Robust information transmission in noisy biochemical reactions

Takehiro Tottori
Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo, Japan

Masashi Fujii
Department of Integrated Sciences for Life, Graduate School of Integrated Sciences for Life, Hiroshima University, Hiroshima, Japan
Department of Bioinformatics and Systems Biology, Faculty of Science, University of Tokyo, Tokyo, Japan
Molecular Genetics Research Laboratory, Graduate School of Science, University of Tokyo, Tokyo, Japan and
Department of Biological Sciences, Graduate School of Science, University of Tokyo, Tokyo, Japan

Shinya Kuroda
Department of Bioinformatics and Systems Biology, Faculty of Science, University of Tokyo, Tokyo, Japan
Molecular Genetics Research Laboratory, Graduate School of Science, University of Tokyo, Tokyo, Japan and
Department of Biological Sciences, Graduate School of Science, University of Tokyo, Tokyo, Japan

Fluctuations of intracellular biochemical reactions (intrinsic noise) reduce the information transmitted from an extracellular input to a cellular response. However, recent studies have demonstrated that the transmitted information barely decreases with respect to extracellular input fluctuations (extrinsic noise) when the intrinsic noise is large. Therefore, it has been suggested that larger intrinsic noise enables more robust information transmission in the presence of extrinsic noise, which is called intrinsic noise-induced robustness (INIR). However, because previous studies have focused on specific scenarios, it is unclear whether INIR appears universally in biochemical reactions. Moreover, the mechanism of INIR remains elusive. In this paper, we address these questions by analyzing simple models. We first analyze a model in which the input–output relation is linear. We show that the robustness against extrinsic noise increases with the intrinsic noise, therefore realizing the INIR phenomenon. Moreover, INIR is particularly strong when the variance of the intrinsic noise is larger than that of the input distribution. Next, we analyze a threshold model in which the output depends on whether the input exceeds the threshold. When the threshold is equal to the mean of the input, INIR is realized, but when the threshold is much larger than the mean, the threshold model exhibits stochastic resonance, and INIR is not always apparent. However, the robustness against extrinsic noise and the transmitted information can be traded off against one another in the linear model and the threshold model without stochastic resonance, whereas they can be simultaneously increased in the threshold model with stochastic resonance.

I. INTRODUCTION

The transmission of cellular information is an important function in various biological processes. Environmental information is transmitted to generate a cellular response by intracellular biochemical reactions. However, these reactions exhibit significant fluctuations [1–4], which can reduce the information transmitted from the environment to the cellular response [5–8]. Therefore, fluctuations in intracellular reactions (intrinsic noise) are thought to be disadvantageous for cellular information transmission. This raises the question of whether there are any advantageous effects of intrinsic noise.

Recent studies simulating biochemical reactions have demonstrated that the transmitted information barely decreases with respect to fluctuations in extracellular molecules (extrinsic noise) when the intrinsic noise is large [9–11]. In other words, larger intrinsic noise enables more robust information transmission in the presence of extrinsic noise. Therefore, it has been suggested that robustness against extrinsic noise is one of the advantageous characteristics of intrinsic noise. In this paper, we call this characteristic intrinsic noise-induced robustness (INIR).

However, previous studies focused on specific biochemical reactions [9–11]; therefore, their models were too complicated to clarify the essence of INIR. Hence, it is still unclear whether INIR appears universally in various biochemical reactions.

In this paper, we reveal the universality and mechanism of INIR by analyzing simple models. This paper is organized as follows. In Sec. [11], we analyze a model in which the input–output relation is linear. This model is an extension of the Gaussian channel, which is a standard model in information theory [12]. We show that the decrease in transmitted information with respect to extrinsic noise is smaller when the intrinsic noise is larger. Therefore, even this simple model reproduces the INIR
FIG. 1: Schematic diagram of cellular information transmission. (a) Environmental information is encoded to extracellular molecules. Cells detect extracellular molecules and determine proper responses via intracellular reactions. (b) Environments and cellular responses correspond to input $X$ and output $Y$, respectively. Furthermore, fluctuations of extracellular molecules and intracellular reactions correspond to extrinsic noise $Z_{ex}$ and intrinsic noise $Z_{in}$, respectively.

phenomenon. Moreover, we demonstrate that INIR becomes stronger when the variance of the intrinsic noise is larger than that of the input distribution.

In Sec. III, we analyze a threshold model in which the output depends on whether the input exceeds the threshold. When the threshold is equal to the mean of the input, INIR is realized in the threshold model (Sec. IIIA). However, when the threshold is much larger than the mean of the input, the threshold model exhibits stochastic resonance (SR), and INIR is not always realized (Sec. IIIB). However, robustness against extrinsic noise and the transmitted information can be traded off against one another in the linear model and the threshold model without SR, whereas they can be simultaneously increased in the threshold model with SR.

In Sec. IV, we explain why INIR is realized in the linear model and in the threshold model without SR, but is not always realized in the threshold model with SR. In Sec. V, we discuss the biological significance of our results and the relevance of other studies in this field.

II. LINEAR MODEL

We consider a channel composed of the input $X$, extrinsic noise $Z_{ex}$, intrinsic noise $Z_{in}$, and output $Y$ (Fig. 1). The input $X$ corresponds to environmental information detected by the cells, and the output $Y$ corresponds to the cellular response against the environment. Furthermore, the extrinsic noise $Z_{ex}$ corresponds to the fluctuation of extracellular molecules, and the intrinsic noise $Z_{in}$ corresponds to the fluctuation of intracellular reactions.

Here, we consider the input $X$, extrinsic noise $Z_{ex}$, and intrinsic noise $Z_{in}$ to obey Gaussian distributions $N(0, 1)$, $N(0, \sigma_{ex}^2)$, and $N(0, \sigma_{in}^2)$, respectively. Furthermore, the output $Y$ is given by the following:

$$Y = X + Z_{ex} + Z_{in}$$

This model is an extension of the Gaussian channel, which is a standard model in information theory [12]. In this paper, we call this the “linear model.”

To quantify the information transmitted from the input $X$ to the output $Y$, we use the mutual information between $X$ and $Y$, $I(X;Y)$ [12]. When the mutual information $I(X;Y)$ is large, the information transmitted from $X$ to $Y$ is large. When $X$ and $Y$ are continuous random variables, $I(X;Y)$ is given by [12]:

$$I(X; Y) = \int_X \int_Y p(x, y) \log_2 \left( \frac{p(x, y)}{p(x)p(y)} \right) dxdy$$

where $p(x, y)$ is the joint probability density function of $X$ and $Y$, and $p(x)$ and $p(y)$ are the marginal probability density functions of $X$ and $Y$, respectively. In the linear model, $I(X;Y)$ can be derived analytically as [12]:

$$I(X; Y) = \frac{1}{2} \log_2 \left( 1 + \frac{1}{\sigma_{ex}^2 + \sigma_{in}^2} \right)$$

The mutual information $I(X;Y)$ decreases with the intrinsic noise $\sigma_{in}^2$ when $\sigma_{ex}^2 = 0$ (Fig. 2, a)). However, the decrease in $I(X;Y)$ with respect to extrinsic noise $\sigma_{ex}^2$ is reduced at higher levels of intrinsic noise $\sigma_{in}^2$ (Fig. 2, b)). In other words, larger intrinsic noise enables more robust information transmission in the presence of extrinsic noise. Therefore, INIR is realized in the linear model.

To investigate the robustness against extrinsic noise in more detail, we define the robustness function $\delta_q(\sigma_{in}^2)$ as the extrinsic noise $\sigma_{ex}^2$ when the normalized mutual information $I(X;Y)/I(X;Y)|_{\sigma_{ex}^2=0}$ decreases to $q$ for $q \in (0, 1)$ (Fig. 2, c)). Therefore, the robustness function $\delta_q(\sigma_{in}^2)$ satisfies the following:

$$\frac{I(X; Y)|_{\sigma_{ex}^2 = \delta_q(\sigma_{in}^2)}}{I(X; Y)|_{\sigma_{ex}^2 = 0}} = q, \quad q \in (0, 1)$$

When $\delta_q(\sigma_{in}^2)$ is large, the robustness against extrinsic noise $\sigma_{ex}^2$ is high.

In terms of $q$, the robustness function $\delta_q(\sigma_{in}^2)$ can be interpreted as follows. When $q$ is close to $1$, the robustness function $\delta_q(\sigma_{in}^2)$ represents the extrinsic noise $\sigma_{ex}^2$ when the mutual information $I(X;Y)$ decreases slightly. In other words, when $q$ is close to $1$, the robustness function $\delta_q(\sigma_{in}^2)$ represents the “local” robustness against extrinsic noise $\sigma_{ex}^2$. When $q$ is much smaller than $1$, the robustness function $\delta_q(\sigma_{in}^2)$ represents the extrinsic noise $\sigma_{ex}^2$ when the mutual information $I(X;Y)$ has decreased significantly. In other words, when $q$ is much smaller than
In the linear model, the robustness function $\delta_q(\sigma^2_{in})$ can be derived analytically as:

$$\delta_q(\sigma^2_{in}) = \frac{1}{(1 + \sigma^2_{in})^q} - 1$$  \hspace{1cm} (5)

(see Appendix [A1]). The robustness function $\delta_q(\sigma^2_{in})$ increases with intrinsic noise $\sigma^2_{in}$ for any $q$ (Fig. 2(d)). Therefore, INIR is clearly demonstrated.

Moreover, the intrinsic noise dependency of the robustness function $\delta_q(\sigma^2_{in})$ changes at $\sigma^2_{in} = 1$ (Fig. 2(d)). The intrinsic noise dependency of the robustness function $\delta_q(\sigma^2_{in})$ in $\sigma^2_{in} \gg 1$ is stronger than that in $\sigma^2_{in} \ll 1$. Indeed, Eq. 5 can be approximated as follows:

$$\delta_q(\sigma^2_{in}) \propto \begin{cases} \sigma^2_{in}^q & (\sigma^2_{in} \ll 1) \\ \sigma^2_{in}^q & (\sigma^2_{in} \gg 1) \end{cases}$$  \hspace{1cm} (6)

(see Appendix [A2]). Considering $q \in (0, 1)$, Eq. 6 indicates that the intrinsic noise dependency of the robustness function $\delta_q(\sigma^2_{in})$ in $\sigma^2_{in} \gg 1$ is stronger than that in $\sigma^2_{in} \ll 1$.

As the variance of the input $X$ is 1, the intrinsic noise dependency of the robustness function $\delta_q(\sigma^2_{in})$ becomes stronger when the intrinsic noise $\sigma^2_{in}$ is much larger than the variance of the input $X$. Therefore, the input distribution plays an important role in INIR. When the intrinsic noise is much larger than the variance of the input, INIR is strongly apparent.

III. THRESHOLD MODEL

A. The case $\theta = 0$

In Sec. II we showed that the linear model realizes INIR. Moreover, we revealed that INIR is stronger when the intrinsic noise is much larger than the variance of the input.

However, the input–output relation in cellular information transmission is often nonlinear. In particular, some form of threshold response is often observed [13–15]. Therefore, in this subsection, we introduce a threshold model, and examine whether this reproduces the same results as the linear model.

We consider the input $X$, extrinsic noise $Z_{ex}$, and intrinsic noise $Z_{in}$ to obey Gaussian distributions $N(0,1)$, $N(0,\sigma_{ex}^2)$, and $N(0,\sigma_{in}^2)$, respectively, as for the linear model. However, the output $Y$ is given by:

$$Y = \begin{cases} 1 & (X + Z_{ex} + Z_{in} \geq \theta) \\ 0 & (X + Z_{ex} + Z_{in} < \theta) \end{cases}$$  \hspace{1cm} (7)

Therefore, the output $Y$ is 1 when $X + Z_{ex} + Z_{in}$ is greater than the threshold $\theta$, whereas the output $Y$ is 0 when $X + Z_{ex} + Z_{in}$ is less than $\theta$. In this paper, we call this the “threshold model.”

In the threshold model, the input $X$ is a continuous random variable, but the output $Y$ is a discrete random variable. Therefore, the mutual information $I(X;Y)$ in the threshold model is given by [12]:

$$I(X;Y) = \sum_Y \int_X p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)} \, dx$$  \hspace{1cm} (8)

For simplicity, in this subsection, we discuss the case of $\theta = 0$ (Fig. 3(a)). As the mean of the input $X$ is 0, $\theta = 0$ corresponds to the case where the mean of the input $X$ is equal to the threshold $\theta$. In Sec. III(B) we will discuss the case of $\theta \neq 0$.

Unlike the linear model, the mutual information $I(X;Y)$ in the threshold model $(\theta = 0)$ cannot be derived analytically. However, it can be approximated as:

$$I(X;Y) \approx 1 - \frac{1}{\sqrt{\frac{2}{\pi \sigma_{in}^2} \left(\sigma_{ex}^2 + \sigma_{in}^2 \right)^{-1} + 1}}$$  \hspace{1cm} (9)

(see Appendix [B3]). This approximate solution matches the numerical solution closely (Fig. 3(b), (c)).

The mutual information $I(X;Y)$ in the threshold model $(\theta = 0)$ decreases with the intrinsic noise $\sigma_{in}^2$, at $\sigma_{ex}^2 = 0$ (Fig. 3(b)). However, the decrease in mutual information $I(X;Y)$ with respect to extrinsic noise $\sigma_{ex}^2$ is reduced as the intrinsic noise $\sigma_{in}^2$ increases (Fig. 3(c)). In other words, larger values of intrinsic noise enable robust information transmission against the extrinsic noise. Therefore, INIR is also realized in the threshold model $(\theta = 0)$. 

![Figure 2: Linear model. (a) Mutual information $I(X;Y)$ against intrinsic noise $\sigma_{in}^2$ at $\sigma_{ex}^2 = 0$. (b) Mutual information $I(X;Y)$ against extrinsic noise $\sigma_{ex}^2$.)](image-url)
As explained in Sec. II, the robustness function \( \delta_q(\sigma^2_{in}) \) represents the extrinsic noise \( \sigma^2_{ex} \) when the normalized mutual information \( I(X;Y)/I(X;Y)_{\sigma^2_{ex}=0} \) decreases to \( q \) for \( q \in (0,1) \). When \( \delta_q(\sigma^2_{in}) \) is large, the robustness against extrinsic noise is high.

The robustness function \( \delta_q(\sigma^2_{in}) \) in the threshold model \( (\theta = 0) \) cannot be derived analytically because we do not have an exact expression for \( I(X;Y) \). However, \( \delta_q(\sigma^2_{in}) \) can be derived approximately using the approximate solution of \( I(X;Y) \) (Eq. (9)). Thus:

\[
\delta_q(\sigma^2_{in}) \approx \frac{(q - I_0)^2}{2q - I_0} \frac{2}{I_0(\pi \ln 2)} - \sigma^2_{in}
\]

where

\[
I_0 = 1 - \frac{1}{\sqrt{\frac{2}{(\pi \ln 2)^2} + 1}}
\]

(see Appendix C1). The robustness function \( \delta_q(\sigma^2_{in}) \) increases with the intrinsic noise \( \sigma^2_{in} \) for any \( q \) (Fig. 3(d)). Therefore, INIR is clearly demonstrated.

B. The case \( \theta \neq 0 \)

In Sec. IIIA, we discussed the threshold model in the case \( \theta = 0 \). In this subsection, we consider the case \( \theta \neq 0 \) (Fig. 3(a)). The mutual information \( I(X;Y) \) is an even function of \( \theta \). Therefore, we can consider the case \( \theta \geq 0 \) without loss of generality.

The mutual information \( I(X;Y) \) in the threshold model cannot be derived analytically, but the following approximate expression can be obtained:

\[
I(X;Y) \approx \left( 1 - \frac{1}{\frac{2}{(\pi \ln 2)^2} + 1} \right)
\times \exp \left( -\frac{\theta^2}{(\pi \ln 2) \left( \frac{2}{(\pi \ln 2)^2} + \sigma^2_{in} \right)} \right)
\]

(see Appendix D1). This approximate solution matches the numerical solution well (Fig. 4(b)).

When \( \theta = 0 \), that is, the threshold \( \theta \) is equal to the mean of the input \( X \), the mutual information \( I(X;Y) \) decreases monotonically with respect to the intrinsic noise \( \sigma^2_{in} \) (Fig. 4(b)(red)). When \( \theta \) is much larger than 0, that is, the threshold \( \theta \) is much larger than the mean of the input \( X \), the mutual information \( I(X;Y) \) is maximized with a moderate value of the intrinsic noise \( \sigma^2_{in} \) (Fig. 4(b)). This phenomenon is called “stochastic resonance” (SR) [16, 17].

Why does SR appear in the threshold model for large \( \theta \)? This is the same as asking why the mutual information
When SR does not appear ($\theta = 0$), the robustness function $\delta_q(\sigma_{\text{in}}^2)$ decreases monotonically with the intrinsic noise $\sigma_{\text{in}}^2$, for any $q$ (Fig. 3(d)). However, when SR appears ($\theta = 3$), the robustness function $\delta_q(\sigma_{\text{in}}^2)$ does not always increase with the intrinsic noise $\sigma_{\text{in}}^2$ (Fig. 4(d)). Therefore, INIR is not always realized in the presence of SR.

We further examine the region where INIR is not realized. Defining $\sigma_{\text{in, max}}^2$ as the intrinsic noise $\sigma_{\text{in}}^2$ that maximize the mutual information $I(X; Y)$, the robustness function $\delta_q(\sigma_{\text{in}}^2)$ decreases with the intrinsic noise $\sigma_{\text{in}}^2$ when $\sigma_{\text{in}}^2 < \sigma_{\text{in, max}}^2$ (Fig. 4(d)(gray)). In other words, the robustness against extrinsic noise $\sigma_{\text{ex}}^2$ decreases with respect to intrinsic noise $\sigma_{\text{in}}^2$ when the mutual information $I(X; Y)$ increases with respect to intrinsic noise $\sigma_{\text{in}}^2$. This is trivial, because the mutual information $I(X; Y)$ increases with the range of the extrinsic noise $\sigma_{\text{ex}}^2$ for which the mutual information $I(X; Y)$ increases will expand. Thus, the robustness against extrinsic noise $\sigma_{\text{ex}}^2$ increases as the intrinsic noise $\sigma_{\text{in}}^2$ decreases when $\sigma_{\text{in}}^2 < \sigma_{\text{in, max}}^2$.

The robustness function $\delta_q(\sigma_{\text{in}}^2)$ decreases with the intrinsic noise $\sigma_{\text{in}}^2$ when $\sigma_{\text{in}}^2 > \sigma_{\text{in, max}}^2$ (Fig. 4(d)). In other words, the robustness against the extrinsic noise $\sigma_{\text{ex}}^2$ decreases with respect to the intrinsic noise $\sigma_{\text{in}}^2$ when the mutual information $I(X; Y)$ decreases with respect to the intrinsic noise $\sigma_{\text{in}}^2$. Defining $\sigma_{\text{in, c}}^2$ as the intrinsic noise $\sigma_{\text{in}}^2$ at which the mutual information $I(X; Y)$ changes from convex upward to convex downward, the robustness function $\delta_q(\sigma_{\text{in}}^2)$ decreases with respect to the intrinsic noise $\sigma_{\text{in}}^2$ when $\sigma_{\text{in, c}}^2 < \sigma_{\text{in}}^2 < \sigma_{\text{in, max}}^2$ and $q$ is close to 1 (Fig. 4(d)(green)). In other words, the (local) robustness against the extrinsic noise $\sigma_{\text{ex}}^2$ decreases with respect to the intrinsic noise $\sigma_{\text{in}}^2$ when the mutual information $I(X; Y)$ is decreasing and convex upward with respect to the intrinsic noise $\sigma_{\text{in}}^2$.

We now prove that the robustness function $\delta_q(\sigma_{\text{in}}^2)$ decreases with respect to the intrinsic noise $\sigma_{\text{in}}^2$ when the mutual information $I(X; Y)$ is decreasing and convex upward against the intrinsic noise $\sigma_{\text{in}}^2$, and $q$ is close to 1.

In the threshold model, the $\sigma_{\text{in}}^2$-dependency of $I(X; Y)$ is the same as the $\sigma_{\text{ex}}^2$-dependency of $I(X; Y)$, and $I(X; Y)$ is decreasing against $\sigma_{\text{in}}^2$. Therefore,

$$\frac{\partial I(X; Y)}{\partial \sigma_{\text{ex}}^2} = \frac{\partial I(X; Y)}{\partial \sigma_{\text{in}}^2} < 0$$

is satisfied. When $\partial I(X; Y)/\partial \sigma_{\text{ex}}^2 < 0$, the robustness function $\delta_q(\sigma_{\text{in}}^2)$ becomes close to 0 as $q \to 1$. Therefore,
when \( \partial I(X;Y)/\partial \sigma_{ex}^2 < 0 \) and \( q \sim 1 \),
\[
I(X;Y)|_{\sigma_{ex}^2=\delta_q} = I(X;Y)|_{\sigma_{ex}^2=0} + \frac{\partial I(X;Y)}{\partial \sigma_{ex}^2} \bigg|_{\sigma_{ex}^2=0} \delta_q + O(\delta_q^2) \tag{15}
\]
is satisfied. Substituting Eq. (15) into Eq. (4), we obtain
\[
\delta_q(\sigma_{in}^2) = -\frac{(1-q)}{\frac{\partial \ln I(X;Y)}{\partial \sigma_{ex}^2} \bigg|_{\sigma_{ex}^2=0}} \tag{16}
\]
\[
\partial \ln I(X;Y)/\partial \sigma_{ex}^2 |_{\sigma_{ex}^2=0}
\]
corresponds to the slope of the normalized mutual information \( I(X;Y)/I(X;Y)|_{\sigma_{ex}^2=0} \)
against the extrinsic noise \( \sigma_{ex}^2 \) at \( \sigma_{ex}^2 = 0 \). When \( \partial \ln I(X;Y)/\partial \sigma_{ex}^2 |_{\sigma_{ex}^2=0} \) is large, the robustness function
\[
\delta_q(\sigma_{in}^2)
\]
at \( q \sim 1 \) is also large.

Differentiating Eq. (16) with respect to the intrinsic noise \( \sigma_{in}^2 \), we have
\[
\frac{\partial \delta_q(\sigma_{in}^2)}{\partial \sigma_{in}^2} = \frac{(1-q)}{\left(\frac{\partial I(X;Y)}{\partial \sigma_{ex}^2} \bigg|_{\sigma_{ex}^2=0}\right)^2} \tag{17}
\]
In the threshold model, \( I(X;Y) \) depends on the summation of \( \sigma_{in}^2 \) and \( \sigma_{ex}^2 \). Therefore,
\[
\frac{\partial}{\partial \sigma_{in}^2} \left( \frac{\partial \ln I(X;Y)}{\partial \sigma_{ex}^2} \bigg|_{\sigma_{ex}^2=0} \right) = \frac{\partial^2 \ln I(X;Y)}{(\partial \sigma_{in}^2)^2} \bigg|_{\sigma_{ex}^2=0} \tag{18}
\]
is satisfied. From Eqs. (17) and (18), and for \( q < 1 \),
\[
\text{sgn} \left( \frac{\partial \delta_q(\sigma_{in}^2)}{\partial \sigma_{in}^2} \right) = \text{sgn} \left( \frac{\partial^2 \ln I(X;Y)}{(\partial \sigma_{in}^2)^2} \bigg|_{\sigma_{ex}^2=0} \right) \tag{19}
\]
is satisfied, where \( \text{sgn}(\cdot) \) is the sign function.
Here, from simple calculations,
\[
\frac{\partial^2 I(X;Y)}{(\partial \sigma_{in}^2)^2} < 0 \Rightarrow \frac{\partial^2 \ln I(X;Y)}{(\partial \sigma_{in}^2)^2} < 0 \tag{20}
\]
is satisfied. Therefore, from Eq. (19),
\[
\frac{\partial^2 I(X;Y)}{(\partial \sigma_{in}^2)^2} < 0 \Rightarrow \frac{\partial \delta_q(\sigma_{in}^2)}{\partial \sigma_{in}^2} < 0 \tag{21}
\]
is satisfied. Therefore, when \( I(X;Y) \) is decreasing and convex upward with respect to \( \sigma_{in}^2 \), the robustness function \( \delta_q(\sigma_{in}^2) \) at \( q \sim 1 \) decreases with \( \sigma_{in}^2 \).

Moreover,
\[
\frac{\partial^2 \ln I(X;Y)}{(\partial \sigma_{in}^2)^2} > 0 \Rightarrow \frac{\partial^2 I(X;Y)}{(\partial \sigma_{in}^2)^2} > 0 \tag{22}
\]
is satisfied. Therefore, from Eq. (19),
\[
\frac{\partial \delta_q(\sigma_{in}^2)}{\partial \sigma_{in}^2} > 0 \Rightarrow \frac{\partial^2 I(X;Y)}{(\partial \sigma_{in}^2)^2} > 0 \tag{23}
\]
is also satisfied. Hence, when \( I(X;Y) \) is decreasing against \( \sigma_{in}^2 \), and \( \delta_q(\sigma_{in}^2) \) at \( q \sim 1 \) is increasing against \( \sigma_{in}^2 \), \( I(X;Y) \) is always convex downward with respect to \( \sigma_{in}^2 \). Indeed, \( I(X;Y) \) is convex downward against \( \sigma_{in}^2 \) in the linear model and in the threshold model (\( \theta = 0 \)).

Therefore, it has been proved that the robustness function \( \delta_q(\sigma_{in}^2) \) at \( q \sim 1 \) decreases with the intrinsic noise \( \sigma_{in}^2 \) when \( I(X;Y) \) is decreasing and convex upward against \( \sigma_{in}^2 \) (Fig. 3(b) (green)). Thus, INIR is not always realized in the threshold model with SR, even when \( I(X;Y) \) is decreasing against \( \sigma_{in}^2 \). However, this is not necessarily bad, and may rather be an advantageous property. In the linear model and the threshold model without SR, reducing the intrinsic noise \( \sigma_{in}^2 \) increases the mutual information \( I(X;Y) \), but decreases the robustness function \( \delta_q(\sigma_{in}^2) \). Hence, there is a tradeoff between the transmitted information and the robustness against extrinsic noise in the linear model and the threshold model without SR. In contrast, in the threshold model with SR, the mutual information \( I(X;Y) \) and the robustness function \( \delta_q(\sigma_{in}^2) \) at \( q \sim 1 \) can be simultaneously increased by reducing the intrinsic noise \( \sigma_{in}^2 \) when \( I(X;Y) \) is decreasing and convex upward with respect to \( \sigma_{in}^2 \) (Fig. 3(d) (green)). Therefore, SR solves the tradeoff problem between the transmitted information and the robustness against extrinsic noise.

IV. MECHANISM

In Sec. III.B, we showed that when \( \partial I(X;Y)/\partial \sigma_{in}^2 < 0 \) and \( q \sim 1 \), the sign of \( \partial \delta_q(\sigma_{in}^2)/\partial \sigma_{in}^2 \) strongly depends on the sign of \( \partial^2 I(X;Y)/(\partial \sigma_{in}^2)^2 \). The robustness against extrinsic noise increases with respect to the intrinsic noise when \( \partial^2 I(X;Y)/(\partial \sigma_{in}^2)^2 > 0 \) (Fig. 5(c) (blue)), whereas it decreases when \( \partial^2 I(X;Y)/(\partial \sigma_{in}^2)^2 < 0 \) (Fig. 5(c) (green)). In this subsection, we explain the relation between \( \partial \delta_q(\sigma_{in}^2)/\partial \sigma_{in}^2 \) and \( \partial^2 I(X;Y)/(\partial \sigma_{in}^2)^2 \) more intuitively by decomposing \( I(X;Y) \). In the following, we explain the INIR mechanism using the threshold model (\( \theta = 3 \)) as an example. A similar explanation applied to the linear model and the threshold model with \( \theta = 0 \). The mutual information \( I(X;Y) \) is decomposed as follows:
\[
I(X;Y) = H(Y) - H(Y|X) \tag{24}
\]
where \( H(Y) \) is the entropy of \( Y \) and \( H(Y|X) \) is the entropy of \( Y \) given by \( X \), which is expressed as:
\[
H(Y) = -\sum_Y p(y) \log_2 p(y) \tag{25}
\]
\[
H(Y|X) = -\int_X p(x) \sum_Y p(y|x) \log_2 p(y|x) dx \tag{26}
\]
where $X$ is a continuous random variable and $Y$ is a discrete random variable.

In the threshold model ($\theta = 3$), both $H(Y)$ and $H(Y|X)$ are basically increasing and convex upward with respect to the intrinsic noise $\sigma_{in}^2$ (Fig. 5(b)), and thus satisfy

$$\frac{\partial H(Y)}{\partial \sigma_{in}^2} > 0,$$

$$\frac{\partial H(Y|X)}{\partial \sigma_{in}^2} > 0 \quad (27)$$

$$\frac{\partial^2 H(Y)}{(\partial \sigma_{in}^2)^2} < 0,$$

$$\frac{\partial^2 H(Y|X)}{(\partial \sigma_{in}^2)^2} < 0 \quad (28)$$

for most values of $\sigma_{in}^2$. These inequalities are also satisfied for the extrinsic noise $\sigma_{ex}^2$.

From Eq. (24), when $\partial H(Y)/\partial \sigma_{in}^2 \neq \partial H(Y|X)/\partial \sigma_{in}^2$, $I(X;Y)$ increases with $\sigma_{in}^2$ (Fig. 5(a)(gray)). However, when $\partial H(Y)/\partial \sigma_{in}^2 < \partial H(Y|X)/\partial \sigma_{in}^2$, $I(X;Y)$ decreases with $\sigma_{in}^2$ (Fig. 5(a)(green, blue)). Therefore, we can interpret $\partial H(Y)/\partial \sigma_{in}^2$ as representing the effect of noise increasing the transmitted information, whereas $\partial H(Y|X)/\partial \sigma_{in}^2$ represents the effect of noise decreasing the transmitted information.

From Eqs. (24) and (28), when $\partial^2 I(X;Y)/(\partial \sigma_{in}^2)^2 > 0$, the following is satisfied:

$$\frac{\partial^2 H(Y)}{(\partial \sigma_{in}^2)^2} < \frac{\partial^2 H(Y|X)}{(\partial \sigma_{in}^2)^2} \quad (29)$$

(Fig. 5(b) blue). Therefore, when $\partial^2 I(X;Y)/(\partial \sigma_{in}^2)^2 > 0$, the change in $H(Y)/\partial \sigma_{in}^2$ is greater than that in $H(Y|X)/\partial \sigma_{in}^2$. Thus, we focus on the change in $H(Y|X)/\partial \sigma_{in}^2$. As $\partial^2 H(Y)/\partial \sigma_{in}^2 < 0$, $H(Y|X)/\partial \sigma_{in}^2$ decreases with $\sigma_{in}^2$. Therefore, when $\sigma_{in}^2$ is larger, the decrease in transmitted information as a result of noise is smaller. Thus, the robustness against extrinsic noise increases with the intrinsic noise.

In contrast, when $\partial^2 I(X;Y)/(\partial \sigma_{in}^2)^2 < 0$, the following is satisfied:

$$\frac{\partial^2 H(Y)}{(\partial \sigma_{in}^2)^2} > \frac{\partial^2 H(Y|X)}{(\partial \sigma_{in}^2)^2} \quad (30)$$

(Fig. 5 green). Therefore, when $\partial^2 I(X;Y)/(\partial \sigma_{in}^2)^2 < 0$, the change in $H(Y)/\partial \sigma_{in}^2$ is greater than that in $H(Y|X)/\partial \sigma_{in}^2$. Thus, we focus on the change in $H(Y|X)/\partial \sigma_{in}^2$. As $\partial^2 H(Y)/\partial \sigma_{in}^2 < 0$, $H(Y|X)/\partial \sigma_{in}^2$ decreases with $\sigma_{in}^2$. Therefore, when $\sigma_{in}^2$ is larger, the increase in transmitted information as a result of noise is smaller. Thus, the robustness against extrinsic noise decreases with the intrinsic noise.

### V. DISCUSSION

In this paper, we first analyzed the robustness against extrinsic noise in a linear model. We showed that the robustness against extrinsic noise increases with the level of intrinsic noise. Therefore, the INIR was realized in the linear model. Moreover, we showed that INIR is realized in a threshold model in which the threshold is equal to the mean of the input. These results suggest that INIR appears universally in various biochemical reactions.

We do not argue that the intrinsic noise should be enhanced to increase the robustness against extrinsic noise. This is because there is a tradeoff between the robustness against extrinsic noise and the transmitted information. The intrinsic noise increases the robustness against extrinsic noise, whereas it decreases the transmitted information. We argue that the high robustness against extrinsic noise is a by-product of the large intrinsic noise.

We further showed that SR appears in the threshold model when the threshold is much larger than the mean of the input. In this model, there is a region where the robustness against extrinsic noise decreases with the intrinsic noise, and therefore, INIR is not always realized. However, in the threshold model with SR, the robustness against extrinsic noise and the transmitted information can be simultaneously increased by reducing the intrinsic noise. Therefore, SR solves the tradeoff problem between the robustness against extrinsic noise and the transmitted information.

SR has been experimentally observed in various living systems, especially neuronal systems [19–30]. Douglass
et al. were the first to report that SR appears in living systems, namely a sensory neuron in a crayfish tail fan [19]. The sensory neuron has mechanoreceptors that fire to detect the flow of the surrounding water. They experimentally revealed that, when the flow of surrounding water is weak, additional noise improves the detection performance of the sensory neuron. Following this study, experiments have revealed that SR appears in the cercal sensory neuron of crickets [20], the hippocampal slice of rats [22–24], electrical sensory organs of paddlefish [21], human vision [25–27], and human audition [28–30]. Moreover, the threshold response is observed in many biochemical reactions [13–15], which may indicate the existence of SR. These reactions not only detect weak signals by utilizing noise, but also realize robust information transmission in the presence of extra noise.

In this paper, we discussed SR in a static system, but SR can appear in dynamical systems too [31]. Moreover, the transmitted information is conventionally quantified by the signal-to-noise ratio (SNR) in dynamical systems, rather than by the mutual information [31]. Even if the transmitted information were quantified by SNR instead of mutual information, our discussion would remain valid because we do not use the inherent properties of mutual information. Therefore, even if the transmitted information was to be quantified by SNR, the robustness against extrinsic noise and the transmitted information could be simultaneously increased by reducing the intrinsic noise when SR appears.

SR appears in the threshold model only when the threshold is much larger than the mean of the input. However, if there are several threshold systems and the entire output is determined by the summation of individual outputs, SR may appear when the mean of the input is equal to the threshold [16] [17]. Therefore, when the mean of the input is equal to the threshold, SR does not appear in the information transmission by a single cell, and appears in the information transmission by multiple cells. For example, tissues and organs composed of multiple cells respond to the same input with multiple cells. Therefore, if individual cells exhibit a nonlinear response such as a threshold response, SR appears regardless of the relation between the input and the threshold. Indeed, several theoretical studies suggest that if multiple neurons transmit the information, SR will appear even when the mean of the input is equal to the threshold [32] [35].

Acknowledgments

This work was supported by the Creation of Fundamental Technologies for Understanding and Control of Biosystem Dynamics, CREST (JPMJCR12W3), of the Japan Science and Technology Agency, and by the Japan Society for the Promotion of Science (JSPS), KAKENHI Grant Numbers 17H03800, 17H62999 and 18H03979. M.F. was funded by the JSPS, KAKENHI Grant Numbers 16K12508 and 19K20382.

Appendix A: Robustness function $\delta_q(\sigma^2_{in})$ in the linear model

1. Derivation of Eq. (5)

From Eqs. (4) and (3),

$$\frac{1}{2} \log_2 \left( 1 + \frac{1}{\delta_q(\sigma^2_{in}) + \sigma^2_{in}} \right) = q$$  \hspace{1cm} (A1)

is obtained. From Eq. (A1), we have

$$\delta_q(\sigma^2_{in}) = \frac{1}{(1 + \sigma^2_{in})^q - 1} - \sigma^2_{in}$$  \hspace{1cm} (A2)

which is the same as Eq. (5).

2. Derivation of Eq. (6)

When $\sigma^2_{in} \gg 1$, Eq. (5) can be approximated as follows:

$$\delta_q(\sigma^2_{in}) = \frac{1}{(1 + \sigma^2_{in})^q - 1} - \sigma^2_{in} \approx 1 - \frac{q}{q \sigma^2_{in}}$$  \hspace{1cm} (A3)

When $\sigma^2_{in} \ll 1$, Eq. (5) can be approximated as:

$$\delta_q(\sigma^2_{in}) = \frac{1}{(1 + \sigma^2_{in})^q - 1} - \sigma^2_{in} \approx \frac{1}{\sigma^2_{in}^{2q} - 1} - \sigma^2_{in}$$  \hspace{1cm} (A4)

When $\sigma^2_{in} \ll 1$,

$$\delta_q(\sigma^2_{in}) \approx \sigma^2_{in} - \sigma^2_{in}$$  \hspace{1cm} (A5)

When $\sigma^2_{in} \gg \sigma^2_{in}$,

$$\delta_q(\sigma^2_{in}) \approx \sigma^2_{in}$$  \hspace{1cm} (A6)

Therefore, from Eqs. (A3) and (A6), we find that

$$\delta_q(\sigma^2_{in}) \propto \begin{cases} \sigma^2_{in} & (\sigma^2_{in} \ll 1) \\ \sigma^2_{in} & (\sigma^2_{tin} \gg 1) \end{cases}$$  \hspace{1cm} (A7)

which is the same as Eq. (6). More precisely, $\delta_q(\sigma^2_{in}) \propto \sigma^2_{in}$ is satisfied not when $\sigma^2_{in} \ll 1$, but when $\sigma^2_{in} \ll \sigma^2_{in}$. 

Therefore, it folowS.
Appendix B: Mutual information $I(X; Y)$ in the threshold model ($\theta = 0$)

1. Derivation of Eq. (9)

In the threshold model, the probability density function of $X$, $p(x)$, the probability mass function of $Y$ given by $X$, $p(y|x)$, and the probability mass function of $Y$, $p(y)$, are given by the following:

$$p(x) = N(x; 0, 1)$$

$$p(y = 1|x) = \Phi \left( \frac{x - \theta}{\sqrt{\sigma^2 + \sigma^2_{in}}} \right)$$

$$p(y = 0|x) = 1 - p(y = 1|x)$$

$$p(y = 1) = \Phi \left( \frac{-\theta}{\sqrt{1 + \sigma^2 + \sigma^2_{in}}} \right)$$

$$p(y = 0) = 1 - p(y = 0|x)$$

where $\Phi(x) = \int^{x}_{-\infty} N(z; 0, 1)dz$ is the cumulative distribution function of the standard normal distribution.

The mutual information $I(X; Y)$ is calculated as follows:

$$I(X; Y) = H(Y) - H(Y|X)$$

(B6)

where

$$H(Y) = - \sum_{y \in \{0, 1\}} p(y) \log_2 p(y)$$

$$H(Y|X) = - \int_{-\infty}^{\infty} p(x) \sum_{y \in \{0, 1\}} p(y|x) \log_2 p(y|x)dx$$

(B8)

$H(Y)$ is the entropy of $Y$, and $H(Y|X)$ is the entropy of $Y$ given $X$.

When $\theta = 0$, $p(y = 1) = p(y = 0) = 1/2$. Therefore,

$$H(Y) = -2 \cdot \frac{1}{2} \log_2 \frac{1}{2} = 1$$

(B9)

$H(Y|X)$ cannot be calculated analytically. Therefore, we calculate the approximate solution of $H(Y|X)$ as follows. First, $H(Y|X)$ is calculated as:

$$H(Y|X) = \int_{-\infty}^{\infty} p(x)H(Y|X = x)dx$$

(B10)

where

$$H(Y|X = x) = - \sum_{y \in \{0, 1\}} p(y|x) \log_2 p(y|x)$$

(B11)

and $H(Y|X = x)$ is the entropy of $Y$ given $X = x$.

As $H(Y|X = x)$ is similar to the Gaussian function $a \exp(-((x - b)^2)/(2c^2))$ (Fig. 6), we approximate $H(Y|X = x)$ by the Gaussian function $a \exp(-((x - b)^2)/(2c^2))$, which gives the following:

$$H(Y|X = x) \approx \exp \left(\frac{-(x - \theta)^2}{(\pi \ln 2)(\sigma^2 + \sigma^2_{in})} \right)$$

(B12)

(see Appendix B2).

From Eqs. (B1), (B10), and (B12), $H(Y|X)$ can be calculated as follows:

$$H(Y|X) \approx \frac{1}{\sqrt{\frac{2}{(\pi \ln 2) \sigma^2 + \sigma^2_{in}}}} + 1 \times \exp \left(\frac{-\theta^2}{(\pi \ln 2) \left(\frac{2}{(\pi \ln 2)^2 \sigma^2 + \sigma^2_{in}}\right)} \right)$$

(B13)

Therefore, from Eqs. (B6), (B9), and (B13), the approximate solution of $I(X; Y)$ in the threshold model ($\theta = 0$) is given by the following:

$$I(X; Y) \approx 1 - \frac{1}{\sqrt{\frac{2}{(\pi \ln 2) \sigma^2 + \sigma^2_{in}}}} + 1$$

(B14)

which is the same as Eq. (9).

2. Derivation of Eq. (B12)

In this subsection, we explain how to approximate $H(Y|X = x)$ in the threshold model as the Gaussian function $a \exp(-((x - b)^2)/(2c^2))$. For simplicity, we define $\sigma^2 := \sigma^2_{ex} + \sigma^2_{in}$.

First, we derive $b$, which is the value of $x$ that satisfies $\partial H(Y|X = x)/\partial x = 0$. $\partial H(Y|X = x)/\partial x$ is given by:

$$\frac{\partial H(Y|X = x)}{\partial x} = -\frac{1}{\ln 2} \sum_{y \in \{0, 1\}} \{p'(y|x) \ln p(y|x) + p'(y|x)\}$$

(B15)
\( p'(y|x) \) represents \( \partial p(y|x)/\partial x \), which is given by:

\[
p'(y = 1|x) = \frac{\partial}{\partial x} \Phi \left( \frac{x - \theta}{\sigma} \right) = \frac{1}{\sigma} N \left( \frac{x - \theta}{\sigma}, 0, 1 \right)
\]

\[
p'(y = 0|x) = -p'(y = 1|x) \tag{B16}
\]

From Eq. \( \text{[B16]} \), \( \partial H(Y|X = x)/\partial x \) is calculated as follows:

\[
\frac{\partial H(Y|X = x)}{\partial x} = - \frac{1}{\ln 2} p'(y = 1|x) \times \{ \ln p(y = 1|x) - \ln p(y = 0|x) \} \tag{B17}
\]

From \( p'(y = 1|x) > 0 \), when \( \partial H(Y|X = x)/\partial x = 0 \),

\[
p(y = 1|x) = p(y = 0|x) \tag{B18}
\]

is satisfied. Accordingly, \( b = \theta \).

Next, we derive \( c \). \( c \) satisfies

\[
\frac{1}{\epsilon^2} = \left. - \frac{\partial^2}{\partial x^2} \ln H(Y|X = x) \right|_{x=\theta} \tag{B19}
\]

\( \partial \ln H(Y|X = x)/\partial x \) is given by the following:

\[
\frac{\partial}{\partial x} \ln H(Y|X = x) = \frac{H'(Y|X = x)}{H(Y|X = x)} \tag{B20}
\]

Therefore, \( \partial^2 \ln H(Y|X = x)/\partial x^2 \) is given by:

\[
\frac{\partial^2}{\partial x^2} \ln H(Y|X = x) = \frac{H''(Y|X = x)H(Y|X = x) - H'(Y|X = x)^2}{H(Y|X = x)^2} \tag{B21}
\]

When \( x = \theta \),

\[
H(Y|X = x) = 1 \tag{B22}
\]

\[
H'(Y|X = x) = 0 \tag{B23}
\]

\[
H''(Y|X = x) = -\frac{2}{(\pi \ln 2)\sigma^2} \tag{B24}
\]

is obtained. Therefore, \( \partial^2 \ln H(Y|X = x)/\partial x^2|_{x=\theta} \) is written as:

\[
\left. \frac{\partial^2}{\partial x^2} \ln H(Y|X = x) \right|_{x=\theta} = -\frac{2}{(\pi \ln 2)\sigma^2} \tag{B25}
\]

Therefore, from Eq. \( \text{[B19]} \),

\[
c^2 = \frac{(\pi \ln 2)\sigma^2}{2} \tag{B26}
\]

is obtained.

Finally, we derive \( a \). \( a \) is given by \( H(Y|X = x)|_{x=\theta} \).

Therefore, \( a = 1 \).

Thus, the approximate solution of \( H(Y|X = x) \) is given by:

\[
H(Y|X = x) \approx \exp \left( -\frac{(x - \theta)^2}{(\pi \ln 2)\sigma^2} \right) \tag{B27}
\]

which is the same as Eq. \( \text{[B12]} \).

### Appendix C: Robustness function \( \delta_q(\sigma^2_{\text{in}}) \) in the threshold model \( (\theta = 0) \)

1. **Derivation of Eq. \( (\text{[10]} \)**

From Eqs. \( (\text{[1]} \) and \( (\text{[9]} \),

\[
1 - \frac{1}{\sqrt{(\pi \ln 2)} \left( \delta_q(\sigma^2_{\text{in}}) + \sigma^2_{\text{in}} \right)^{-1} + 1} \approx qI_0 \tag{C1}
\]

is satisfied, where

\[
I_0 = 1 - \frac{1}{\sqrt{(\pi \ln 2)} \sigma^2_{\text{in}} + 1} \tag{C2}
\]

From Eq. \( (\text{[C1]} \),

\[
\delta_q(\sigma^2_{\text{in}}) \approx \left\{ 1 - qI_0 \right\}^2 \frac{2}{\left( 2 - qI_0 \right) qI_0 (\pi \ln 2) - \sigma^2_{\text{in}}} \tag{C3}
\]

is obtained, which is the same as Eq. \( (\text{[10]} \).

2. **Derivation of Eq. \( (\text{[12]} \)**

When \( \sigma^2_{\text{in}} \gg \frac{2}{\pi \ln 2} \), Eq. \( (\text{[11]} \) can be approximated as follows:

\[
I_0 = 1 - \frac{1}{\sqrt{(\pi \ln 2)} \sigma^2_{\text{in}} + 1} \approx 1 - \frac{1}{(\pi \ln 2)} \sigma^2_{\text{in}} \tag{C4}
\]

Therefore, Eq. \( (\text{[10]} \) can be approximated as:

\[
\delta_q(\sigma^2_{\text{in}}) \approx \frac{\left( 1 - qI_0 \right)^2}{\left( 2 - qI_0 \right) qI_0 (\pi \ln 2) - \sigma^2_{\text{in}}} \approx \frac{1}{\sqrt{(\pi \ln 2)} \sigma^2_{\text{in}}} \tag{C5}
\]

When \( \sigma^2_{\text{in}} \ll \frac{2}{\pi \ln 2} \), Eq. \( (\text{[11]} \) can be approximated as follows:

\[
I_0 = 1 - \frac{1}{\sqrt{(\pi \ln 2)} \sigma^2_{\text{in}} + 1} \approx 1 - \sqrt{(\pi \ln 2)} \frac{\sigma^2_{\text{in}}}{2} \tag{C6}
\]

Therefore, Eq. \( (\text{[10]} \) can be approximated as:

\[
\delta_q(\sigma^2_{\text{in}}) \approx \frac{\left( 1 - qI_0 \right)^2}{\left( 2 - qI_0 \right) qI_0 (\pi \ln 2) - \sigma^2_{\text{in}}} \approx \left\{ \frac{1 - q}{q} + \frac{1}{\sqrt{(\pi \ln 2)} \sigma^2_{\text{in}}} \right\}^2 \tag{C7}
\]

\[
\frac{2 - q}{q} \left( \frac{\sigma^2_{\text{in}}}{(\pi \ln 2)} \right)^2 - \sigma^2_{\text{in}}
\]
Organizing the equation above based on \( \sigma_{in} \) gives:

\[
\delta_q(\sigma_{in}^2) = \frac{4}{(\pi \ln 2)} \left( \frac{1-q}{2-q} \right) \left( \frac{1-q}{2} \right) + \frac{\sqrt{(\pi \ln 2)}}{2} \sigma_{in} + \left( \frac{\pi \ln 2}{2} \right) \sigma_{in}^2
\]

(C8)

From \( \sigma_{in}^2 \ll \frac{2}{\pi \ln 2} \), the equation above can be approximated as follows:

\[
\delta_q(\sigma_{in}^2) \approx 2 \frac{2}{(\pi \ln 2)} \left( \frac{1-q}{2-q} \right) \sigma_{in}
\]

(C10)

When \( \frac{(1-q)^2}{4q} \frac{2}{\pi \ln 2} \ll \sigma_{in}^2 \ll \frac{2}{\pi \ln 2} \), Eq. (C9) can be approximated as:

\[
\delta_q(\sigma_{in}^2) \approx 2 \frac{(1-q)^2}{q(2-q)}
\]

(C11)

Therefore, from Eqs. (C9), (C10), and (C11), we obtain

\[
\delta_q(\sigma_{in}^2) \approx \begin{cases} 
\{1-q \theta \}^2 \frac{2}{\{2-q \theta \} q \theta (\pi \ln 2)} - \sigma_{in}^2 \\
\sigma_{in}^2 \sigma_{in} \left( \frac{(1-q)^2}{4q} \frac{2}{\pi \ln 2} \ll \sigma_{in}^2 \ll \frac{2}{\pi \ln 2} \right) \\
\const \left( \sigma_{in}^2 \ll \min \left( \frac{(1-q)^2}{4q} \frac{2}{\pi \ln 2} ; \frac{2}{\pi \ln 2} \right) \right)
\end{cases}
\]

(C12)

which is the same as Eq. (12).

Appendix D: Mutual information \( I(X;Y) \) in the threshold model \( (\theta \neq 0) \)

1. Derivation of Eq. (13)

The mutual information \( I(X;Y) \) is calculated as follows:

\[
I(X;Y) = H(Y) - H(Y|X)
\]

(D1)

where

\[
H(Y) = - \sum_{y \in \{0,1\}} p(y) \log_2 p(y)
\]

(D2)

\[
H(Y|X) = - \int_{-\infty}^{\infty} p(x) \sum_{y \in \{0,1\}} p(y|x) \log_2 p(y|x) dx
\]

(D3)

\( H(Y) \) is the entropy of \( Y \), and \( H(Y|X) \) is the entropy of \( Y \) given \( X \).

In Appendix [12] by approximating \( H(Y|X = x) \) using a Gaussian function, we obtained the approximate solution of \( H(Y|X) \) as:

\[
H(Y|X) \approx \frac{1}{\sqrt{\frac{2}{(\pi \ln 2)} \left( \sigma_{ex}^2 + \sigma_{in}^2 \right)^{-1} + 1}}
\]

(D4)

\[
\times \exp \left( - \frac{\theta^2}{(\pi \ln 2) \left( \frac{2}{(\pi \ln 2)} \left( \sigma_{ex}^2 + \sigma_{in}^2 \right) \right)} \right)
\]

(D5)

In the same way, we can approximate \( H(Y) \) using a Gaussian function to give:

\[
H(Y) \approx \exp \left( - \frac{\theta^2}{(\pi \ln 2) \left( 1 + \sigma_{ex}^2 + \sigma_{in}^2 \right)} \right)
\]

(D6)

Therefore, \( I(X;Y) \) can be approximated as follows:

\[
I(X;Y) \approx \exp \left( - \frac{\theta^2}{(\pi \ln 2) \left( 1 + \sigma_{ex}^2 + \sigma_{in}^2 \right)} \right)
\]

(D7)

\[
- \frac{1}{\sqrt{\frac{2}{(\pi \ln 2)} \left( \sigma_{ex}^2 + \sigma_{in}^2 \right)^{-1} + 1}} \times \exp \left( - \frac{\theta^2}{(\pi \ln 2) \left( \frac{2}{(\pi \ln 2)} \sigma_{ex}^2 + \sigma_{in}^2 \right) \right)
\]

(D8)

Here, \( \frac{2}{\pi \ln 2} = 0.91844 \cdots \approx 1 \). Therefore, \( H(Y) \) can be approximated as:

\[
H(Y) \approx \exp \left( - \frac{\theta^2}{(\pi \ln 2) \left( \frac{2}{(\pi \ln 2)} \sigma_{ex}^2 + \sigma_{in}^2 \right) \right)
\]

(D9)

Thus, the mutual information \( I(X;Y) \) can be approximated as follows:

\[
I(X;Y) \approx \left( 1 - \frac{1}{\sqrt{\frac{2}{(\pi \ln 2)} \left( \sigma_{ex}^2 + \sigma_{in}^2 \right)^{-1} + 1}} \right)
\]

(D10)

\[
\times \exp \left( - \frac{\theta^2}{(\pi \ln 2) \left( \frac{2}{(\pi \ln 2)} \sigma_{ex}^2 + \sigma_{in}^2 \right) \right)
\]

(D11)

which is the same as Eq. (13).

2. Derivation of the condition under which SR appears: \( |\theta| > 2 \)

We can derive the condition for the appearance of SR from the approximate solution of the mutual information
\( I(X;Y) \). For simplicity, we consider the case \( \sigma_{xx}^2 = 0 \). Note that we can expand this to the case \( \sigma_{xx}^2 \geq 0 \) easily.

When SR appears, the mutual information \( I(X;Y) \) has a local maximum with respect to the intrinsic noise \( \sigma_{in}^2 \). Therefore, when SR appears, there is some value of \( \sigma_{in}^2 \) for which \( \partial I(X;Y)/\partial \sigma_{in}^2 = 0 \). Therefore, we derive the condition that ensures \( \partial I(X;Y)/\partial \sigma_{in}^2 = 0 \).

\[ \partial I(X;Y)/\partial \sigma_{in}^2 \text{ is given by the following:} \]

\[ \approx -\left\{ \left( \frac{2}{(\pi \ln 2)^2} \sigma_{in}^{-2} + 1 \right) - \theta^2 \sqrt{\frac{2}{(\pi \ln 2)^2} \sigma_{in}^{-2} + 1 + \theta^2} \right\} \]

\[ \times \sigma_{in}^{-4} \exp \left( -\frac{\theta^2}{(\pi \ln 2)^2} \frac{\frac{2}{(\pi \ln 2)^2} \sigma_{in}^{-2} + 1}{\sigma_{in}^{-2}} \right) \]

(D9)

Therefore, when \( \partial I(X;Y)/\partial \sigma_{in}^2 = 0 \) holds,

\[ \left( \frac{2}{(\pi \ln 2)^2} \sigma_{in}^{-2} + 1 \right) - \theta^2 \sqrt{\frac{2}{(\pi \ln 2)^2} \sigma_{in}^{-2} + 1 + \theta^2} = 0 \]

(D10)

is satisfied. Defining \( A = \sqrt{\frac{2}{(\pi \ln 2)^2} \sigma_{in}^{-2} + 1} \),

\[ A^2 - \theta^2 A + \theta^2 = 0 \]

(D11)

is the necessary condition for \( A \) to satisfy Eq. (D11).

As \( A \) is a real number, \( D \) must satisfy \( D \geq 0 \). Therefore,

\[ \theta = 0, \quad |\theta| \geq 2 \]

(D13)

is the necessary condition for \( \theta \) to satisfy Eq. (D11).

When \( \theta = 0 \), the solution of Eq. (D11) is \( A = 0 \). This is contrary to \( A > 1 \). Therefore, when \( \theta = 0 \), there is no value of \( \sigma_{in}^2 \) that satisfies Eq. (D11), and SR does not appear.

When \( |\theta| \geq 2 \), the solutions of Eq. (D11) are given by the following:

\[ A_{\pm} = \frac{\theta^2 \pm \sqrt{\theta^4 - 4 \theta^2}}{2} = \frac{\theta^2}{2} \left\{ 1 \pm \sqrt{1 - \frac{4}{\theta^2}} \right\} \]

(D14)

From simple calculations, we can confirm that \( A_{\pm} \) satisfies \( A_{\pm} > 1 \). Therefore, \( |\theta| > 2 \) is the necessary and sufficient condition for the appearance of SR.

Furthermore, we can derive the intrinsic noise \( \sigma_{in}^2 \) that gives the local maximum of \( I(X;Y) \), defined as \( \sigma_{in,\text{max}}^2 \).

\[ \sigma_{in,\text{max}}^2 = \frac{2}{(\pi \ln 2)^2} \left[ A_{\pm}^2 - 1 \right]^{-1} \]

(D15)

As \( \sigma_{in}^2 = \sigma_{in,\pm}^2 \) satisfies \( \partial I(X;Y)/\partial \sigma_{in}^2 = 0 \), \( \sigma_{in,\pm}^2 \) gives the local maximum or local minimum of \( I(X;Y) \).

From numerical calculations, it can be clarified that \( \sigma_{in,\pm}^2 \) gives the local maximum of \( I(X;Y) \) and \( \sigma_{in,-}^2 \) gives the local maximum of \( I(X;Y) \). Therefore,

\[ \sigma_{in,\text{max}}^2 = \sigma_{in,-}^2 := \frac{2}{(\pi \ln 2)^2} \left[ A_{-}^2 - 1 \right]^{-1} \]

(D16)

is satisfied. This approximate solution matches the numerical solution well (Fig. 7).

[1] M. B. Elowitz, Science 297, 1183 (2002), ISSN 00368075, URL http://www.sciencemag.org/cgi/doi/10.1126/science.1070919
[2] A. Bar-Evren, J. Paulsson, N. Maheshri, M. Carmi, E. O'Shea, Y. Pilpel, and N. Barkai, Nature Genetics 38, 636 (2006), ISSN 10614036.
[3] J. R. Newman, S. Ghaemmaghami, J. Ihmels, D. K. Breslow, M. Noble, J. L. DeRisi, and J. S. Weissman, Nature 441, 840 (2006), ISSN 14764687.
[4] Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie, Science 329, 533 (2010), ISSN 0036-8075, URL http://www.sciencemag.org.
