Supplemental Materials

S1. Approximation Fokker-Planck Equation (aFPE) for General Langevin Equation

For general Langevin equation [1]:

$$\frac{dx}{dt} = f(x) + g_1(x)\varepsilon(t) + g_2(x)\eta(t) \tag{S1}$$

where $\varepsilon(t)$ and $\eta(t)$ are all colored noise, with statistical characters

$$\begin{aligned}
&\langle \varepsilon(t) \rangle = \langle \eta(t) \rangle = 0 \\
&\gamma_{11}(t,t') = \left\langle \varepsilon(t)\varepsilon\left(t'\right) \right\rangle = \frac{D}{\tau_1}\exp\left(-\frac{|t-t'|}{\tau_1} \right) \\
&\gamma_{22}(t,t') = \left\langle \eta(t)\eta\left(t'\right) \right\rangle = \frac{\alpha}{\tau_2}\exp\left(-\frac{|t-t'|}{\tau_2} \right) \\
&\gamma_{12}(t,t') = \gamma_{21}(t,t') = \left\langle \varepsilon(t)\eta\left(t'\right) \right\rangle \\
&= \left\langle \eta(t)\varepsilon\left(t'\right) \right\rangle = \frac{\lambda D\alpha}{\tau_3}\exp\left(-\frac{|t-t'|}{\tau_3} \right)
\end{aligned} \tag{S2}$$

where τ_1 and D are ML and NI of the multiplicative noise $\varepsilon(t)$, respectively. τ_2 and α are ML and NI of the additive noise $\eta(t)$, respectively. τ_3 and λ denote the memory and intensity of cross-correlation noise between $\varepsilon(t)$ and $\eta(t)$.

According to the stochastic Liouville equation [2] corresponding to Equation (S1) and van Kampen's lemma[3]: $P(x,t) =< \delta(x(t)-x) >$, the evolution equation for the probability density $P(x,t)$ is obtained:

$$\begin{aligned}
&\frac{\partial}{\partial t}P(x,t) = -\frac{\partial}{\partial x}f(x)P(x,t) \\
&-\frac{\partial}{\partial x}g_1(x)\left\langle \varepsilon(t)\delta(x(t)-x) \right\rangle - \frac{\partial}{\partial x}g_2(x)\left\langle \eta(t)\delta(x(t)-x) \right\rangle
\end{aligned} \tag{S3}$$

To calculate the average $\left\langle \varepsilon(t)\delta(x(t)-x) \right\rangle$ and $\left\langle \eta(t)\delta(x(t)-x) \right\rangle$ in Equation (S3), we use the Novikov theorem[4-5]:

$$\left\langle \xi_k \Phi\left[\xi_1,\xi_2\right] \right\rangle = \int_0^t \mathrm{d}t' \gamma_{k,l} \left\langle \frac{\partial \Phi\left[\xi_1,\xi_2\right])}{\delta \xi_l} \right\rangle \tag{S4}$$

where $\Phi\left[\xi_1,\xi_2\right]$ is a function of ξ_1,ξ_2 , $\gamma_{k,l}$ is the correlation function between them. From Equation (S4), we can get [6]

$$
\begin{aligned}
\langle \varepsilon(t)\delta(x(t)-x)\rangle = & -\int_0^t \mathrm{d}t' \gamma_{11}(t,t') \left\langle \frac{\delta(\delta(x(t)-x))}{\delta x} \frac{\delta x}{\delta \varepsilon(t')} \right\rangle \\
& -\int_0^t \mathrm{d}t' \gamma_{12}(t,t') \left\langle \frac{\delta(\delta(x(t)-x))}{\delta x} \frac{\delta x}{\delta \eta(t')} \right\rangle \\
= & -\frac{\partial}{\partial x} \int_0^t \mathrm{d}t' \gamma_{11}(t,t') \left\langle \delta(x(t)-x) \frac{\delta x}{\delta \varepsilon(t')} \right\rangle \\
& -\frac{\partial}{\partial x} \int_0^t \mathrm{d}t' \gamma_{12}(t,t') \left\langle \delta(x(t)-x) \frac{\delta x}{\delta \eta(t')} \right\rangle
\end{aligned}
\tag{S5}
$$

After integrating equation (S1), we have

$$
x(t) = x(0) + \int_0^t \mathrm{d}s \left[f(x(s)) + g_1(x(s))\varepsilon(s) + g_2(x(s))\eta(s) \right] \tag{S6}
$$

The first $\varepsilon(t')$ functional-derivative of equation (S6) is

$$
\frac{\delta x(t)}{\partial \varepsilon(t)} = g_1\left(x\left(t'\right)\right) + \int_{t'}^t \mathrm{d}s \left\{ \left[f'(x(s)) + g_1'(x(s))\eta(s) + g_2'(x(s))\varepsilon(s) \right] \frac{\delta x(s)}{\delta \varepsilon\left(t'\right)} \right\} \tag{S7}
$$

Here $t > t'$, the derivative of the above formula with respect to evolution-time t is

$$
\frac{\partial}{\partial t} \left[\frac{\delta x(t)}{\delta \varepsilon\left(t'\right)} \right] = \left[f'(x(t)) + g_1'(x(t))\varepsilon(t) + g_2'(x(t))\eta(t) \right] \frac{\delta x(t)}{\delta \varepsilon(t)} \tag{S8}
$$

The initial condition in solving the above derivative equation is given by

$$
\left. \frac{\delta x(t)}{\delta \varepsilon(t)} \right|_{t'=t} = g_1\left(x\left(t'\right)\right)
$$

Combining the above condition with equation (S8) and performing time integration, we obtain

$$
\frac{\delta x(t)}{\delta \varepsilon(t)} = g_1\left(x\left(t'\right)\right) \exp\left(\int_{t'}^t \mathrm{d}s \left[f'(x(s)) + g_1'(x(s))\varepsilon(s) + g_2'(x(s))\eta(s) \right] \right) \tag{S9}
$$

In addition, the derivative of $g_1(x(t))$ to evolution-time t reads as

$$
\frac{\mathrm{d}}{\mathrm{d}t} g_1(x(t)) = g_1'(x(t)) \frac{\mathrm{d}x(t)}{\mathrm{d}t} = \frac{g_1'(x(t))}{g_1(x(t))} \mathrm{f}(x(t)) + g_1(x(t))\varepsilon(t) + g_2(x(t))\eta(t) \Big] g_1(x(t)) \tag{S10}
$$

The integral of equation (S10) gives

$$
g_1\left(x\left(t'\right)\right) = g_1(x(t)) \exp\left(\int_t^{t'} \mathrm{d}s \left[\frac{g_1'(x(s))}{g_1(x(s))} \left[f(x(s)) + g_1(x(s))\varepsilon(s) + g_2(x(s))\eta(s) \right] \right) \right) \tag{S11}
$$

Inserting equation (S11) into equation (S9), we get

$$
\frac{\delta x(t)}{\delta \varepsilon(t)} = g_1(x(t)) \exp\left(\int_{t'}^t \mathrm{d}s \left[f'(x(s)) - \frac{g_1'(x(s))}{g_1(x(s))} f(x(s)) + g_2'(x(s))\eta(s) - \frac{g_1'(x(s))}{g_1(x(s))} g_2(x(s))\eta(s) \right] \right) \tag{S12}
$$

In a similar fashion,

$$\frac{\delta x(t)}{\delta \eta(t)} = g_2(x(t)) \exp\left(\int_{t'}^{t} \mathrm{d}s \left[f'(x(s)) - \frac{g_2'(x(s))}{g_2(x(s))} f(x(s)) + g_1'(x(s))\varepsilon(s) - \frac{g_2'(x(s))}{g_2(x(s))} g_1(x(s))\varepsilon(s) \right] \right) \tag{S13}$$

Under the condition, $t >> \tau_1, \tau_2, \tau_3$, inserting Equations (S12) and (S13) into equation (S15), we can get

$$\begin{aligned}
\langle \varepsilon(t)\delta(x(t)-x)\rangle &= -\frac{\partial}{\partial x} g_1(x) \int_0^t \mathrm{d}t' \gamma_{11}(t,t') \left\langle \delta(x(t)-x) \exp\left(\left(f'(x_s) - \frac{g_1'(x_s)}{g_1(x_s)} f(x_s) \right)(t-t') \right) \right\rangle \\
&\quad -\frac{\partial}{\partial x} g_1(x) \int_0^t \mathrm{d}t' \gamma_{12}(t,t') \left\langle \delta(x(t)-x) \exp\left(\left(f'(x_s) - \frac{g_2'(x_s)}{g_2(x_s)} f(x_s) \right)(t-t') \right) \right\rangle \\
&\approx \frac{D}{1-\tau_1\left[f'(x_s) - f(x_s) g_1'(x_s)/g_1(x_s) \right]} \frac{\partial}{\partial x} g_1(x) P(x,t) \\
&\quad -\frac{\lambda\sqrt{D\alpha}}{1-\tau_3\left[f'(x_s) - f(x_s) g_2'(x_s)/g_2'(x_s) \right]} \frac{\partial}{\partial x} g_2(x) P(x,t)
\end{aligned} \tag{S14}$$

In a similar fashion

$$\begin{aligned}
\langle \eta(t)\delta(x(t)-x)\rangle &= -\frac{\partial}{\partial x} g_2(x) \int_0^t \mathrm{d}t' \gamma_{22}(t,t') \left\langle \delta(x(t)-x) \exp\left(\left(f'(x_s) - \frac{g_2'(x_s)}{g_2(x_s)} f(x_s) \right)(t-t') \right) \right\rangle \\
&\quad -\frac{\partial}{\partial x} g_2(x) \int_0^t \mathrm{d}t' \gamma_{21}(t,t') \left\langle \delta(x(t)-x) \exp\left(\left(f'(x_s) - \frac{g_1'(x_s)}{g_1(x_s)} f(x_s) \right)(t-t') \right) \right\rangle \\
&\approx \frac{D}{1-\tau_2\left[f'(x_s) - f(x_s) g_2'(x_s)/g_2(x_s) \right]} \frac{\partial}{\partial x} g_2(x) P(x,t) \\
&\quad -\frac{\lambda\sqrt{D\alpha}}{1-\tau_3\left[f'(x_s) - f(x_s) g_1'(x_s)/g_1'(x_s) \right]} \frac{\partial}{\partial x} g_1(x) P(x,t)
\end{aligned} \tag{S15}$$

Here, according to Hanggi Ansatz, x_s denotes the steady-state value of deterministic theory[7]. $f'(x)$, $g_1'(x)$ and $g_2'(x)$ denote the first x derivatives of f, g_1 and g_2.

$$\begin{aligned}
&\frac{\partial}{\partial t} P(x,t) = -\frac{\partial}{\partial t} f(x) P(x,t) \\
&+\frac{D}{1-\tau_1\left[f'(x_s) - f(x_s) g_1'(x_s)/g_1(x_s) \right]} \frac{\partial}{\partial x} g_1(x) \frac{\partial}{\partial x} g_1(x) P(x,t) \\
&+\frac{\lambda\sqrt{D\alpha}}{1-\tau_3\left[f'(x_s) - f(x_s) g_2'(x_s)/g_2(x_s) \right]} \frac{\partial}{\partial x} g_1(x) \frac{\partial}{\partial x} g_2(x) P(x,t) \\
&+\frac{\alpha}{1-\tau_2\left[f'(x_s) - f(x_s) g_2'(x_s)/g_2(x_s) \right]} \frac{\partial}{\partial x} g_2(x) \frac{\partial}{\partial x} g_2(x) P(x,t) \\
&+\frac{\lambda\sqrt{D\alpha}}{1-\tau_3\left[f'(x_s) - f(x_s) g_1'(x_s)/g_1(x_s) \right]} \frac{\partial}{\partial x} g_2(x) \frac{\partial}{\partial x} g_1(x) P(x,t)
\end{aligned} \tag{S16}$$

This AFPE is valid for

$$1-\tau_i\left[f'(x_s)-f(x_s)g_j'(x_s)/g_j(x_s)\right]>0 \tag{S17}$$

and $i=1,2,3$, $j=1,2$. Equation (S17) limits the range of possible values of the correlation times.

S2. Linear Mapping Approximation (LMA)

At this point, the entire nonlinear complex network can be described as the following equation system:

$$\begin{aligned} &off \xrightarrow{\bar{\sigma}_b} on,\ on+p \xrightarrow{\sigma_u} off+p,\\ &on \xrightarrow{\rho_b} on+p, off \xrightarrow{\rho_u} off+p\\ &p \xrightarrow{\delta} \varnothing \end{aligned} \tag{S18}$$

Here, the above nonlinear GRN can be transformed into linear GRN by removing the second-order reaction between the protein and the state off [8,9]. Specifically, this is refactoring the reaction $off+P \rightleftharpoons on+P$ to $off \rightleftharpoons on$. Note that all the parameters between the two models are the same except for σ_b. Then how to choose the free parameter $\bar{\sigma}_b$ in the linear GRN under the given nonlinear GRN with certain parameters, so that the solution of the system is very close to the original nonlinear GRN solution?

Next, the mean-field hypothesis of LMA is used to determine the value of the effective parameter $\bar{\sigma}_b$, and the linear moment equation below the third order can be obtained:

$$\mathcal{M}_{\mathrm{FL}}(\bar{\sigma}_b):\begin{cases} \partial_t\langle n_p\rangle=\rho_u\langle n_g\rangle+\rho_b\left(1-\langle n_g\rangle\right)-\langle n_p\rangle,\\ \partial_t\langle n_g\rangle=-\bar{\sigma}_b\langle n_g\rangle+\sigma_u\left(1-\langle n_g\rangle\right),\\ \partial_t\langle n_p^2\rangle=2(\rho_u-\rho_b)\langle n_p n_g\rangle+(2\rho_b+1)\langle n_p\rangle-2\langle n_p^2\rangle+(\rho_u-\rho_b)\langle n_g\rangle+\rho_b,\\ \partial_t\langle n_p n_g\rangle=\rho_u\langle n_g\rangle+\sigma_u\langle n_p\rangle-(1+\bar{\sigma}_b+\sigma_u)\langle n_p n_g\rangle,\\ \partial_t\langle n_p^2 n_g\rangle=(2\rho_u+1)\langle n_p n_g\rangle-(\sigma_u+\bar{\sigma}_b+2)\langle n_p^2 n_g\rangle+\rho_u\langle n_g\rangle+\sigma_u\langle n_p^2\rangle. \end{cases} \tag{S19}$$

Note that if $\bar{\sigma}_b$ is given, then $\mathcal{M}_{\mathrm{FL}}(\bar{\sigma}_b)$ is a closed set of ordinary differential equations. LMA parameterizes $\bar{\sigma}_b$ by means of conditional averaging. Specifically, for a non-cooperative nonlinear feedback loop, if the nonlinear response is $off+P \rightleftharpoons on+P$, then $\bar{\sigma}_b=\sigma_b\langle n_p|n_g=1\rangle=\sigma_b\langle n_p n_g\rangle/\langle n_g\rangle$, and this is using the fact that n_g is a Boolean variable.

In the steady-state case, the solution is simpler. When the time derivative is zero, the moment equation can be solved, that is, $\mathcal{M}_{\mathrm{FL}}\left(\bar{\sigma}_b=\sigma_b\langle n_p n_g\rangle/\langle n_g\rangle\right)=0$, and an explicit expression of $\langle n_g\rangle\langle n_p n_g\rangle$ can be obtained. Thus, the effective parameter $\bar{\sigma}_b=\sigma_b\langle n_p n_g\rangle/\langle n_g\rangle$ in the non-collaborative feedback loop can be obtained as:

$$\bar{\sigma}_b=\frac{-1+\rho_b\sigma_b-\sigma_u+\sqrt{(1-\rho_b\sigma_b+\sigma_u)^2+4\rho_u\sigma_b(1+\sigma_u)}}{2} \tag{S20}$$

Therefore, the whole equation system is simplified into the following form:

$$
\begin{aligned}
&off \xrightarrow{\bar{\sigma}_b} on,\ on \xrightarrow{\sigma_u} off,\\
&on \xrightarrow{\rho_b} on + p, off \xrightarrow{\rho_u} off + p\\
&p \xrightarrow{\delta} \varnothing
\end{aligned}
\tag{S21}
$$

Next, $\bar{\sigma}_b$ in the equation is replaced by $\bar{\sigma}_b$ to obtain the approximate probability distribution of the system at the time t, namely:

$$
P^*\left(n_p,t\right)=\frac{1}{n_p!}\frac{\mathrm{d}^{n_p}}{\mathrm{d}w^{n_p}}\left(G_0^*\left(w,t\right)+G_1^*\left(w,t\right)\right)\Big|_{w=-1} \tag{S22}
$$

where

$$
\begin{aligned}
G_0^*\left(w,t\right)=\exp\left(\rho_b w\right)\Big[& f\left(we^{-t}\right)\left(-\rho_\Delta w\right)^{1-\Sigma} M\left(1-\bar{\sigma}_b^*,2-\Sigma,-\rho_\Delta w\right)\\
&+\mathrm{g}\left(we^{-t}\right)M\left(1+\sigma_u,\Sigma,-\rho_\Delta w\right)\Big]
\end{aligned}
\tag{S23}
$$

$$
\begin{aligned}
G_1^*\left(w,t\right)=\sigma_u^{-1}\exp\left(\rho_b w\right)\Big[& -\sigma_u f\left(we^{-t}\right)\left(-\rho_\Delta w\right)^{1-\Sigma} M\left(-\bar{\sigma}_b^*,2-\Sigma,-\rho_\Delta w\right)\\
&+\bar{\sigma}_b\mathrm{g}\left(we^{-t}\right)M\left(\sigma_u,\Sigma,-\rho_\Delta w\right)\Big]
\end{aligned}
\tag{S24}
$$

and $G(w)=\sum_{n=0}^{\infty}P(n)w^n=\sum_{n=0}^{\infty}P_0(n)w^n+\sum_{n=0}^{\infty}P_1(n)w^n=G_0(w)+G_1(w)$ are all generation functions of the corresponding probability distribution, respectively.

S3. Analysis of Sensitivity Effect

In complex environments, the life development of phage λ is regulated by changeable parameters NI/AI and ML/AM (i.e., $D,\alpha,\lambda,\tau_1,\tau_2$ and τ_3) to achieve phase switching or phenotype diversity. Therefore, it is necessary to evaluate the sensitivity effects of changeable parameters on state switching. Therefore, we define the so-called sensitivity index as $\varphi=\mu_1*\frac{1}{MFPT}+\mu_2*\mathrm{Z}$, here μ_1,μ_2 represents the weight coefficient and satisfies the conservation condition $\mu_1+\mu_2=1$, Z is inverse of the variable magnitudes encompassed within the probabilistic distribution region, and MFPT represents the mean first passage time. To obtain the sensitivity index, it is important to calculate the components at the same levels, respectively.

Considering the probabilistic distribution region for a changeable parameter Z, referring to Figure 5 in the main text and Figure S1, the following table is given.

Parameter	Relative Fold	Sensitivity (Maximal Change Relative)
D	$10.321\to7.155\to4.762\to6.868\to10.318$	10.321
α	$5.0545\to5.0485\to5\to5.005\to5.054$	5.0545
λ	$1.25\to1.37\to1.52\to1.39\to1.25$	1.52
τ_1	$1.11\to0.81\to0.67\to0.77\to1.1$	1.11
τ_2	$2.04\to2.06\to2.16\to2.14\to2.04$	2.16
τ_3	$0.22\to0.21\to0.2\to0.21\to0.22$	0.22

For MFPT, we consider the change rate of its relative fold, that is, MFPT is going to be divided by the minimum value of every trajectory, referring to Figure S2. Here, we take the maximal change relative fold to be the sensitivity effect, as following table.

Parameter	Relative Fold	Sensitivity (Maximal Change RElative)
D	$4 \to 7.4328 \to 8.4168$	8.4168
α	$4.082 \to 6.507 \to 8.498$	8.498
λ	$5 \to 5.8 \to 6.52$	6.52
τ_1	$1.93 \to 1.67 \to 1.25$	1.93
τ_2	$3.33 \to 3.46 \to 3.622$	3.622
τ_3	$1.49 \to 1.43 \to 1.01$	1.49

Therefore, if the weight coefficients are all equal to each other($\mu_1 = \mu_2 = 1/2$), we can obtain the absolute value of sensitivity effect for every changeable parameter as following table.

Parameter	Sensitivity (Absolute)	Sensitivity (Relative Fold)	Parameter	Sensitivity (Absolute)	Sensitivity (Relative Fold)
D	9.3689	10.96	τ_1	1.52	1.78
α	6.7763	7.93	τ_2	2.891	3.38
λ	4.02	4.70	τ_3	0.855	1

S4. Supplementary Graphics

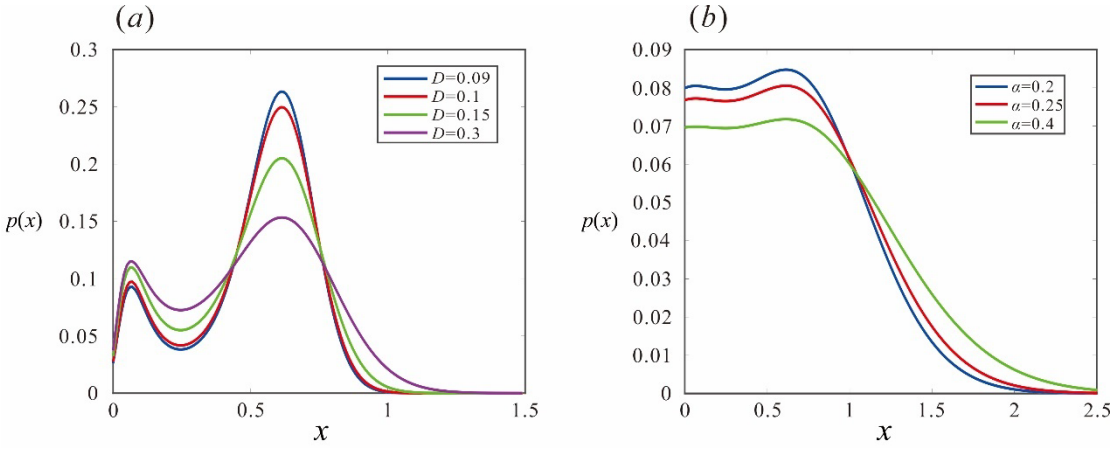


Figure S1. The phenotype switching is regulated by the NI(D/α). (**a**) NI(D) reshapes the cell phenotype; (**b**) NI(α) regulates the cell phenotype. Other parameters are the same as Figure 5 in the main text.

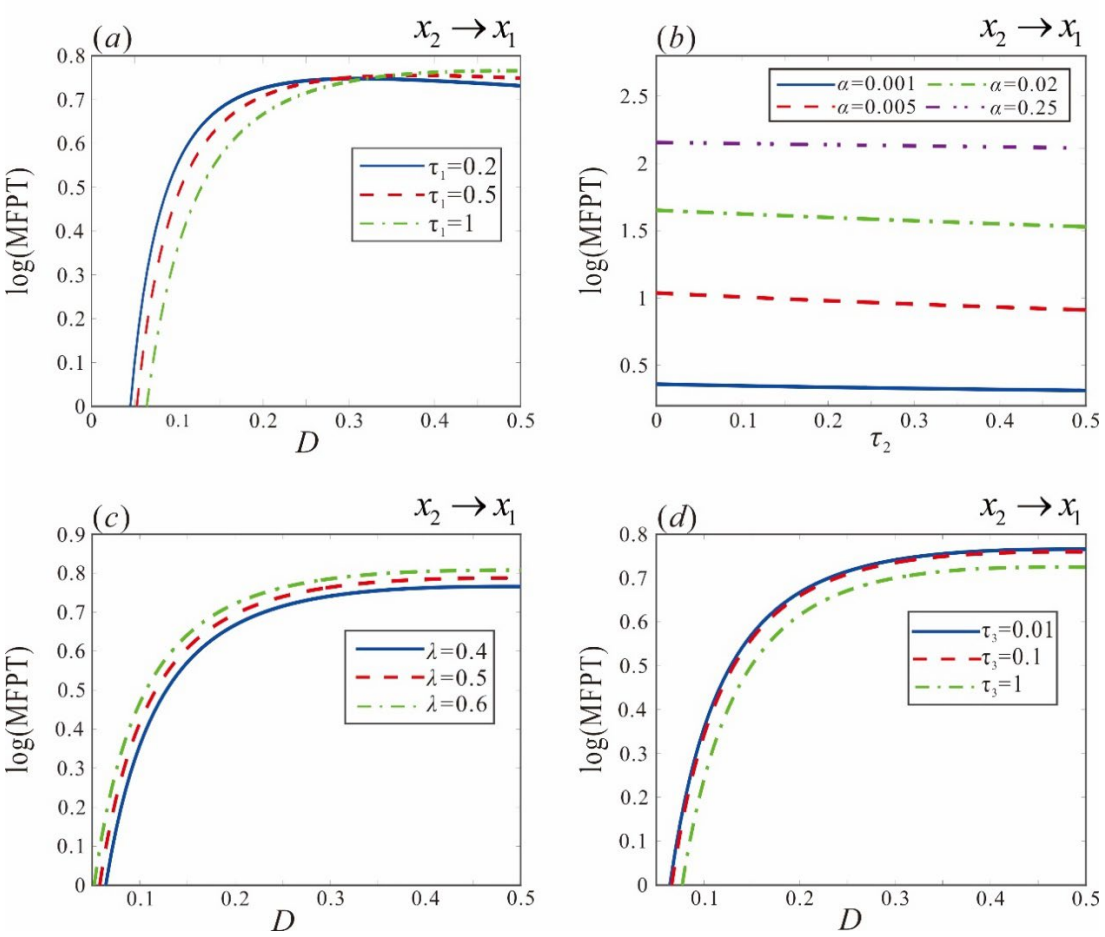


Figure S2. The MFPT from the high steady-state switching to the low steady-state is regulated by different characteristic parameters. (**a**,**c**–**d**) the MFPT is as the function of NI(D) with ML/AM(τ_1 / τ_3) and AI(λ) changing;(**b**) MFPT is as the function of ML(τ_2) with NI(α). Other parameters are the same as Figure 5 in the main text.

References

1. Liang, G.Y.; Cao, L.; D. Wu, J. Approximate Fokker–Planck equation of system driven by multiplicative colored noises with colored cross-correlation. *Phys. A-Stat. Mech.* **2004**, *335*, 371–384.
2. Chen, F.F. Introduction to Kinetic Theory. In *Introduction to Plasma Physics*; Springer: Boston, MA, USA, 1995; Volume 7, pp. 199–239.
3. Shenton, L.R.; Bowman, K.O. A bivariate model for the distribution of b1 and b2. *J. Am. Stat. Assoc.* **1977**, *72*, 206–211.
4. Frank, T.D. Delay Fokker-Planck equations, Novikov's theorem, and Boltzmann distributions as small delay approximations. *Phys. Rev. E* **2005**, *72*, 11112.
5. Zhu, P. Associated relaxation time and intensity correlation function of a bistable system driven by cross-correlation additive and multiplicative colored noise sources. *Eur. Phys. J. B* **2007**, *55*, 447–452.
6. Wu, D.J.; Li, C.; Kang, S.Z. Bistable kinetic model driven by correlated noises: Steady-state analysis. *Phys Rev E.* **1994**, *4*, 2496.
7. Hanggi, P.; Mroczkowski, T.J.; Moss, F.; Mcclintock, P. Bistability driven by colored noise: theory and experiment. *Phys. Rev. A Gen. Phys.* **1985**, *32*, 695–698.
8. Cao, Z.; Grima, R. Linear Mapping Approximation of Gene Regulatory Networks with Stochastic Dynamics. *Nat. Commun.* **2018**, *9*, 3305.
9. Liu, P.; Wang, H.; Huang, L.; Zhou, T. The dynamic mechanism of noisy signal decoding in gene regulation. *Sci. Rep.* **2017**, *7*, 42128–42128.