Supplementary Information: A universal indicator of critical state transitions in noisy complex networked systems

Junhao Liang, Yanqing Hu, Guanrong Chen & Tianshou Zhou

Contents

A  Theory of critical transition .................................................................................................................. 2
   A.1 One-dimensional case ................................................................................................................... 2
   A.2 Higher-dimensional case ............................................................................................................. 7
B  Dynamics of a simple system ............................................................................................................. 12
C  Numerical simulation methods ......................................................................................................... 15
   C.1 Network structure setting ........................................................................................................... 15
   C.2 Simulating the effect of random failures .................................................................................... 17
D  Extended validation ............................................................................................................................. 18
   D.1 Simulations of artificial networks ............................................................................................... 18
   D.2 Effect of noise on the characteristic returning time .................................................................. 20
References .................................................................................................................................................. 22
A Theory of critical transition

In this section, we introduce the mathematical approach to analyzing the critical transition\textsuperscript{1-3}, focusing on the critical slowing down phenomenon\textsuperscript{1,2,4-8} in a noisy setting. The analytical methods are based on the theory on stochastic differential equations and some known results from complex networks.

A.1 One-dimensional case

First, consider the following one-dimensional (1-D) nonlinear dynamical system subjected to small noise:

\[
\frac{dx}{dt} = F(x, \alpha) + \sigma \xi(t)
\]

(1)

where \( \alpha \) is a tunable real parameter, which captures the changing environmental conditions; \( \xi(t) \) is a standard Gaussian white noise with zero mean, satisfying \( \langle \xi(t) \xi(t') \rangle = \delta(t-t') \), representing a perturbation; \( \sigma \) represents the noise intensity and function \( F(x, \alpha) \) represents the system dynamics.

In the absence of noise, i.e., \( \sigma = 0 \), assume that the system is at one of the stable fixed points, \( x_s \), given by

\[
\begin{align*}
F(x_s, \alpha) &= 0 \\
\lambda_s &= \frac{\partial}{\partial x} F(x_s, \alpha) < 0
\end{align*}
\]

(2)

The solution of this equation provides a resilience function, \( x(\alpha) \), which represents all possible states of the unperturbed system as a function of \( \alpha \). The shape of this function is uniquely determined by the functional form of \( F(x, \alpha) \). In contrast, the momentary state of the system is determined by the tunable parameter \( \alpha \). At some critical point \( \alpha_{critical} \), the resilience function may feature a bifurcation (\( \alpha_1 \) of Fig. S1) or become non-analytic (\( \alpha_2 \) of Fig. S1), indicating\textsuperscript{2,9} that the system loses its resilience via undergoing a sudden transition to a different, often undesirable, fixed...
Figure S1. Possible shapes for the resilience function of the 1-D system (1), where the solid lines represent stable states whereas the dash lines represent unstable ones.

In the presence of noise, i.e., $\sigma \neq 0$, the Fokker-Planck equation corresponding to Eq. (1) is given by

$$\frac{\partial P(x, t)}{\partial t} = -\frac{\partial}{\partial x} \left( F P \right) + \frac{\sigma^2}{2} \frac{\partial^2 P}{\partial x^2}$$

(3)

where $P = P(x, t)$ is the probability distribution that the system is at state $x$ at time $t$. Recall that by letting $\Omega$ be the size of a chemical reaction system, the intensity of the noise due to thermal movement of the species molecules is proportional to $1/\sqrt{\Omega}$ (called the noise factor) $^{10}$, giving $\sigma^2 \approx 1/\Omega$. Thus, similarly to the case of the linear noise approximation$^{10}$, if the system size $\Omega$ is sufficiently large or if the intensity $\sigma^2$ is sufficiently small (which will always be assumed in the following analysis), then one may set $x(t) = \phi(t) + \left(1/\sqrt{\Omega} \right) \mu(t)$, where $\phi(t)$ is the deterministic part whereas $\mu(t)$ is the stochastic fluctuation part satisfying $\{\mu(t)\} = 0$. To calculate the autocorrelation of the stochastic variable $\mu(t)$, denote by $Q(\mu, t)$ its probability density function. Then, the linear noise approximation yields the following Fokker-Planck equation:

$$\frac{\partial Q(\mu, t)}{\partial t} = -\lambda_0 \frac{\partial}{\partial \mu} \left( \mu Q \right) + \frac{\sigma^2}{2} \frac{\partial^2 Q}{\partial \mu^2}$$

(4)

At steady state, the autocorrelation is given by

$$\langle \mu^2(t) \rangle_{ss} = \frac{\sigma^2}{-2\lambda_0}$$

(4a)
and the lag- \( t \) autocorrelation by

\[
\langle \mu(0)\mu(t) \rangle_{st} = \frac{\sigma^2}{-2\lambda} e^{-\lambda t}
\]  

(4b)

Next, to show the perturbation effect of the noise, adopt the effective stability approximation procedure\(^1\) and set

\[
x(t) = x_s + x_p(t) + \sigma \mu(t) = (x_s + \sigma \mu) + x_p
\]

(5)

where \( x_p(t) \) represents a deterministic deviation from the steady state \( x_s \). Note that, in the case where the perturbation is very small, the evolution of \( x_p(t) \) is governed by the following linear equation with a stochastic coefficient:

\[
\frac{dx_p(t)}{dt} = \lambda x_p(t)
\]

(6)

where \( \lambda = \frac{\partial F}{\partial x} \big|_{x=x_s,\sigma\mu(t)} \). Considering \( \lambda \) as a function of \( \sigma \), making the Taylor expansion, and noting the smallness of \( \sigma \), one obtains an approximation

\[
\lambda(\sigma) \approx \lambda(0) + \sigma \lambda'(0)
\]

(7)

where \( \lambda(0) = \lambda_0 = \frac{\partial F}{\partial x} \big|_{x=x_s} = F'(x_s, k_0) \), \( \lambda'(0) = \mu(t) \frac{\partial^2 F}{\partial x^2} \big|_{x=x_s} = \mu(t) F''(x_s, k_0) \) (denoted by \( \lambda_i(t) \) for convenience). With such an approximation, Eq. (6) can be rewritten as

\[
\frac{dx_p(t)}{dt} = [\lambda_0 + \sigma \lambda_i(t)] x_p(t)
\]

(8)

where \( \lambda_i(t) \) depends on the steady-state correlation \( \langle \mu(0)\mu(t) \rangle_{st} \) given by Eq. (4b). Note that the stability of the zero solution to Eq. (8) is equivalent to that of \( x_s \) in the average sense. However, we are interested in the average stability. In this case, the equation for the average of \( x_p(t) \) is governed by

\[
\frac{d\langle x_p(t) \rangle}{dt} = \lambda_0 \langle x_p(t) \rangle + \sigma \langle \lambda_i(t) x_p(t) \rangle
\]

(9)

To study this equation, one may assume various forms of \( \lambda_i(t) \). In the case that \( |\lambda_0| \gg |\lambda_i(t)| \), we adopt the Bourret approximation\(^2\) for Eq. (9), obtaining
\[
\frac{d\langle x_p(t)\rangle}{dt} \approx \lambda_0 \langle x_p(t)\rangle + \sigma^2 \int_0^t J_\epsilon(t-\tau) \langle x_p(\tau)\rangle d\tau
\]

where \( J_\epsilon(t-\tau) = \langle \lambda_\epsilon(t) \exp(\lambda_0(t-\tau)) \lambda_\epsilon(\tau) \rangle \) is the correlation function for fluctuations. At steady state, using Eq. (4b), one obtains

\[
J_\epsilon(t-\tau) = \left[ \frac{\sigma F_\epsilon'(x_s, k_0)}{2\lambda_0} \right] e^{2\lambda_0(t-\tau)}
\]

Substituting Eq. (11) into Eq. (10) yields

\[
\frac{d\langle x_p(t)\rangle}{dt} = \lambda_0 \langle x_p(t)\rangle + a \int_0^t e^{2\lambda_0(t-\tau)} \langle x_p(\tau)\rangle d\tau
\]

where \( a = \left[ \frac{\sigma^2 F_\epsilon'(x_s, k_0)^2}{(-2\lambda_0)} \right] > 0 \).

Next, use the Laplace transformation method to solve Eq. (12). For this purpose, by setting \( \tilde{x}_p(s) = \int_0^\infty x_p(t)e^{-st} dt \), one has

\[
\langle \tilde{x}_p(s) \rangle = \left( s - \frac{a}{2\lambda_0 - s} \right)^{-1} \langle x_p(0) \rangle
\]

The averaged perturbation mode \( \langle x_p(t) \rangle \) is asymptotically stable if and only if the roots of the equation \( s - \frac{a}{2\lambda_0 - s} = 0 \) all have negative real parts. Hence, one gets the effective eigenvalue \( \tilde{\lambda}_0 \) expressed as

\[
\tilde{\lambda}_0 = \frac{3\lambda_0 + \sqrt{\lambda_0^2 + 4a}}{2} = \frac{3\lambda_0 + \sqrt{\lambda_0^2 + 2 \left[ \sigma^2 F_\epsilon'(x_s, k_0)^2 \right] / (-\lambda_0)}}{2}
\]

The fact of \( \lambda_0 + \sqrt{\lambda_0^2 + 4a} > 0 \) implies that the noise positively contributes to the effective characteristic value of the stochastic system. In particular, for a small negative value of \( \tilde{\lambda}_0 \), although the deterministic system is stable, the stochastic system may approach and even cross the critical state if the noise is large. In order to better characterize such a critical transition, define

\[
T = -\frac{1}{\tilde{\lambda}_0}
\]

This index actually represents the characteristic returning time of the 1-D system because
\( \langle x_p(t) \rangle - \exp(\tilde{\lambda}_0 t) = \exp(-t/T) \). We will show that in contrast to \( \tilde{\lambda}_0 \), \( T \) is a better indicator to predict when the system goes to a critical state.

First, critical transition implies that a dynamic system has two potential stable states\(^{1,3}\), which usually originates from the so-called fold-type bifurcation. According to the bifurcation theory for 1-D deterministic systems\(^{13,14}\), when a system tends to go through a fold catastrophic bifurcation at \( x_j \), the absolute value of the eigenvalue, \( |\tilde{\lambda}_0| \), decreases to zero (namely, \( \tilde{\lambda}_0 \) increases to zero since it is negative), which corresponds to the increasing of both the lag- \( \tau \) autocorrelation \( K(\tau) \) defined by \( K(\tau) = \langle \langle x(0) x(\tau) \rangle \rangle_{ss} \) and the variance \( \text{Var} \) defined by \( \langle \langle x^2(t) \rangle \rangle_{ss} \) evaluated at steady state, since both are monotonic increasing functions of \( \tilde{\lambda}_0 \) as can be seen from Eq. (4a) and Eq. (4b). This is the same as the early-warning signals of critical slowing down: slower recovery rate from perturbations\(^{4,5}\), higher autocorrelation\(^{6,7}\) and higher variance\(^{8}\).

Second, one can see from Eq. (14) that noise always has a positive effect on the characteristic value \( \tilde{\lambda}_0 \). When the system moves close to the critical state, \( \tilde{\lambda}_0 \) will approach zero from the negative direction. In particular, large noise can make \( \tilde{\lambda}_0 \) become positive and, in this case, \( T \) is negative. A negative characteristic returning time means that the system’s recovery to the stable state is impossible.

Finally, notice that an approximate solution to Eq. (8) is given by \( x_p(t) \approx x_p(0) e^{\tilde{\lambda}_0 t} \). From this expression, one can see that, for an initial deviation from the equilibrium, \( d_0 = x(0) - x_j \) and a fixed \( \tilde{\lambda}_0 < 0 \), the time-evolutional deviation \( d_t = x(t) - x_j \) will always return to zero exponentially after the time becomes sufficiently large. From the viewpoint of dynamical systems, the dominant eigenvalue (i.e., \( \tilde{\lambda}_0 \) in our case) characterizes the rate/speed of the recovery to the stable equilibrium. In applications, \( \tilde{\lambda}_0 \) can actually be considered as a recovery speed, which can be estimated from empirical data through simple linear regression (in this case, \( \tilde{\lambda}_0 \) is commonly called the Lyapunov exponent). Typically, one takes \( x_p(t) \approx 0 \) when \( t > -1/\tilde{\lambda}_0 \). Thus, the quantity \( T = -1/\tilde{\lambda}_0 \) characterizes the time returning to the equilibrium. In summary, compared
with the recovery rate \( \tilde{\lambda}_0 \), the index \( T \) serves as a better indicator for forecasting critical transitions in a highly nonlinear dynamical system subjected to noise.

### A.2 Higher-dimensional case

Previous studies\(^2\)\(^3\)\(^15\)\(^16\) showed that, when external environments change (captured by parameters), heterogeneous complex networks (nodes being non-similar) with high modularity tend to shift gradually, while homogeneous complex networks (nodes being similar) with high connectivity tend to occur at critical transitions up to some critical thresholds. This is because the homogeneity of nodes in a network affects the way in which the system with local alternative states responds to changing conditions\(^2\). Local damages of the homogeneous network tend to be repaired by subsidiary inputs from connected units until the system collapses at a critical stress level. In contrast, heterogeneous networks where components differ but are modular tend to have an adaptive capacity, so they can gradually adjust themselves to changes.

Here, we consider the case of homogeneous networks subjected to internal and external fluctuations (or noise). Similarly to the deterministic case\(^17\), we first derive a one-dimensional (1-D) indicative equation from a higher-dimensional system, and then use the above developed theory to analyze various critical transitions in complex networked systems.

We begin by describing the complex networks in interest. A complex network can be described as a simple graph with \( N \) nodes, where there is at most one directed edge with positive weight from node \( i \) to node \( j \) for any \( i \) and \( j \), and a self-edge from a node \( i \) to itself is permissible. Thus, this network can be characterized by an adjacency matrix \( A = (a_{ij})_{N \times N} \), where element \( a_{ij} \) represents the weight of the directed edge from node \( j \) to node \( i \). If there is no directed edge from node \( j \) to node \( i \), then \( a_{ij} = 0 \). If \( a_{ij} > 0 \), \( j \) is a neighbor of \( i \), meaning that \( i \) can be influenced by \( j \). Since an undirected edge between \( i \) and \( j \) is equivalent in effect to both a directed edge from \( i \) to \( j \) and vice versa, the above adjacency matrix \( A \) is also suitable for the description of an undirected graph and in this case it is symmetric.
Define the in-degree and the out-degree of node $i$ by $k_i^\text{in} = \sum_{j=1}^{N} a_{ij}$ and $k_i^\text{out} = \sum_{k=1}^{N} a_{ki}$, respectively. It is easy to show the relation of $\langle k_i^\text{in} \rangle = \langle k_i^\text{out} \rangle$, where the symbol $\langle \cdot \rangle$ represents the mean (i.e., average) over all the network nodes (i.e., the sum over index $i$).

![Graph representation of network nodes](image)

**Figure S2.** Graph representations of networks.

Suppose that the homogeneity of a network is so high that the dynamics of every node in the network can be approximately described by the following equation in the unified form:

$$\frac{dx_i}{dt} = f(x_i) + \sum_{j=1}^{N} a_{ij} \left[ g(x_i, x_j) + \eta_j(t) \right] + \xi_i(t) \quad (16)$$

where $f(x_i)$ describes the primitive deterministic dynamics of node $i$ and $\xi_i(t)$ is a Gaussian white noise from fluctuations in the external environment; $a_{ij} g(x_i, x_j)$ represents the effective impact of node $j$ on node $i$, and $\eta_j(t)$ represents the effective noise in this impacting process. In Eq. (16), $x_i$ describing the dynamics of node $i$ may represent the quantity of a species, the expressional level of a gene, and so on.

In most of the previously-studied dynamical models, functions $f$ and $g$ were assumed to have some simple forms. Here, we assume that the form of function $f$ may be complex but is of high linearity, namely $0 \neq |f''| \ll 1$. Similarly, function $g(x, y)$ is assumed to be of high linearity in variable $y$ and is less affected by variable $x$, namely $|\partial^2 g / \partial y^2| \ll 1$ and $|g(x, y) - g(\bar{x}, y)| \ll 1$. These assumptions are reasonable for many interactive processes in real-world complex networks. In addition, we assume that the amplitude of noise is much smaller than the size of the deterministic part, so that the conditions in the “critical slowing down” theory\textsuperscript{1,2}, the linear noise approximation\textsuperscript{10} and the effective stability approximation\textsuperscript{11}, can be satisfied.
Note that the average of \( \{ x_i \} \) reflects the global properties of the whole network. The normalized average \( \sum_{i=1}^{N} x_i / N \) does not seem to be a good description because the influence power of each node is not the same in the network. In order to highlight the influential differences between different nodes, we endow a weight \( w_j \) to each node with the conservative condition \( \sum_{i=1}^{N} w_i = 1 \). In our case, we set \( w_j = k_{ij}^\text{out} / \sum_{j} k_{ij}^\text{out} \). We will use the weighted average \( \langle x_i \rangle_w \), namely \( \langle x_i \rangle_w = \sum_{j} k_{ij}^\text{out} x_j / \sum_{j} k_{ij}^\text{out} \equiv \langle k_{ij}^\text{out} x_j \rangle / \langle k_{ij}^\text{out} \rangle \), to describe the mean behavior. Since we have assumed that the network is of a high homogeneity, the weight \( w_j \) can be approximated as \( a_{ij} / \sum_{j} a_{ij} \) for all \( i \).

Under the above assumptions, similarly to the deterministic case\(^{17}\), we can derive the following equation from Eq. (16):

\[
\frac{dx_i}{dt} = f(x_i) + k_{ij}^\text{in} g(x_i, \langle x_i \rangle_w) + \zeta_i(t)
\]  

(17)

where \( \zeta_i(t) = \sum_{j=1}^{N} a_{ij} \eta_j(t) + \xi_i(t) \) is integrated noise, depending on the properties of the original noisy sources.

We are interested in the time-evolutional equation for the average quantity \( \langle x_i \rangle_w \). From the linearity of \( \langle \cdot \rangle_w \) with respect to the argument, we have

\[
\frac{d\langle x_i \rangle_w}{dt} = \left\langle \frac{dx_i}{dt} \right\rangle_w = \left\langle f(x_i) \right\rangle_w + \left\langle k_{ij}^\text{in} g(x_i, \langle x_i \rangle_w) \right\rangle_w + \left\langle \zeta_i(t) \right\rangle_w
\]  

(18)

According to our hypothesis on the function \( f \), we have the approximation

\[
\left\langle f(x_i) \right\rangle_w \approx f\left( \langle x_i \rangle_w \right)
\]  

(19)

With our hypothesis on the function \( g(x, y) \), we also have

\[
\left\langle k_{ij}^\text{in} g(x_i, \langle x_i \rangle_w) \right\rangle_w = \frac{\sum_{ij} k_{ij}^\text{in} k_{ij}^\text{out} g(x_i, \langle x_i \rangle_w)}{\sum_{ij} k_{ij}^\text{in} k_{ij}^\text{out}} \approx \frac{\sum_{ij} k_{ij}^\text{in} k_{ij}^\text{out} g(x_i, \langle x_i \rangle_w)}{\sum_{ij} k_{ij}^\text{in} k_{ij}^\text{out}} = \langle k_{ij}^\text{in} \rangle \langle g(x_i, \langle x_i \rangle_w) \rangle
\]  

(20)

Now, denote \( x = \langle x_i \rangle_w \), \( k_0 = \langle k_{ij}^\text{in} \rangle \) and \( \zeta(t) = \langle \zeta_i(t) \rangle_w \). Then, we arrive at the following 1-D system
\[ \frac{dx}{dt} = F(x, k_0) + \zeta(t) \]  

where \( F(x, k_0) = f(x) + k_0 g(x, x) \), and \( \zeta(t) \) is the effective noise, which is expressed as a linear combination of internal and external noisy sources in the original large-scale system. Equation (21) is completely similar to Eq. (1), which we have analyzed above, so it has a well-defined characteristic returning time \( T \).

Next, we analyze how the topological properties of the network affect the characteristic returning time \( T \). For this purpose, we derive another expression of the parameter \( k_0 \). In fact, denote by \( \langle k \rangle \) the average weight of the network, namely \( \langle k \rangle = \langle k^{in} \rangle = \langle k^{out} \rangle \). Then, we have the following factorization \( ^{17} \)

\[
k_0 = \frac{\langle k^{out} \rangle}{\langle k^{in} \rangle} = \langle k \rangle + \text{PCC}(\langle k^{in} \rangle, \langle k^{out} \rangle) \sqrt{FF_{in} \cdot FF_{out}}
\]

where \( \text{PCC} \) is the Pearson Correlation Coefficient, \( \text{PCC}(\langle k^{in} \rangle, \langle k^{out} \rangle) = \frac{\langle k^{in} \rangle - \langle k^{in} \rangle \langle k^{out} \rangle}{\sigma_{in} \sigma_{out}} \); \( FF_{in} = \frac{\sigma_{in}^2}{\langle k^{in} \rangle} \) and \( FF_{out} = \frac{\sigma_{out}^2}{\langle k^{out} \rangle} \) are the Fano Factors of the average in-degree and out-degree, where \( \sigma_{in}^2 = \sum_{i=1}^{N} (k^{in}_i - \langle k^{in} \rangle)^2 / N \) and \( \sigma_{out}^2 = \sum_{i=1}^{N} (k^{out}_i - \langle k^{out} \rangle)^2 / N \). Note that the value of \( \langle k \rangle \) reflects the density of the network and a dense/sparse network tends to have a large/small \( \langle k \rangle \). In contrast, \( FF_{in} \) and \( FF_{out} \) reflect the otherness of nodes and a network with greater degree differences between nodes tend to have greater \( FF_{in} \) and \( FF_{out} \). The coefficient

\( \text{PCC}(\langle k^{in} \rangle, \langle k^{out} \rangle) \in [-1, 1] \) describes the correlation between in-degree and out-degree. If \( \text{PCC}(\langle k^{in} \rangle, \langle k^{out} \rangle) \) is positive or negative, then the underlying network is called symmetric or asymmetric. Therefore, the sign of the in-out degree correlation reflects the network symmetry.

From the factorization (22), one can see that \( k_0 \) is a monotonically increasing function of \( \langle k \rangle \) and \( \text{PCC}(\langle k^{in} \rangle, \langle k^{out} \rangle) \). A positive or negative \( \text{PCC}(\langle k^{in} \rangle, \langle k^{out} \rangle) \) implies that \( k_0 \) is a monotonically increasing or decreasing function of \( FF_{in} \) and \( FF_{out} \).
The last issue is to determine the relationship between $T$ and $k_0$. Note that such a relationship can be in principle analyzed by using Eq. (21). Note also that for fold catastrophic bifurcation, there are two distinct situations of the dynamic curves for the 1-D system described by Eq. (21), referring to Fig. S3. The difference between these two situations is caused by the type of the interaction between the nodes. Since the edges between nodes represent the interactive effects, the topological characteristic parameter $k_0$ represents the strength of this effect. In case (A) of Fig. S3, a greater $k_0$ (representing a stronger interactive effect) leads to a larger $x_s$, implying that the interactive effect is enhanced (mutualistic). On the contrary, the interactive effect is repressed (competitive) in case (B) of Fig. S3. Also, note that the type of the interactive effect is determined by the function $g(x, y)$, that is, $\partial g(x, y)/\partial y > 0$ corresponds to the mutualistic effect whereas $\partial g(x, y)/\partial y < 0$ to the competitive effect.

![Figure S3](image)

**Figure S3.** Two types of fold catastrophic bifurcation in system (21).

Before the critical transition occurs, the system has its stable state with a large value of $x_s$. In the mutualistic case, with the decrease of $k_0$, the system moves closer to the bifurcation point and $\lambda_0$ (thus $\lambda_0'$) becomes greater (a universal law revealed by Ref. 13). Therefore, $T$ is a monotonically decreasing function of $k_0$. Similarly, in the competitive case $T$ is a monotonically increasing function of $k_0$.

From the above analysis, the relationship between the characteristic returning time $T$ and the topological structure of the network can be summarized in Table S1.
<table>
<thead>
<tr>
<th>Interaction type</th>
<th>Topological characteristic</th>
<th>Density</th>
<th>Difference</th>
<th>Topological characteristic</th>
<th>Symmetry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutualistic</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>Symmetric</td>
<td></td>
</tr>
<tr>
<td>Competitive</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>Asymmetric</td>
<td></td>
</tr>
</tbody>
</table>

Table S1. Relationships between the characteristic returning time $T$ and the topological structure of a network. Here, ‘+’/’-’ represents increase/decrease of $T$. For example, in the mutualistic case, one can increase the network’s density to decrease $T$ (similarly in the competitive case). However, the influence of the difference on the characteristic returning time depends on the symmetry of the network. Note that these hold only before the critical state transition that corresponds to $T > 0$.

B Dynamics of a simple system

Now, consider Eq. (16) with $f(x) = b + x \left(1 - \frac{x}{K}\right) \left(x - \frac{1}{C}\right)$. The reasons for choosing this form of $f$ are as below. First, such a choice has realistic significance since $f$ represents the Logistic growth with Allee effects in ecology\(^\text{18,19}\), which describes the dynamic characteristics of many realistic processes. Second, this function is a cubic polynomial. For a general complicated function, one may only consider the first several terms in its Taylor expansion as approximation. Third, $|f''(x)| = \left|2\left(\frac{1}{C + 1}\right) - \frac{6x}{CK}\right| \ll 1$ holds under some parametric conditions. In this case, the linearity of $f$ is strong, which meets the requirement of the theory.

Reasons for the setting of $g(x, y) = \frac{\beta_1 x^n + \beta_2 y^n}{1 + \alpha_1 x^m + \alpha_2 y^m}$ are below. The function $g$ of this form, also called a generalized Hill function, has been confirmed to be a good description for many interactive processes\(^\text{20}\). When the Hill coefficients $n, m$ are small (we will set $n = 2, m = 1$ in the simulations below), $g(x, y)$ is of high linearity in variable $y$. If we choose the parameters to satisfy $\beta_1 \ll \beta_2$ and $\alpha_1 \ll \alpha_2$, then $g(x, y)$ is less affected by variable $x$. All these settings also meet the conditions of the theory.

As for the primitive noise at node $i$, the specific form of $\xi_i(t)$ is determined by stochastic
sources. Since there is no knowledge of these noise sources, they are often considered as rapidly fluctuating forces, and their effect is captured by Gaussian white noise. Specifically, it is assumed that $\xi_i(t)$ are mutually independent Gaussian white noise sources, satisfying

$$\langle \xi_i(t) \rangle = 0, \quad \langle \xi_i(t) \xi_j(t') \rangle = \Gamma_{ij} \delta(t-t')$$

(23a)

where $\{\Gamma_{ij}\}$ are constants characterizing the sizes of the fluctuations $\{\xi_i(t)\}$, $\delta(t)$ is the Kronecker delta, and $\delta(t)$ is the common delta function. Similarly, assume that

$$\langle \eta_{ij}(t) \rangle = 0, \quad \langle \eta_{ij}(t) \eta_{kl}(t') \rangle = \Gamma_{ij} \delta_{i\ell} \delta_{j\ell} \delta(t-t')$$

(23b)

where $\{\Gamma_{ij}\}$ are constants characterizing the sizes of the fluctuations $\{\eta_{ij}(t)\}$. In addition, assume that these two kinds of noises are mutually independent, that is,

$$\langle \xi_i(t) \eta_{ij}(t') \rangle = 0$$

(23c)

The integrated noise at node $i$, given by $\zeta_i(t) = \sum_{j=1}^{\infty} a_j \eta_{ij}(t) + \xi_i(t)$, is a linear combination of some Gaussian white noise sources, so $\{\zeta_i(t)\}$ are Gaussian, satisfying

$$\langle \zeta_i(t) \rangle = 0, \quad \langle \zeta_i(t) \zeta_j(t') \rangle = \left( \sum_j a_j^2 \Gamma_{ij} + \Gamma_i \right) \delta_{ij} \delta(t-t')$$

(23d)

According to existing analytical results (see Section A in this Supplementary Information), the key to understand the global dynamics of the whole network is to study the 1-D equation

$$\frac{dx}{dt} = b + x \left( 1 - \frac{x}{K} \right) \left( \frac{x}{C} - 1 \right) + k_0 \frac{\beta x^n}{1 + \alpha x^m} + \zeta(t)$$

(24)

where $\beta = \beta_1 + \beta_2$, $\alpha = \alpha_1 + \alpha_2$; $k_0 = \langle k_{ij}^{out} \rangle_w$ is a topological characteristic parameter, and $\zeta(t) = \left( \zeta_i(t) \right)_w$ is the effective noise, which is Gaussian and satisfies

$$\langle \zeta(t) \rangle = 0, \quad \langle \zeta(t) \zeta(t') \rangle = \sum_j \left( \frac{k_{ij}^{out}}{\sum_j k_{ij}^{out}} \right)^2 \left( \sum_k a_k^2 \Gamma_{ik} + \Gamma_i \right) \delta(t-t')$$

(25)

Equation (24) is a quasilinear Langevin equation, and its deterministic dynamics is described as follows.

When $n = 2, m = 1$, denote by $F(x, k_0)$ the deterministic part of the right-hand side of Eq. (24). Then, $F(x, k_0) = 0$ is equivalent to
\[
  k_0(x) = -\frac{b}{\beta x^2} + \frac{1-b\alpha}{x} + \left(\frac{1}{CK} - \frac{\alpha}{C} - \frac{\alpha}{K}\right)x + \frac{\alpha x^2}{CK\beta} + \frac{1}{\beta} \left(\alpha - \frac{1}{C} - \frac{1}{K}\right)
\]  

(26)

Differentiating both sides of Eq. (26) yields

\[
  k_0'(x) = \frac{2b}{\beta x^3} + \frac{b\alpha - 1}{\beta x^2} + \frac{2\alpha x}{CK\beta} + \frac{1}{\beta} \left(\frac{1}{CK} - \frac{\alpha}{C} - \frac{\alpha}{K}\right)
\]  

(27)

From this, it follows that

\[
  \lim_{x \to 0^+} k_0(x) = -\infty, \quad \lim_{x \to +\infty} k_0(x) = +\infty
\]

\[
  \lim_{x \to 0^+} k_0'(x) = +\infty, \quad \lim_{x \to +\infty} k_0'(x) = +\infty
\]

(28)

Under some parametric conditions, there exists an interval \([A, B] \subseteq (0, +\infty)\) such that \(k_0'(x)_{|_{(A, B)}} < 0\). Analysis on the shape of the curve \(F(x, k_0) = 0\) shows that, under certain parametric conditions, system (24) goes through a fold catastrophe bifurcation\(^1\)\(^\text{,}^\text{14}\) (belonging to the mutualistic type). See Fig.1 in the main text for more details.

Note that the characteristic returning time \(T\) determined by Eq. (14) becomes

\[
  -\frac{1}{\bar{T}} = \bar{\lambda}_0 = \lambda_0 + \frac{\lambda_0^2 + \lambda_0^2 - 2(\Gamma \cdot F^*_s(x_s, k_0))^2 / \lambda_0}{2}
\]

(29)

where

\[
  \lambda_0 = -1 + 2 \left(\frac{1}{C} + \frac{1}{K}\right)x_s - \frac{3x_s^2}{CK} + \frac{k_0\beta x_s (2 + \alpha x_s)}{(1 + \alpha x_s)^2}
\]

(30a)

\[
  F^*_s(x_s, k_0) = 2 \left(\frac{1}{C} + \frac{1}{K}\right) - \frac{6x_s^2}{CK} + \frac{2k_0\beta}{(1 + \alpha x_s)^3}
\]

(30b)

\[
  \Gamma = \sum_j \left(\sum_{j'} k_{ji}^{out}\right)^2 \left(\sum_{i} a_{ji}^2 \Gamma_{jk} + \Gamma_{ij}\right)
\]

(30c)

with \(x_s\) satisfying \(F(x_s, k_0) = 0\).
C Numerical simulation methods

C.1 Network structure setting

In order to confirm the wide applicability of our new theory, we choose 9 typical and representative complex networks to implement numerical simulations. Among them, 3 networks are chosen from empirical data and the other 6 networks are artificially constructed, which include directed and undirected networks with different degree distributions and weight distributions. See Table S2 for more details.

<table>
<thead>
<tr>
<th>Type</th>
<th>Network ID</th>
<th>Node Number $N$</th>
<th>Mean degree $\langle k \rangle$</th>
<th>Effective parameter $k_0$</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empirical</td>
<td>Undirected</td>
<td>Net1</td>
<td>96</td>
<td>13.5</td>
<td>36.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Net2</td>
<td>276</td>
<td>6</td>
<td>18.2</td>
</tr>
<tr>
<td></td>
<td>Directed</td>
<td>Net3</td>
<td>1550</td>
<td>20.9</td>
<td>11.7</td>
</tr>
<tr>
<td>Artificial</td>
<td>Undirected</td>
<td>Net4</td>
<td>200</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Net5</td>
<td>300</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Net6</td>
<td>200</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Net7</td>
<td>300</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Directed</td>
<td>Net8</td>
<td>200</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Net9</td>
<td>300</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

Table S2. Complex network data used in numerical simulations. Net1~Net3 are empirical networks from experiments in ecology and molecular biology. The empirical data are download from Refs. 21,22, which were also used in the numerical simulation in Ref. 17. Net4~Net9 are representative random models extensively studied in the complex network theory.
For mutualistic networks Net1 and Net2, there is an initial bipartite $n \times m$ matrix $M_{ik}$, where the index $i$ represents a plant and $k$ represents a pollinator from the data of Montane forest and grassland\textsuperscript{23}. By projecting the bipartite networks on their two node sets, the $n \times n$ plant network Net1 and the $m \times m$ pollinator network Net2 are constructed, where the nodes are connected by mutualistic interactions. The weight representing the mutualistic strength between nodes $i$ and $j$ in Net1 is determined by the following method: (i) the more mutual pollinators $k$ that plants $i$ and $j$ share are, the stronger the mutualistic interaction between them is; (ii) the more plants pollinated by $k$ are, the smaller its contribution to each plant is. For more detailed explanations, see Ref. 17. The weight construction for Net2 is similar to that for Net1, both in a symmetric way.

Net3 is the gene regulatory network in E. coli\textsuperscript{21}, a directed network where nodes represent the genes and edges represent the regulatory relations between them. The weights of edges are set initially to $a_{ij} = 10$ for all linked $i, j$ pairs, just for simplicity.

The artificial networks are representative models studied in the complex network theory\textsuperscript{24-26}.

The Net4 (200 nodes) and Net5 (300 nodes) are random graphs (or the so-called Erdos-Renyi model) with mean degrees 5 and 10 (for initial nodes), respectively. The weights of edges in Net4 are all set to be 2 and the weights of edges in Net5 are chosen from a power-law distribution with mean 1 and power-law exponent 2.5.

The Net6 (200 nodes) and Net7 (300 nodes) are scale-free networks (the so-called Barabasi-Albert models) with mean degree 5 and 12 (for initial nodes), respectively, and their degree distributions are both power-law with exponent 2.5. The weights of edges in Net6 are all set to be 2 and the weights of edges in Net7 are chosen from a power-law distribution with mean 1 and power-law exponent 2.6.

The Net8 (200 nodes) and Net9 (300 nodes) are directed scale-free networks with mean (in and out) degrees 5 and 10 (for initial nodes), respectively, and their degree distributions are power laws with exponent 2.5, while Net9 has negative in-out degree correlations (nodes with large out-degree often have small in-degree) and Net8 is the one without in-out degree correlations. The weights of edges of them are set to be 4 and 2, respectively.
C.2 Simulating the effect of random failures

To simulate the effect of random failures or attacks, we adopt the following settings. For random failures, we delete nodes from the network with a fraction $\rho$ randomly (to simulate the internal failures). For random attacks, we randomly reduce a fraction $\rho$ of the weights of the initial network (to simulate the external environmental deteriorations), that is, we let $a_{ij} \rightarrow (1-r)a_{ij}$ for non-zero elements in the adjacency matrix $A$ of the network, where $r$ is chosen from a Beta distribution with mean $\rho$. It is known that the Beta distribution $Be(a, b)$ with parameters $a, b$ has the mean $\frac{a}{a+b}$ and the variance $\frac{ab}{(a+b)^2(a+b+1)}$. Note that this distribution is unimodal for $a > 1, b > 1$. For a fixed $\rho$, one can choose two suitable parameters $a > 1, b > 1$ to construct the needed Beta distribution. Here, we set $\rho = \frac{a}{a+b}$ and the variance can be reduced by increasing $a, b$ simultaneously. For an undirected network, the adjacency matrix is kept symmetric. After the failing, we use the given network data to perform simulation and calculation. When a fraction $\rho$ of the network weights is randomly removed, the network changes its topology in a stochastic way, thus influencing its topological characteristic parameter $k_0$.

By using the given network data $\{a_{ij}\}$, the stochastic differential equations (i.e., Eq. (1) in the main text) can be solved by the Euler method$^{27,28}$. In simulation, we set the system’s parameters as

$$b = 0.1, \quad C = 1, \quad K = 5, \quad n = 2, \quad m = 1,$$
$$\alpha_1 = \beta_1 = 0.01, \quad \alpha_2 = \beta_2 = 0.19,$$

and adjust $\Gamma_i, \Gamma_{ij}$ to obtain a desired noise intensity.

When numerically solving Eq. (1) in the main text, we set two initial conditions: a low initial state value with all $x_i(0) = 0.01$ and a high initial state value with all $x_i(0) = 5$. This allows us to test whether the critical transition takes place. In fact, at the before-transition state, the system has its stationary position around a unique stable state with large magnitudes of components. The appearance of the stable state with a large magnitude implies that the critical transition has occurred.

To get a better representation of the changes in the system’s state, we calculate the average
\[ \langle x \rangle = \sum_{i \in GC} x_i / |GC|, \]
where GC is a giant component of the network. This quantity reflects the
global resilience of the system and can describe whether critical transition occurs. See Fig. 2 in the
main text for more details.

Finally, we calculate the average of the weighted means \( \langle x \rangle_w \) of the network components
and use the stationary state data (from which we can get the stationary average \( E[\langle x \rangle_w] \)) to test
whether the theoretical prediction of the 1-D equation is effective. See Fig. 3a in the main text for
more details.

D Extended validation

D.1 Simulations of artificial networks

In the main text we have shown that the higher-dimensional stochastic dynamic behavior of a
complex networked system can be captured by a 1-dimensional stochastic equations, Eq. (21). The
form of the 1-D equation is solely determined by the initial dynamics of the original large-scale
system, namely \( f(\cdot) \) and \( g(\cdot, \cdot) \) in Eq. (16). The complex networked system has an important
topological characteristic parameter, \( k_0 = \left\langle k^n \right\rangle_w \), which can be calculated based on the network
adjacency matrix \( \{a_{ij}\} \). Note that this parameter preserves the topological structure properties of
the complex network. Note also that different network topological structures give rise to different
dynamic patterns.

To validate our prediction theory, we also conducted a set of numerical tests on artificial
complex networks, which are of different scales, levels, degree differences and structural
correlations. All these simulated complex networks are classified into 6 types. See Net4-9 in Section
C.1 in this Supplementary Information for details.

For the 6 artificial networks, we performed numerical simulations with numerical results shown
in Fig. S4, where the row diagrams from top to bottom correspond to the random graph with uniform
weights, the random graph with power-law weights, the scale-free network with uniform weights,
the scale-free network with power-law weights, the directed scale-free network with uniform
weights and without in-out degree correlations, and the directed scale-free network with uniform
weights and with negative in-out degree correlations, respectively. The column diagrams from left to right correspond to low loss, high loss, high loss with large noise, universal dynamics, and the characteristic returning time of the complex networks, respectively. These numerical results also have well verified our theoretical predictions.

![Diagram](image)

**Figure S4.** Critical transitions in artificial networks with different topological properties. The numerical results of these networks confirm the universality of the indicator that we introduced in the main text and in this Supporting Information. Here, low loss is set as 40% (corresponds to $\rho = 0.4$), high loss is set as 90% (corresponds to $\rho = 0.9$), high noise corresponds to the setting of $\Gamma_{ij} = 0.05, \Gamma_i = 0.1$, the noise in the other cases corresponds to the setting of $\Gamma_{ij} = 0.001, \Gamma_i = 0.01$. Other settings are the same as Fig. 2 and Fig.3 in the main text.
D.2 Effect of noise on the characteristic returning time

We furthermore compare the characteristic returning time in the case considering the noise effect, $T = -\bar{\lambda}_0^{-1}$, and the case without considering the noise effect, $T = -\lambda_0^{-1}$. See section A in this Supplementary Information for more details. It is found that noise has an enhancing effect on, or has a positive contribution to $T$, remarkably when the system is close to the critical transition state, referring to Fig. S5. In a word, the index $T$ incorporating the noise effects is a good indicator for the critical transitions in the networks since noise can amplify the critical signals.

![Graph](image)

**Figure S5.** An example showing the effect of noise on the characteristic returning time $T$ (which has been magnified by 10 times), where dashed(blue)/solid(red) curves represent the characteristic returning time in the cases with/without considering the noise effects. We enhance the loss fraction of the network Net1 and calculate the corresponding characteristic returning time. It can be seen that noise has a remarkable effect on $T$ when the system is close to a critical state. The parametric settings are the same as Fig. 3 in the main text.

References


