Supplementary Information for Stochastic effects as a force to increase the complexity of signaling networks

Hiroyuki Kuwahara[1] and Xin Gao[1*]

1 Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, 23955-6900, Saudi Arabia

∗ Corresponding author (email: xin.gao@kaust.edu.sa)

S1 Steady state behavior of the enzymatic circuit motif

Our enzymatic circuit motif has the following set of reactions:

$$\begin{aligned} X &\xrightarrow{k_1} X^* \\ X^* &\xrightarrow{k_2} X \\ X + E_f &\xrightarrow{k_f} X^* + E_f \\ X^* + E_b &\xrightarrow{k_b} X + E_b. \end{aligned}$$

The time evolution of X^* in the mass-action kinetic model of this motif can be expressed in the follows ordinary differential equation:

$$\frac{dX^*(t)}{dt} = k_1 X(t) + k_f X(t) E_f(t) - k_2 X^*(t) - k_b X^*(t) E_b(t). \tag{S1}$$

Solving this for the steady state by setting the left-hand side to be 0 gives the expression of X^*, the deterministic steady state of X^{ds*} as follows:

$$X^{ds*} = \frac{(k_1 + k_f E_f^{ds}) X_{tot}}{k_1 + k_f E_f^{ds} + k_2 + k_b E_b^{ds}}, \tag{S2}$$

where $X_{tot} = X + X^*$ is the total molecular count of the protein and E_f^{ds} and E_b^{ds} are the steady state of E_f and E_b, respectively.

The stochastic model of the motif is based on stochastic chemical kinetics. Given $X^*(t) = x^*$, $X(t) = x$, $E_f(t) = e_f$, and $E_b(t) = e_b$, $X^*(t + dt)$ is expressed as

$$X^*(t + dt) = x^* + \Xi(dt; x^*, x, e_f, e_b), \tag{S3}$$

where $\Xi(dt; x^*, x, e_f, e_b)$ is a random variable with density function $\Pi(v \mid dt; x^*, x, e_f, e_b)$:

$$\Pi(v \mid dt; x^*, x, e_f, e_b) = \begin{cases} (k_1 x + k_f e_f x)dt & \text{for } v = 1, \\ (k_2 x^* + k_b e_b x^*)dt & \text{for } v = -1, \\ 1 - [(k_1 x + k_f e_f x + k_2 x^* + k_b e_b x^*]dt & \text{for } v = 0. \end{cases} \tag{S4}$$

We can then express the mean time evolution of $X^*(t)$ as

$$\langle X^*(t + dt)\rangle = \langle X^*(t)\rangle + \langle \Xi(dt; X^*(t), X(t), E_f(t), E_b(t))\rangle. \tag{S5}$$

$$\begin{aligned} \langle \Xi(dt; X^*(t), X(t), E_f(t), E_b(t))\rangle &= 1 * \left[k_1 \sum_{x \in X} x P_X(x) + k_f \sum_{\substack{x \in X(t) \\ e_f \in E_f(t)}} e_f x P_{X, E_f}(x, e_f) \right] dt \\ &\quad + -1 * \left[k_2 \sum_{x^* \in X^*(t)} x^* P_{X^*}(x^*) + k_b \sum_{\substack{x^* \in X^*(t) \\ e_f \in E_b(t)}} e_b x^* P_{X^*, E_b}(x^*, e_b) \right] dt \\ &= \left[k_1 \langle X(t)\rangle + k_f \langle X(t) E_f(t)\rangle - k_2 \langle X^*(t)\rangle - k_b \langle X^*(t) E_b(t)\rangle \right] dt \end{aligned} \tag{S6}$$

where $P_X(x)$ is the probability that $X(t) = x$, $P_{X,E_f}(x, e_f)$ is the joint probability that $X(t) = x$ and $E_f(t) = e_f$, $P_X(x)$ is the probability that $X^*(t) = x^*$, and $P_{X^*,E_b}(x^*, e_b)$ is the joint probability that $X^*(t) = x^*$ and $E_f(t) = e_f$. From Eqs. S4 and S6, we can obtain the average behavior of $X^*(t)$ as

$$\begin{aligned}\frac{d\langle X^*(t)\rangle}{dt} &= k_1\langle X(t)\rangle + k_f\langle X(t)E_f(t)\rangle - k_2\langle X^*(t)\rangle - k_b\langle X^*(t)E_b(t)\rangle \\ &= k_1\langle X(t)\rangle + k_f\langle X(t)\rangle\langle E_f(t)\rangle - k_2\langle X^*(t)\rangle - k_b\langle X^*(t)\rangle\langle E_b(t)\rangle \\ &\quad + k_f Cov(X(t), E_f(t)) - k_b Cov(X^*(t), E_b(t)).\end{aligned} \tag{S7}$$

Since $Cov(X(t), E_f(t)) = Cov(X_{tot} - X^*(t), E_f(t)) = -Cov(X^*(t), E_f(t))$, the steady state of $\langle X^*(t)\rangle$ becomes

$$\langle X^{ss*}\rangle = X^{ds*} - \frac{k_f Cov(X^{ss*}, E_f^{ss}) + k_b Cov(X^{ss*}, E_b^{ss})}{k_1 + k_f\langle E_f^{ss}\rangle + k_2 + k_b\langle E_b^{ss}\rangle}. \tag{S8}$$

Since the molecular counts of $E_f(t)$ and $E_b(t)$ do not change in this motif in isolation, both $Cov(X^{ss*}, E_f^{ss})$ and $Cov(X^{ss*}, E_b^{ss})$ become 0. Thus, we have $\langle X^{ss*}\rangle = X^{ds*}$. This also means here that when there are only very small correlation between X^{ss*} and E_f^{ss} and between X^{ss*} and E_b^{ss},

S2 Enzymatic circuit motif with Michaelis-Menten kinetics

With enzymatic reaction kinetics, the reaction to activate X^* has the following rate function:

$$\frac{(k_1 + k_f E_f)X}{1 + K_E X} \tag{S9}$$

whereas the catalytic reaction to inhibit X^* has the following rate function:

$$\frac{(k_2 + k_b E_b)X^*}{1 + K_E X^*}. \tag{S10}$$

Here, both k_1 and k_2 were set to be 0.01 to make the basal rates small. To have the half velocity at the molecular count of 25, we set K_E to be 1/25. We chose to have the value of k_f and k_b to be 20/25 or 100/25 to make the affinities of the protein binding comparable and to keep their ratio the same as the reaction kinetics used in the main text.

S3 A system of equations for N4A

The 4-node network N4A has the following list of reactions:

$$
\begin{aligned}
R_1 &: X_1 \xrightarrow{k_1} X_1^* \\
R_2 &: X_1 + X_2^* \xrightarrow{k_{f1}} X_1^* + X_2^* \\
R_3 &: X_1^* \xrightarrow{k_2} X_1 \\
R_4 &: X_1^* + X_3^* \xrightarrow{k_{b1}} X_1 + X_3^* \\
R_5 &: X_2 \xrightarrow{k_1} X_2^* \\
R_6 &: X_2 + E_f^2 \xrightarrow{k_{f2}} X_2^* + E_f^2 \\
R_7 &: X_2^* \xrightarrow{k_2} X_2 \\
R_8 &: X_2^* + X_4^* \xrightarrow{k_{b2}} X_2 + X_4^* \\
R_9 &: X_3 \xrightarrow{k_1} X_3^* \\
R_{10} &: X_3 + E_f^3 \xrightarrow{k_{f3}} X_3^* + E_f^3 \\
R_{11} &: X_3^* \xrightarrow{k_2} X_3 \\
R_{12} &: X_3^* + E_b^3 \xrightarrow{k_{b3}} X_3 + E_b^3 \\
R_{13} &: X_4 \xrightarrow{k_1} X_4^* \\
R_{14} &: X_4 + E_f^4 \xrightarrow{k_{f4}} X_4^* + E_f^4 \\
R_{15} &: X_4^* \xrightarrow{k_2} X_4 \\
R_{16} &: X_4^* + E_b^4 \xrightarrow{k_{b4}} X_4 + E_b^4
\end{aligned}
$$

A system of these reactions can be modeled via stochastic chemical kinetics, and the following system of equations can be derived for the mean time evolution of the model:

$$
\begin{aligned}
\frac{d\langle X_1^*(t)\rangle}{dt} &= k_1\langle X_1(t)\rangle + k_{f1}\langle X_1(t)\cdot X_2^*(t)\rangle - k_2\langle X_1^*(t)\rangle - k_{b1}\langle X_1^*(t)\cdot X_2^*(t)\rangle, \\
\frac{d\langle X_2^*(t)\rangle}{dt} &= k_1\langle X_2(t)\rangle + k_{f2}\langle X_2(t)\cdot E_f^2(t)\rangle - k_2\langle X_2^*(t)\rangle - k_{b2}\langle X_2^*(t)\cdot X_4^*(t)\rangle, \\
\frac{d\langle X_3^*(t)\rangle}{dt} &= k_1\langle X_3(t)\rangle + k_{f3}\langle X_3(t)\cdot E_f^3(t)\rangle - k_2\langle X_3^*(t)\rangle - k_{b3}\langle X_3^*(t)\cdot E_b^3(t)\rangle, \\
\frac{d\langle X_4^*(t)\rangle}{dt} &= k_1\langle X_4(t)\rangle + k_{f4}\langle X_4(t)\cdot E_f^4(t)\rangle - k_2\langle X_4^*(t)\rangle - k_{b4}\langle X_4^*(t)\cdot E_b^4(t)\rangle,
\end{aligned}
$$

Supporting Tables

Table S1. Parameter/network combinations with 17 highest deviation levels in 4-node networks.

N4A					N4B				
N1[a]	N2	N3	N4	Deviation	N1	N2	N3	N4	Deviation
3[b]	3	2	2	1.803313	3	2	3	2	1.65566
0	3	2	2	0.8921112	2	2	1	2	0.8874556
2	1	2	2	0.8807242	0	2	3	2	0.8437993
3	0	2	2	0.7959255	3	2	0	2	0.7887775
3	1	1	2	0.7479988	3	2	1	1	0.7293608
3	1	2	1	0.692863	3	1	1	2	0.6925644
1	3	1	2	0.6299344	1	2	2	2	0.6332205
1	2	2	2	0.5883827	1	1	3	2	0.6310045
					1	2	3	1	0.5198858

[a]NX means node X. For example, N1 means node 1.

[b]Here, the values of N1-4 indicate parameter combinations. 0 means $k_f = 1, k_b = 1$; 1 means $k_f = 1, k_b = 5$; 2 means $k_f = 5, k_b = 1$; and 3 means $k_f = 5, k_b = 5$.

Supporting Figures

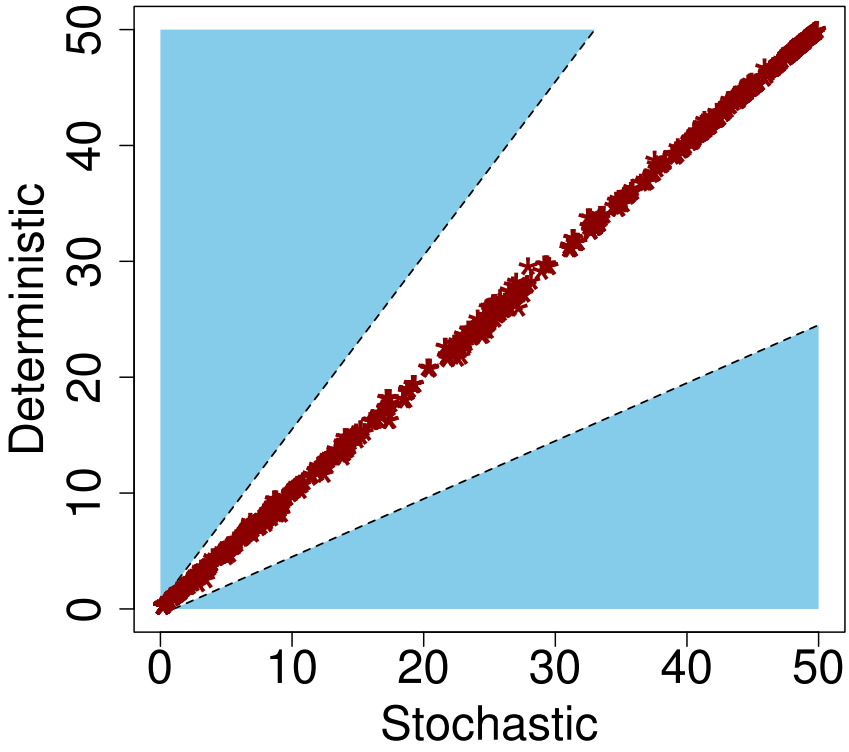

Figure S1. A scatter plot showing the stochastic steady state (x-axis) and the deterministic steady state (y-axis) of each network/parameter combinations in the three-node networks. The total number of network/parameter combinations is 1,664. The area colored in sky blue represents the region where the deviation level is greater than or equal to 0.5.

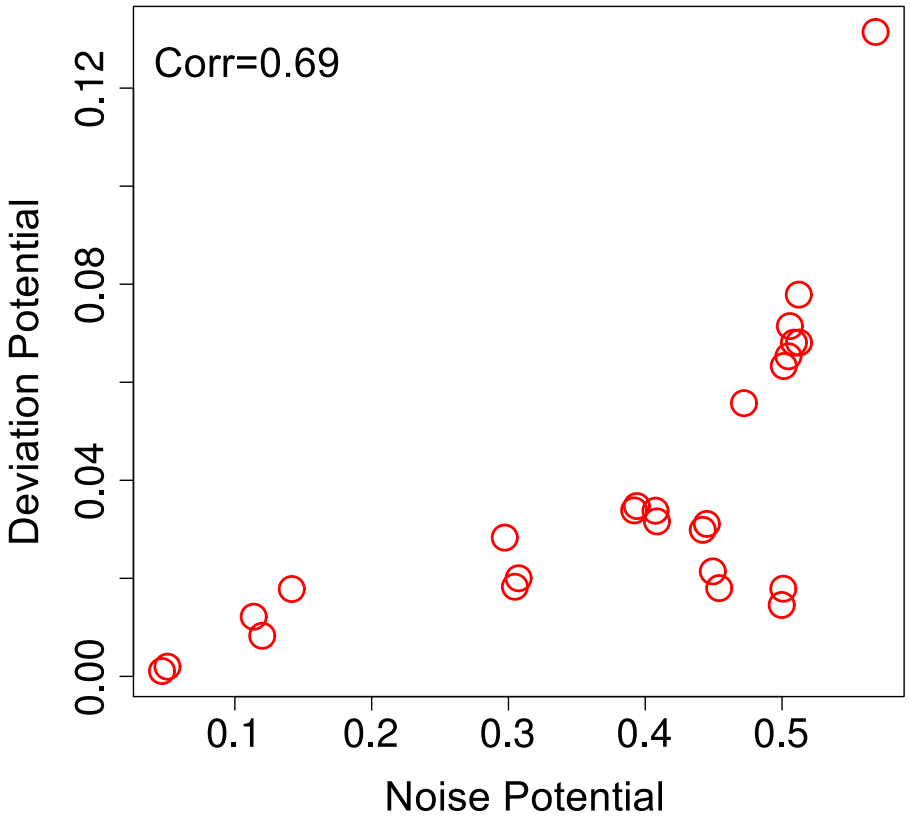

Figure S2. A scatter plot showing the correlation between the potential for stochastic variation (x-axis) and the potential for deviation level (y-axis) of each network in the three-node networks. In each network, the potential for stochastic variation is computed by calculating the average of the five highest noise levels, whereas the potential for deviation level is computed by calculating the average of the five highest deviation levels. The correlation coefficient of the two data is 0.69, suggesting a high positive relation between the potential for deviation level and the potential for variation.

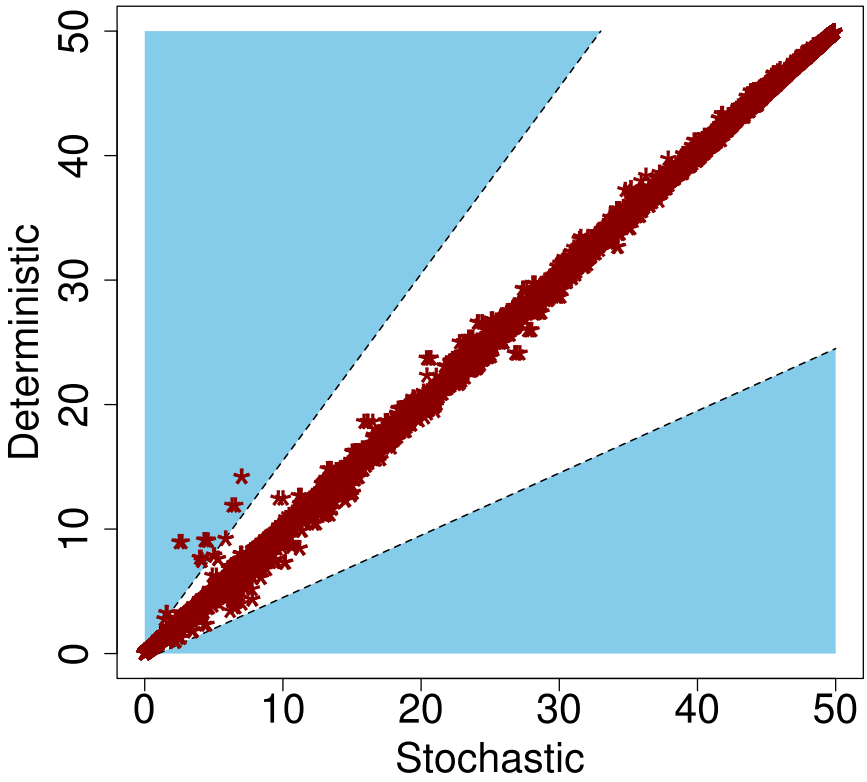

Figure S3. A scatter plot showing the correlation between the stochastic steady state (x-axis) and the deterministic steady state (y-axis) of each network/parameter combinations in the four-node networks. The area colored in sky blue represents the region where the deviation level is greater than or equal to 0.5. A very small fraction of the combinations (17 out of 27,648) have deviation levels higher than 0.5.

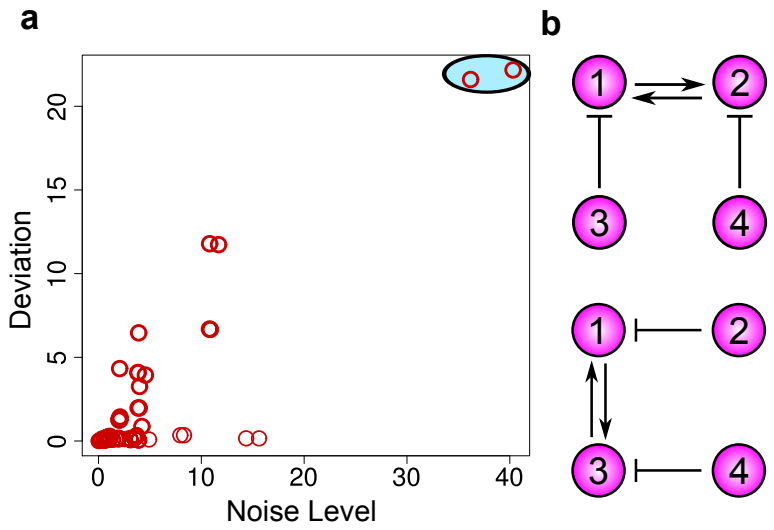

Figure S4. The relation between the potential for variation and the potential for deviation level of four-node networks with Michaelis-Menten kinetics. (a) A scatter plot showing the correlation between the potential for stochastic variation (x-axis) and the potential for deviation level (y-axis) of each network. Each red open circle represents the potential for deviation and variation of a network. In each network, the potential for stochastic variation is computed by calculating the average of the ten highest noise levels, whereas the potential for deviation level is computed by calculating the average of the ten highest deviation levels. The two networks surrounded by the oval in sky blue are the ones with the highest potential for deviation and variation. (b) The structures of the two networks with the highest potential for deviation and variation. These are the same as N4A and N4B in the main text.

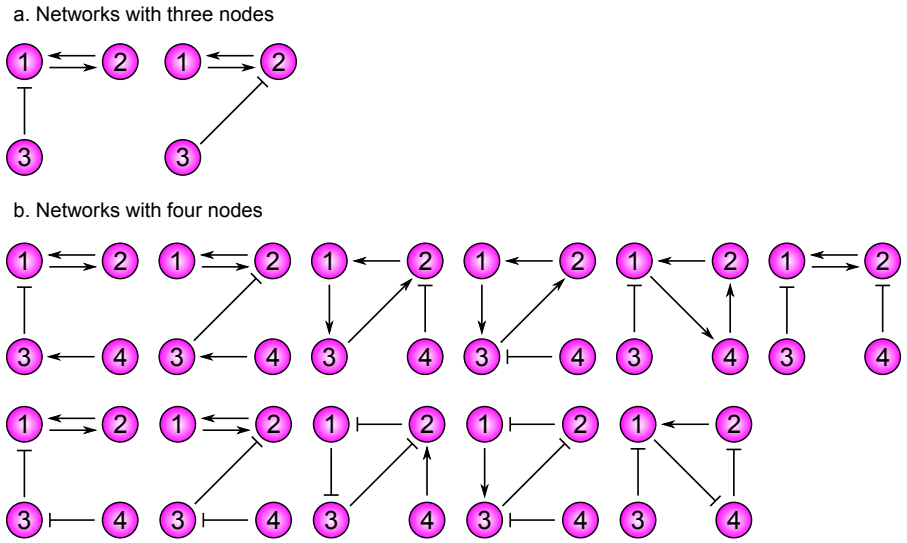

Figure S5. Networks capable of exhibiting bistability assuming enzymatic reaction kinetics. Stability of each network in our three-node and four-node networks is checked using CRNT Toolbox. (a) three-node networks with bistability. (b) four-node networks with bistability.

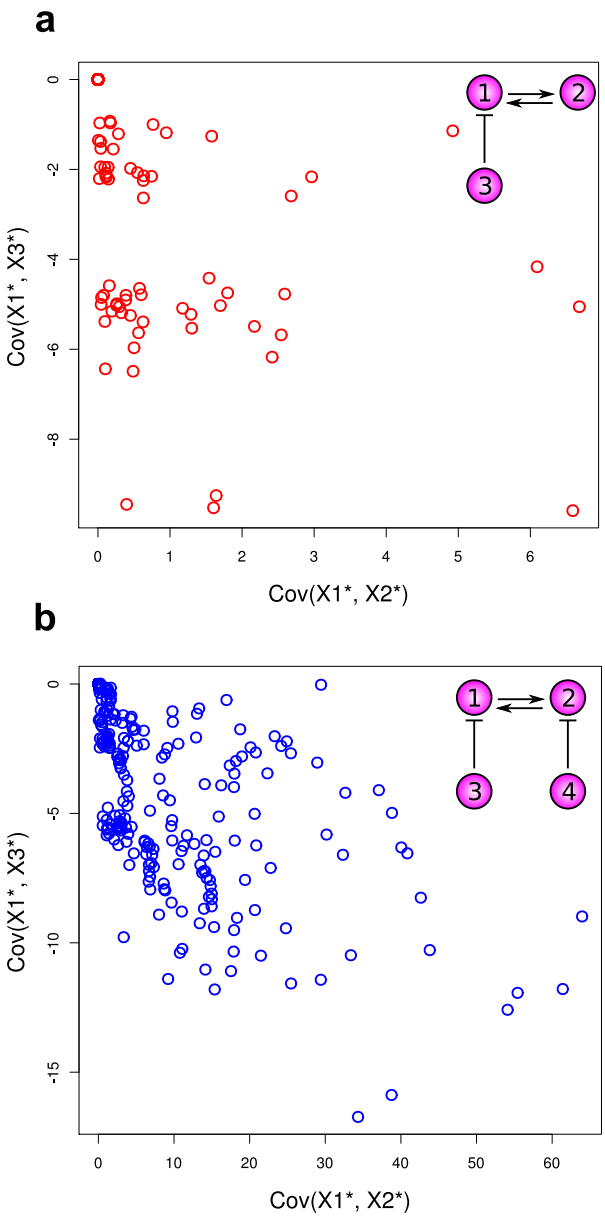

Figure S6. Changes in the distributions of the covariances, $Cov(X_1^{ss*}, X_2^{ss*})$ and $Cov(X_1^{ss*}, X_3^{ss*})$, based on the addition of node 4 in N4A. (a) A scatter plot showing the covariance between X_1^{ss*} and X_2^{ss*} (x-axis) and the covariance between X_1^{ss*} and X_3^{ss*} (y-axis) of each parameter combination for the three-node network showing in the upper right-hand corner of the panel. (b) A scatter plot showing the covariance between X_1^{ss*} and X_2^{ss*} (x-axis) and the covariance between X_1^{ss*} and X_3^{ss*} (y-axis) of each parameter combination for N4A. Each covariance is computed by simulating each stochastic model for 100 time units. Data were uniformly sampled at 100,000 time points. The first 10,000 samples were discarded, and the rest were used to compute the covariances.

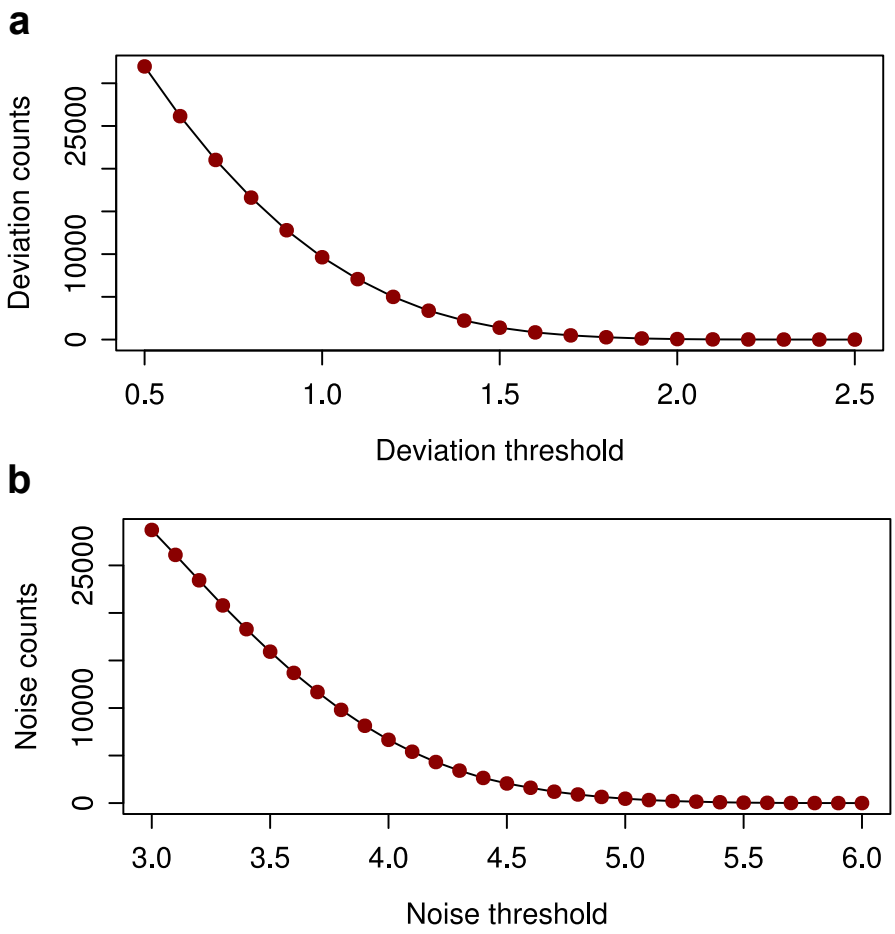

Figure S7. The distributions of the number of parameter combinations in N4A that reach given deviation and noise levels. Here, each node has 25 different parameter combinations, resulting in 390,625 distinct parameter combinations. (a) The distribution based on given deviation levels. (b) The distribution based on given noise levels.

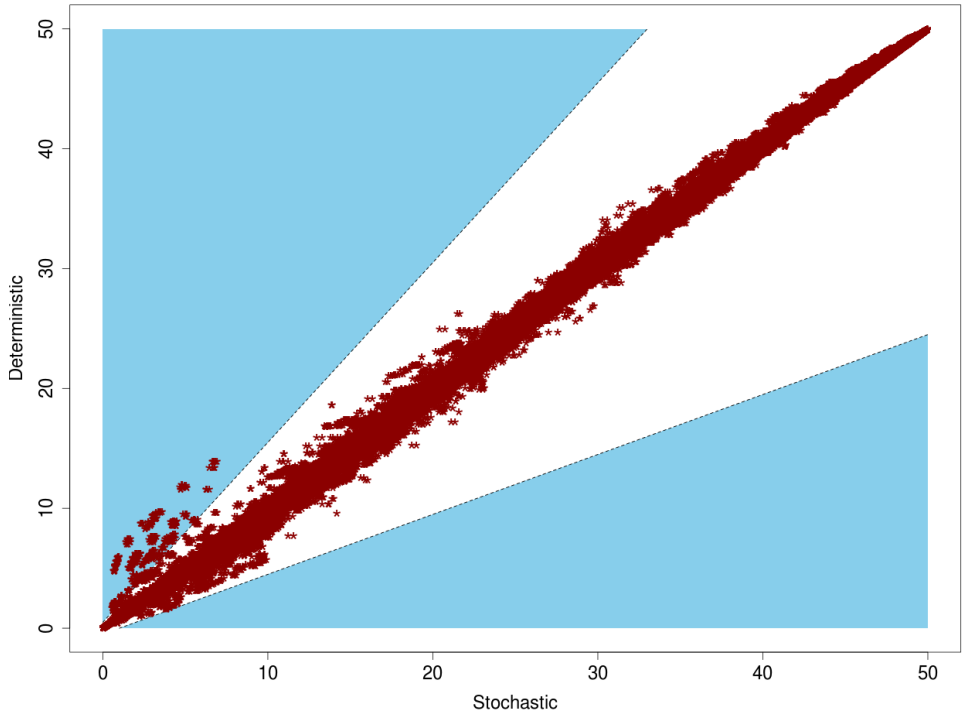

Figure S8. A scatter plot showing the stochastic steady state (x-axis) and the deterministic steady state (y-axis) of each network/parameter combinations in the five-node networks. There are 436 different five-node networks, each of which has 1,024 different parameter combinations, resulting in 446,464 distinct combinations. The area colored in sky blue represents the region in which the deviation level is greater than or equal to 0.5. 474 combinations had deviation levels greater than or equal to 0.5.

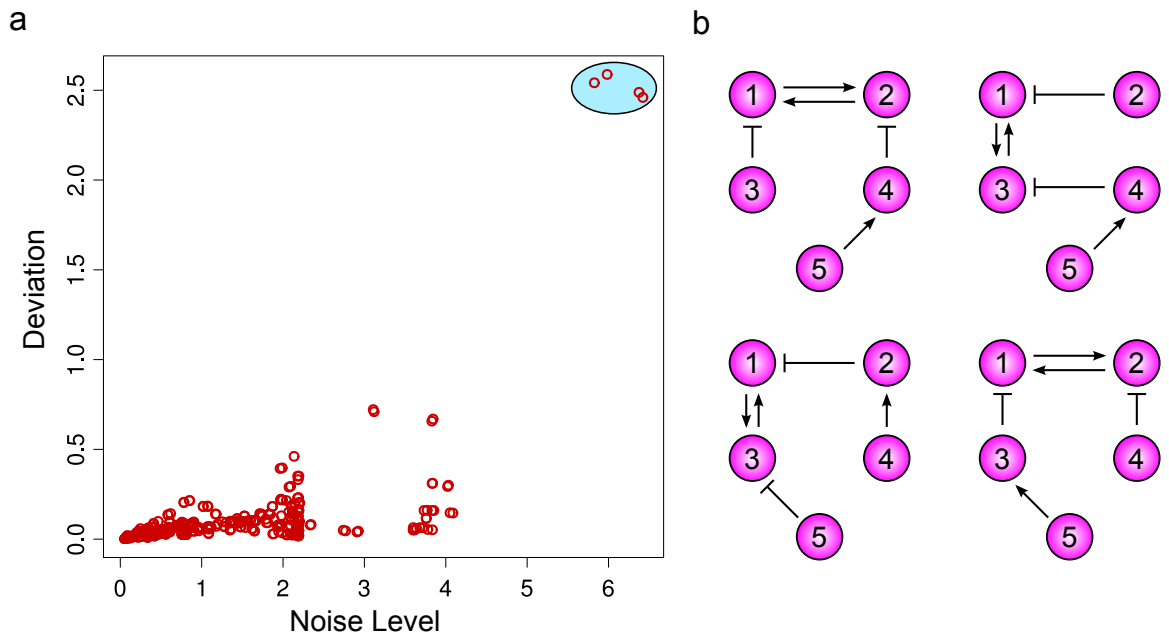

Figure S9. The potnetial for deviation and stochastic variation in five-node networks. (a) A scatter plot showing the correlation between the potential for stochastic variation (x-axis) and the potential for deviation level (y-axis) of each network in the five-node networks. In each network, the potential for stochastic variation is computed by calculating the average of the 10 highest noise levels, whereas the potential for deviation level is computed by calculating the average of the 10 highest deviation levels. The fourred circles surrounded by the light blue oval represent the four networks with the highest deviation and noise potential. (b) The strcutures of the four five-node networks with the highest deviation and noise potential.

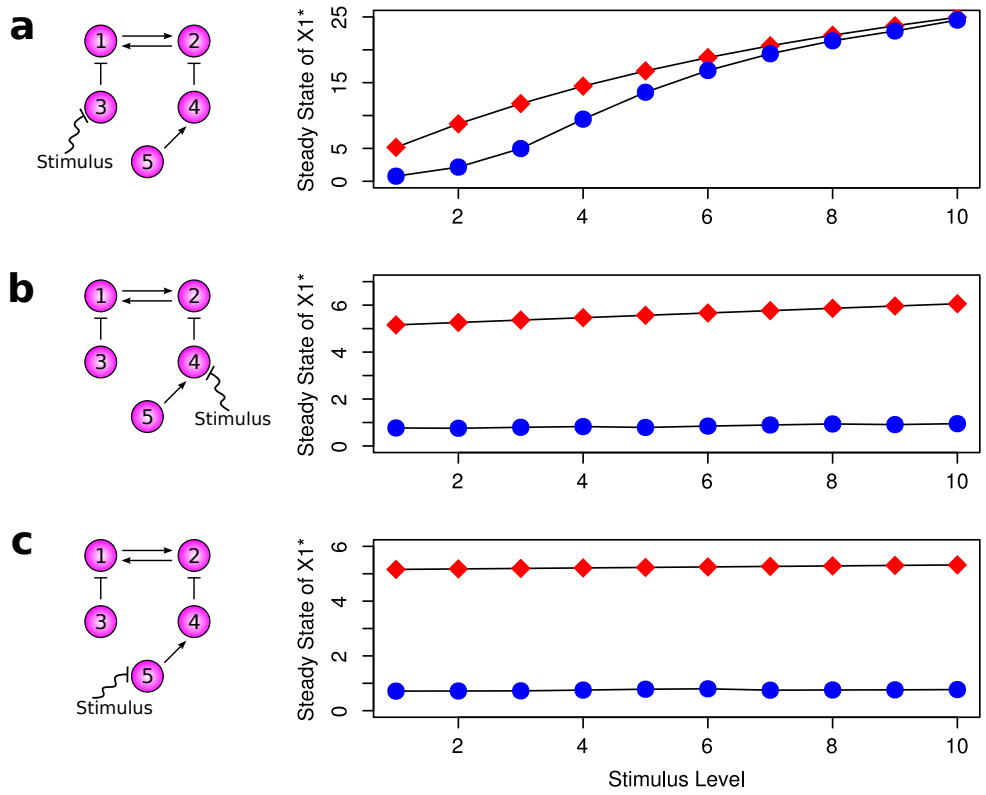

Figure S10. Steady-state response curves based on inhibitory stimuli to various nodes. The steady-state level of X_1^* is measured given various levels of an inhibitory stimulus to (a) node 3; (b) node 4; and (c) node 5. Here, the red diamonds represent data from the deterministic model, while the blue circles represent data from the stochastic model.

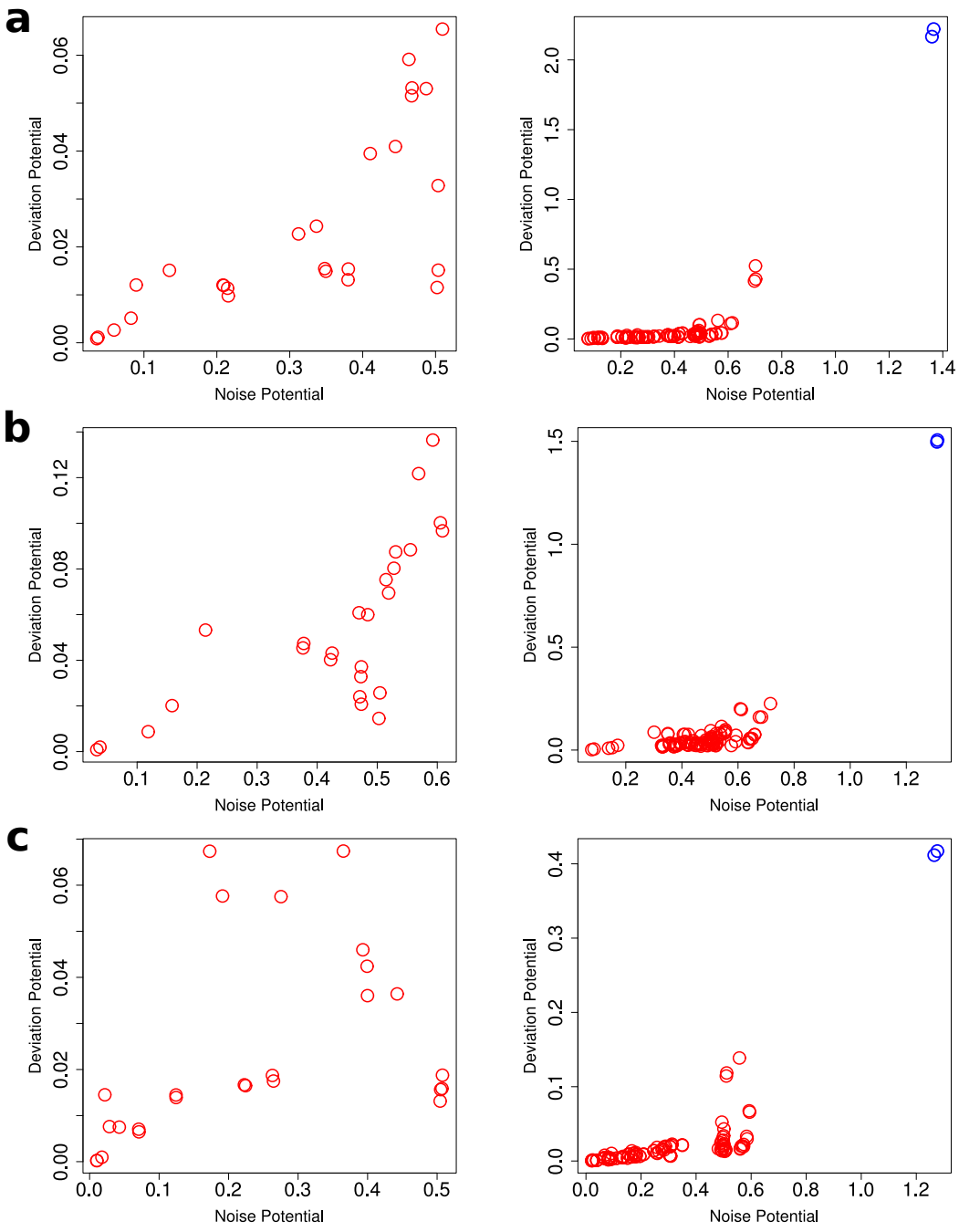

Figure S11. The potential for deviation and noise levels for various model conditions. The left panels show results from three-node networks, whereas the right panels show results from four-node networks that correspond to the three-node network on the same row. On the right panels, the blue circles represent the data from N4A and N4B. In each of the models simulated here, the parameter condition is changed from the original one in the main text as follows: (a) $k_f, k_b \in \{3, 9\}$. (b) $X_{tot} = 100$ and $k_f, k_b \in \{1, 10\}$. (c) $X_{tot} = 200$ and $k_f, k_b \in \{0.5, 2\}$.

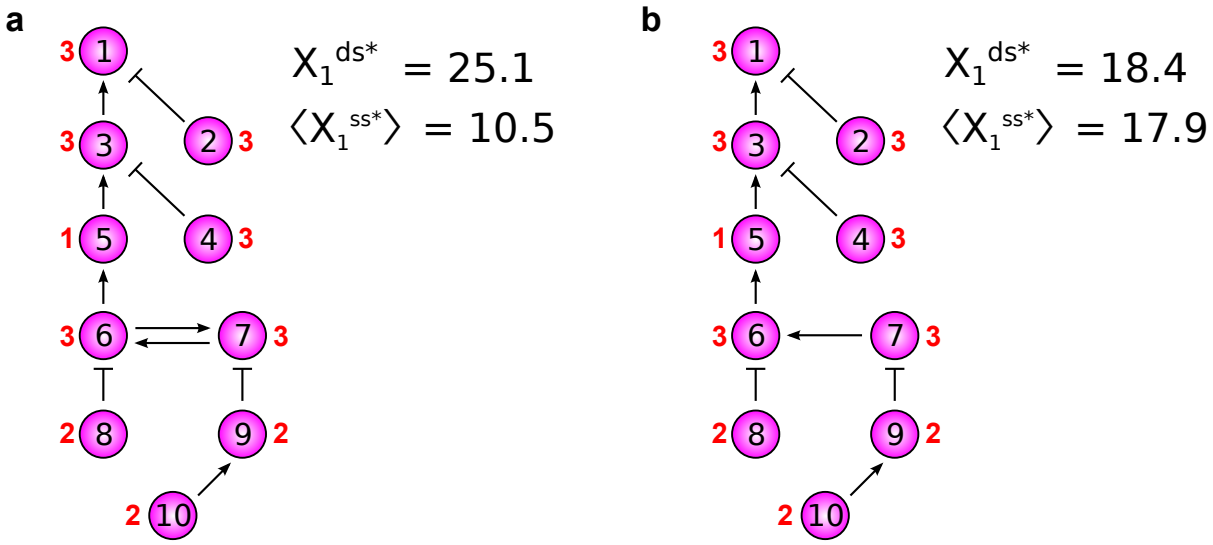

Figure S12. Propagation of deviations in a large-network context. (a) Discrepancy between $\langle X_1^{ss*}\rangle$ and X_1^{ds*} in a 10-node network propagated from a 5-node subnetwork with high deviation level upstream. (b) Insignificant difference between $\langle X_1^{ss*}\rangle$ and X_1^{ds*} in a 10-node network with a 5-node subnetwork with insignificant level of deviation upstream. Here, similar to Table S1, the number in red by each node indicates parameter combinations. 0 means $k_f = 1, k_b = 1$; 1 means $k_f = 1, k_b = 5$; 2 means $k_f = 5, k_b = 1$; and 3 means $k_f = 5, k_b = 5$. .

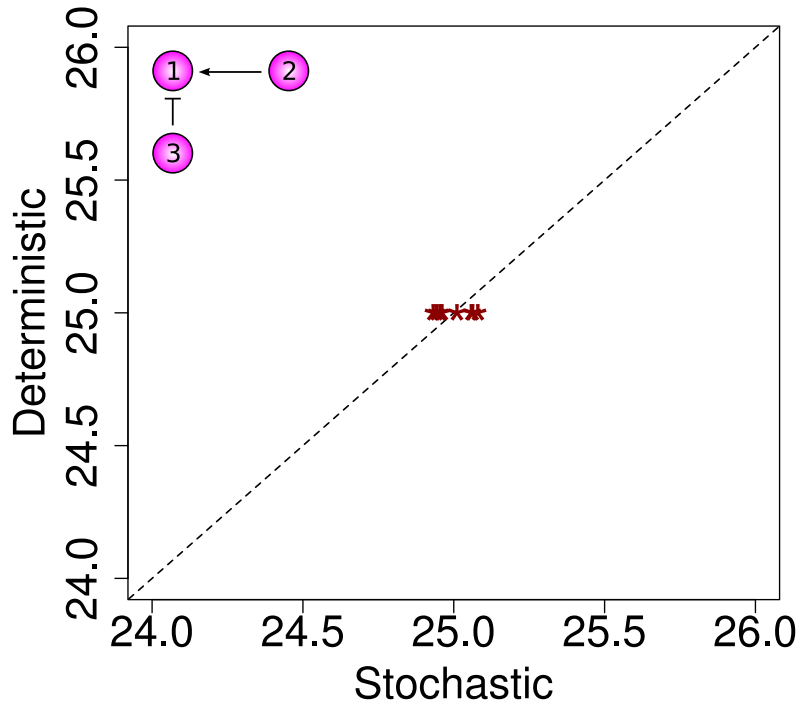

Figure S13. A scatter plot showing the stochastic steady state, $\langle X_1^{ss*} \rangle$, (x-axis) and the deterministic steady state, X_1^{ds*}, (y-axis) of three-node networks with parameter combinations that are known to be deviation free. Since the enzymes for nodes 2 and 3 are constant, we have $\langle X_2^{ss*} \rangle = X_2^{ds*}$ and $\langle X_3^{ss*} \rangle = X_3^{ds*}$. Here, we have eight combinations, and for each the parameters for nodes 2 and 3 are set to be the same so that we have $X_2^{ss*} = X_3^{ss*}$. Furthermore, we set $k_{f1} = k_{b1}$ so that $k_{f1} Cov(X_1^{ss*}, X_2^{ss*}) = -k_{b1} Cov(X_1^{ss*}, X_3^{ss*})$. Thus, from Eq. S7, this network is deviation-free in these parameter conditions, and this control experiment shows the correctness of our simulations.