2011).
[13] R. M. May, S. A. Levin, and G. Sugihara, Nature **451**, 893 (2008).
[14] L. Lin and D. Sornette, arXiv:1510.08162 (2015).
[15] T. Lenton, V. Livina, V. Dakos, E. Van Nes, and M. Scheffer, Philos. Trans. R. Soc. Lon. A: Math. Phys. Eng. Sci. **370**, 1185 (2012).
[16] V. Dakos, S. R. Carpenter, E. H. van Nes, and M. Scheffer, Philos. Trans. R. Soc. Lond. B: Bio. Sci. **370**, 20130263 (2015).

[17] H. L. D. de S. Cavalcante, M. Oriá, D. Sornette, E. Ott, and D. J. Gauthier, Phys. Rev. Lett. **111**, 198701 (2013).
[18] C. Kuehn, Physica D **240**, 1020 (2011).
[19] L. Chen, R. Liu, Z.-P. Liu, M. Li, and K. Aihara, Sci. Rep. **2**, 342 (2012).
[20] R. Liu, M. Li, Z.-P. Liu, J. Wu, L. Chen, and K. Aihara, Sci. Rep. **2**, 813 (2012).
[21] R. Liu, K. Aihara, and L. Chen, Quant. Biol. **1**, 105 (2013).
[22] R. Liu, X. Wang, K. Aihara, and L. Chen, Med. Res. Rev. **34**, 455 (2014).
[23] X. Yu, G. Li, and L. Chen, Bioinformatics **30**, 852 (2014).
[24] R. Liu, P. Chen, K. Aihara, and L. Chen, Sci. Rep. **5**, 17501 (2015).
[25] R. Liu, X. Yu, X. Liu, D. Xu, K. Aihara, and L. Chen, Bioinformatics **30**, 1579 (2014).
[26] J. M. T. Thompson and J. Sieber, in *IUTAM Symposium on Nonlinear Dynamics for Advanced Technologies and Engineering Design* (Springer, Berlin, 2013), pp. 3–16.
[27] V. Dakos, E. H. van Nes, and M. Scheffer, Theor. Ecol. **6**, 309 (2013).
[28] W. Horsthemke and R. Lefever, *Noise Induced Transitions* (Springer, Berlin, 1984).
[29] L. Arnold, *Random Synamical Systems* (Springer, Berlin, 1995).
[30] P. Metzner, C. Schütte, and E. Vanden-Eijnden, Multis. Model. Simul. **7**, 1192 (2009).
[31] N. Berglund and B. Gentz, *Noise-Induced Phenomena in Slow-Fast Dynamical Systems: A Sample-Paths Approach* (Springer Science & Business Media, London, 2006).
[32] A. Zaikin and J. Kurths, Chaos **11**, 570 (2001).
[33] D. Sornette, *Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools* (Springer Science & Business Media, Berlin, 2004).
[34] J. J. Binney, N. Dowrick, A. Fisher, and M. Newman, *The Theory of Critical Phenomena: An Introduction to the Renormalization Group* (Oxford University Press, Oxford, 1992).
[35] P. Ao, H. Qian, Y. Tu, and J. Wang, arXiv:1310.5585 (2013).
[36] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* (Springer, Berlin, 1983).
[37] M. I. Freidlin, J. Szücs, and A. D. Wentzell, *Random Perturbations of Dynamical Dystems*, 3rd ed., Vol. 260 (Springer, Heidelberg, 2012).
[38] C. Lv, X. Li, F. Li, and T. Li, PLoS ONE **9**, e88167 (2014).
[39] W. E, W. Ren, and E. Vanden-Eijnden, Phys. Rev. B **66**, 052301 (2002).
[40] W. E, W. Ren, and E. Vanden-Eijnden, J. Phys. Chem. B **109**, 6688 (2005).
[41] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction* (Springer, New York, 2009).
[42] C. Van den Broeck, J. M. R. Parrondo, and R. Toral, Phys. Rev. Lett. **73**, 3395 (1994).
[43] R. Müller, K. Lippert, A. Kühnel, and U. Behn, Phys. Rev. E **56**, 2658 (1997).
[44] S. Wieczorek, P. Ashwin, C. M. Luke, and P. M. Cox, Proc. R. Soc. London A: Math. Phys. Sci. **467**, 1243 (2011).
[45] C. Luke and P. Cox, Eur. J. Soil. Sci. **62**, 5 (2011).
[46] P. Ashwin, S. Wieczorek, R. Vitolo, and P. Cox, Philos. Trans. R. Soc. Lond. A: Math. Phys. Eng. Sci. **370**, 1166 (2012).
[47] P. Ashwin, C. Perryman, and S. Wieczorek, arXiv:1506.07734 (2015).
[48] P. Bak and C. Tang, J. Geophys. Res. **94**, 15635 (1989).
[49] Y. Huang, H. Saleur, C. Sammis, and D. Sornette, Europhys. Lett. **41**, 43 (1998).
[50] L. Gil and D. Sornette, Phys. Rev. Lett. **76**, 3991 (1996).
[51] W. Chen, M. Schröder, R. M. DSouza, D. Sornette, and J. Nagler, Phys. Rev. Lett. **112**, 155701 (2014).
[52] J. V. Andersen and D. Sornette, Phys. Rev. E **72**, 056124 (2005).
[53] D. Sornette and J. V. Andersen, Europhys. Lett. **74**, 778 (2006).
[54] S. Kéfi, V. Dakos, M. Scheffer, E. H. Van Nes, and M. Rietkerk, Oikos **122**, 641 (2013).