Information-transfer characteristics in network motifs

AUTHOR(S):

Mori, Fumito; Okada, Takashi

PHYSICAL REVIEW RESEARCH **5**, 013037 (2023)

Information-transfer characteristics in network motifs

Fumito Mori *
Faculty of Design, Kyushu University, Fukuoka 815-8540, Japan
and Education and Research Center for Mathematical and Data Science, Kyushu University, Fukuoka 819-0395, Japan

Takashi Okada †
Institute for Life and Medical Sciences, Kyoto University, Kyoto 606-8507, Japan;
RIKEN Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), Wako 351-0198, Japan;
and Department of Physics, and Department of Integrative Biology, University of California, Berkeley, California 94720, USA

(Received 27 July 2022; accepted 5 January 2023; published 24 January 2023)

Information processing in biological systems is realized by the appropriate transmission of information flows over complex networks, such as gene regulatory, signal transduction, and neural networks. These information flows are affected by the input-signal characteristics and structural properties of network systems, such as the network topology, regulation rules, and intrinsic and environmental noise. Many biological networks frequently include several typical patterns called network motifs, which are considered to play important roles in biological functions. However, their information-theoretic properties, particularly the dependence of the information flows in each network on the input signal, remain poorly understood. In our previous study [Mori and Okada, Phys. Rev. Res. **2**, 043432 (2020)], we developed a graphical expansion method to describe transfer entropy (TE), a measure of information flow, in Boolean networks in terms of multiple information pathways. There, the input signal was limited to a simple case, and the effect of the input-signal characteristics on TE was not clarified. In this paper, we improve our method to render it applicable to Boolean networks that receive input signals with arbitrary stochastic characteristics. Our formula expresses how TE is determined by the input-signal characteristics, the assignment of Boolean functions, and the noise magnitude. We find that, in both positive and negative feedback loops, TE hardly depends on the signal timescale. In contrast, coherent and incoherent feedforward loops show low- and high-pass filtering properties, respectively, for a time-varying signal, which is consistent with previous reports. The emergence of either low- or high-pass filtering is determined by the Fourier components of the Boolean functions on specific pathways transmitting information flows. Thus our formula reveals the mechanism of information transfer in network motifs and provides insights into the origin of information processing in biological networks.

DOI: 10.1103/PhysRevResearch.5.013037

I. INTRODUCTION

Network motifs are basic units that express the local connection patterns in directed complex networks, such as gene regulatory, signal transduction, and neural networks [1–4]. Feedforward and feedback loops are representative motifs, which are considered to provide various types of biological functions [5,6]. For example, the positive feedforward loop (PFFL) is a persistent detector that filters high-frequency input [5,7,8], whereas the negative feedforward loop (NFFL) is a pulse generator that blocks long-timescale signals [5,7]. In addition, positive feedback and negative feedback loops (PFBL and NFBL) are employed as a lock-on switch [5,9,10] and oscillation generator [5,11–13], respectively. Attempts are being made to clarify the relationships between the static structure of the motifs and their dynamical properties, as one of the central issues in network and biological sciences [14–20].

Although the biological functions performed by motifs have been much discussed, many aspects of their information-transfer characteristics remain unexplained theoretically and experimentally. For instance, it is difficult to determine whether a certain motif can process information from an input stimulus with a short timescale or from an environmental signal with slow variation. Moreover, the relationship between the information flows in motifs and their network topology has not been clarified. If each motif has a respective information-transfer characteristic (Fig. 1), a question arises as to what is the internal structure of the motif that generates this characteristic? Are the characteristics robust against partial modification of the regulation rules? Can we provide optimal design principles for maximizing the information flows in motifs? To address these issues, theoretical predictions are necessary, using mathematical models.

*mori@design.kyushu-u.ac.jp
†okada.takashi.3z@kyoto-u.ac.jp

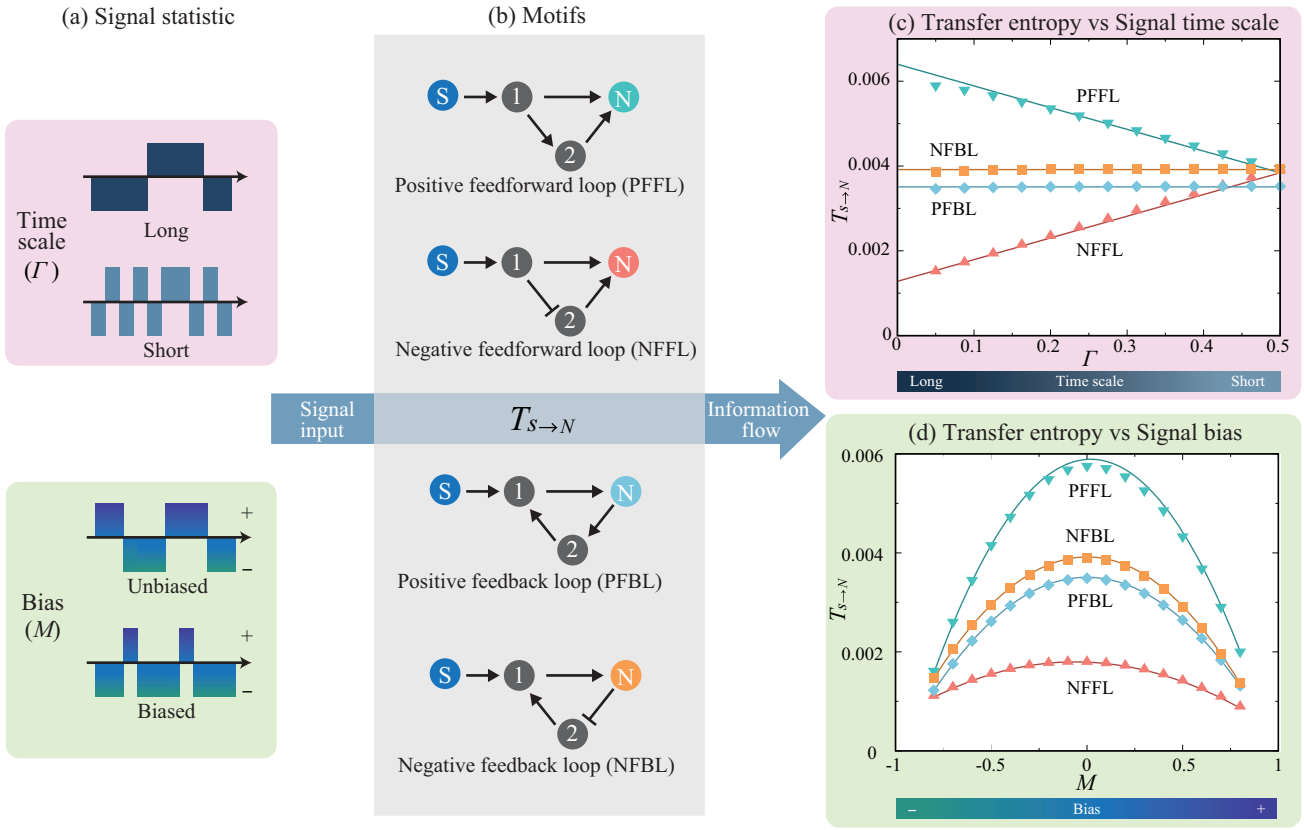

FIG. 1. (a) Schematic of the statistics of the signal input to the motifs. The signal timescale Γ and bias M are illustrated. (b) Representative network motifs. In each network motif, we consider the transfer entropy (TE) $T_{s\to N}$ from the signal source s to an output vertex labeled N. (c) TE comparison among the motifs as a function of the signal timescale Γ when signal bias is absent ($M = 0$). The error rates are uniformly distributed ($\psi_i = 0.3$). The symbols represent the numerically obtained TE, whereas the solid lines represent the theoretical predictions through our pathway expansion method. (d) TE as a function of the signal bias M when $\Gamma = 0.1$ and $\psi_i = 0.3$.

Due to their high degree of abstraction, Boolean network models [21] that approximately describe the dynamics in gene regulatory networks and neural networks enable theoretical investigation of the information properties in networks [22–31]. However, analytical understanding of the information flow mechanism remains limited. In particular, a systematic formula for understanding how input signals having certain characteristics are transmitted in a network with specific topology and regulations is still lacking. For example, in Ref. [32], the information flows and related thermodynamic quantities were computed exhaustively for small motifs, but through numerical computations. In Refs. [33–36], information propagation was analyzed using annealed approximations, which randomly reassign inputs and functions to all nodes at each time step. In Ref. [37], the influence of nodes was analyzed using a quenched approximation. However, as the annealed approximations dispose of network connectedness, focusing on the ensemble properties of random networks, and the quenched approximation neglects dynamical correlations among variables, focusing on the average dynamical behavior, these analyses provide little insight into the information flow mechanism in an individual network.

The major obstacles preventing analytical understanding of the information flows over networks are summarized as follows: (i) Because information flow is a nonlinear phenomenon, the total amount of information flow from the signal source to the output vertex cannot be considered as the sum of flows along the routes connecting the signal source to the output vertex. Collective behaviors or interactions must be analyzed [38–40]. (ii) The computation of information flows generally involves a large state space and high-dimensional matrices. It is difficult to extract the essence of the complex phenomena from matrix calculations. (iii) Because various factors, such as input-signal characteristics [Fig. 1(a)], network topology [Fig. 1(b)], regulation rules, and noise effects, are involved in information flows [Figs. 1(c) and 1(d)], systematic and comprehensive analysis and elucidation of the roles of these factors are challenging. Although an analytical graphical expansion method was developed in Ref. [41], the input signal was limited to a simple case, and the relationship between the signal properties and information flows was not clarified.

Thus this study improves the graphical expansion method to render it applicable to Boolean networks that receive input signals with arbitrary stochastic characteristics. Our formula expresses information flows in terms of pathways (precisely defined in Sec. IV). We show that only combinations of pathways that satisfy certain graphical conditions can contribute to the information flows, and each contribution is determined separately from the input-signal characteristics, the assignment of Boolean functions, and the noise magnitude (Fig. 2). Note that the definition of pathways and the graphical

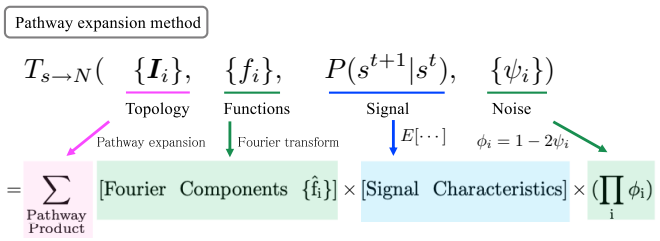

FIG. 2. Pathway expansion method of TE in a stochastic Boolean network with an arbitrary signal. Although TE in the network depends on various model assumptions, our formula can be separately described in terms of the signal characteristics, Fourier components of the Boolean functions, and noise parameters when the network topology is fixed.

conditions are slightly changed from those in our previous study [41]. Applying the method to network motifs, the origins of the information-transfer characteristics in the PFFL, NFFL, PFBL, and NFBL are revealed. In particular, Fourier transformation of the Boolean functions participating in the pathway interaction indicates whether a low- or high-pass filter property emerges in feedforward loops (FFLs).

II. MODEL AND DEFINITION

We consider a stochastic Boolean network model that receives the input signal from a signal source that provides a binary signal [Figs. 1(a) and 1(b)]. It is assumed that the signal characteristics are arbitrary and that the signal can be generated by either a discrete-time Markov or non-Markov process. We denote the signal state at time t as s^t, where $s^t = 1$ represents the ON or up state and $s^t = 0$ represents the OFF or down state. As an example, we employ a signal generated by a Markov information source:

$$s^{t+1} = \begin{cases} s^t & \text{with probability } 1-\gamma(s^t) \\ \bar{s}^t & \text{with probability } \gamma(s^t), \end{cases} \tag{1}$$

where the transition probability γ $(0 < \gamma \leqslant \frac{1}{2})$ can depend on the current state s^t. The negation of Boolean variable s is denoted by $\bar{s}$. Let $\gamma(0) = \gamma_0$ and $\gamma(1) = \gamma_1$; then, the stationary distribution of s^t is given by $P_{\rm st}(s^t = s) = \frac{\gamma(\bar{s})}{\gamma_1+\gamma_0}$. The difference between the probabilities of the up and down states indicates the signal bias:

$$M \equiv \frac{\gamma_0 - \gamma_1}{\gamma_0 + \gamma_1} \quad (-1 < M < +1), \tag{2}$$

i.e., the up state occurs more frequently than the down state when $M > 0$ and less frequently when $M < 0$. Moreover, the harmonic mean of γ_0 and γ_1 gives the signal timescale:

$$\Gamma \equiv \frac{2\gamma_0\gamma_1}{\gamma_0+\gamma_1} \quad \left(0 < \Gamma \leqslant \frac{1}{2}\right). \tag{3}$$

Note that $\frac{1}{\Gamma}$ is equivalent to half the average period of the stochastic-signal oscillation. Thus the signal state changes with the long timescale when $\Gamma \gtrsim 0$ and with the short timescale when $\Gamma \lesssim \frac{1}{2}$.

In a stochastic Boolean network, the state of vertex i is synchronously updated according to

$$x_i^{t+1} = \begin{cases} f_i\left(\boldsymbol{x}_{\boldsymbol{I}_i}^t\right) & \text{with probability } 1-\psi_i \\ \bar{f}_i\left(\boldsymbol{x}_{\boldsymbol{I}_i}^t\right) & \text{with probability } \psi_i, \end{cases} \tag{4}$$

where x_i^t is a Boolean variable of vertex i at time t and f_i is a Boolean function assigned to vertex i. We assign N to the output vertex and $1, \ldots, (N-1)$ to the other vertices arbitrarily [Fig. 1(b)]. A set of input vertices to vertex i is denoted as $\boldsymbol{I}_i$, and a set of Boolean variables of $\boldsymbol{I}_i$ at t is denoted as $\boldsymbol{x}_{\boldsymbol{I}_i}^t$. Note that $\boldsymbol{x}_{\boldsymbol{I}_i}^t$ can include s^t. When $|\boldsymbol{I}_i| = 1$, f_i is either the identity function or the negation function. Parameter ψ_i $(0 < \psi_i \leqslant \frac{1}{2})$ represents the error rate of vertex i, i.e., the dynamics of vertex i become deterministic in the limit of $\psi_i \to 0$, whereas they are completely random when $\psi_i = \frac{1}{2}$.

The entropy and conditional entropy are defined as $H[\boldsymbol{x}] = -\sum_{\boldsymbol{x}} P(\boldsymbol{x}) \ln P(\boldsymbol{x})$ and $H[\boldsymbol{x}|\boldsymbol{y}] = H[\boldsymbol{x},\boldsymbol{y}] - H[\boldsymbol{y}]$, respectively, where $P(\boldsymbol{x})$ is the probability distribution of a set of variables $\boldsymbol{x}$. We assume a stationary signal $P(s^t) = P_{\rm st}(s^t)$. Then, we employ transfer entropy (TE) for quantifying the information flows [42] from the signal source to the output vertex in the stationary state, which is defined as

$$T_{s\to N} = H\left[x_N^t \middle| \boldsymbol{x}_N^-\right] - H\left[x_N^t \middle| \boldsymbol{x}_N^-, \boldsymbol{s}^-\right], \tag{5}$$

where $\boldsymbol{s}^- \equiv (s^{t-l}, \ldots, s^{t-1})$ and $\boldsymbol{x}_N^- \equiv (x_N^{t-l}, \ldots, x_N^{t-1})$ for sufficiently large l. Although $T_{s\to N}$ depends on l for small l, it will be confirmed later that the l dependency disappears for large l. TE generally depends on signal statistics, network topology, assignment of Boolean functions, and error rates. This complicated dependence structure is dismantled in our formulation (Fig. 2).

III. NUMERICAL RESULTS

We employ four representative network motifs, namely, PFFL (coherent motif), NFFL (incoherent motif), PFBL, and NFBL, as shown in Fig. 1(b). In this section, we assign the AND function to vertices with two input variables, in all the motifs. Vertex $i = 2$ in the NFFL and NFBL has the negation function $f_2(x) = \bar{x}$, and the other vertices with an input variable have the identity function $f(x) = x$. The error rates are assumed to be homogeneously distributed: $\psi_i = 0.3$ for all i. The input signal is assumed to be generated by a Markov information source [Eq. (1)].

We numerically calculated Eq. (5) with $l = 5$ for the motifs. It will be shown later that $l = 5$ is sufficient for l independency of TE in the motifs. The calculation steps are as follows: The steady-state probability distribution was first obtained from the eigenvector of the transition matrix corresponding to the eigenvalue of 1. Next, $P(s^{t-l}, \boldsymbol{x}^{t-l})$ was assumed as the steady state, where $\boldsymbol{x}^t \equiv (x_1^t, \ldots, x_N^t)$. Furthermore, the transition matrix of the system and $P(s^{t-l}, \boldsymbol{x}^{t-l})$ were multiplied to obtain $P(s^{t-l}, \boldsymbol{x}^{t-l}, s^{t-l+1}, \boldsymbol{x}^{t-l+1})$. Repeating this, $P(s^{t-l}, \boldsymbol{x}^{t-l}, \ldots, s^t, \boldsymbol{x}^t)$ was obtained. Finally, $P(s^{t-l}, \boldsymbol{x}^{t-l}, \ldots, s^t, \boldsymbol{x}^t)$ was marginalized to obtain $P(x_N^t, \boldsymbol{x}_N^-, \boldsymbol{s}^-)$, and substituting $P(x_N^t, \boldsymbol{x}_N^-, \boldsymbol{s}^-)$ into Eq. (5) yielded $T_{s\to N}$.

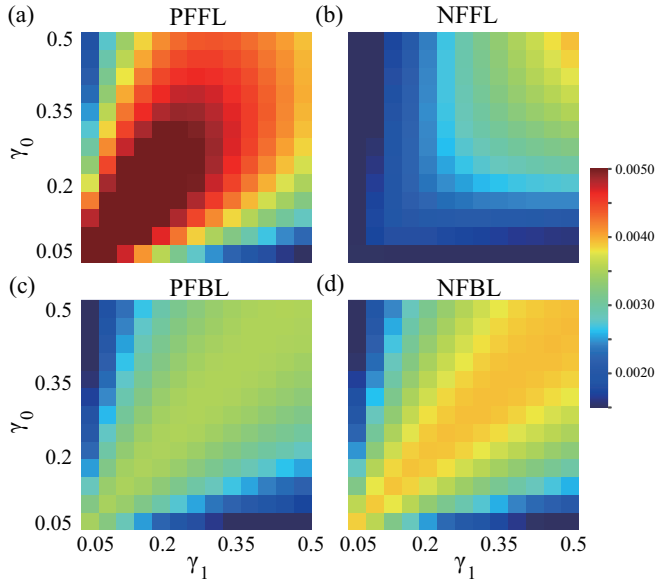

FIG. 3. Two-dimensional heat map of the numerically obtained TE as a function of γ_0 and γ_1 in the (a) PFFL, (b) NFFL, (c) PFBL, and (d) NFBL.

Figures 3(a)–3(d) show the numerically obtained TE as a function of γ_0 and γ_1 for the four motifs. While the PFFL and NFFL exhibit obviously different heat maps, the PFBL and NFBL have similar signal-statistics dependence of the TE.

We first focus on the diagonal line $\gamma_0 = \gamma_1$ in these figures and inspect the signal-timescale dependence of TE when there is no signal bias. In Fig. 1(c), the symbols represent the numerically obtained $T_{s\to N}$ as a function of Γ for $M = 0$ for the four motifs. With the increase in Γ, TE decreases in the PFFL, whereas it increases in the NFFL. In the PFBL and NFBL, TE appears to be independent of Γ.

Next, we extract only the signal-bias effect on TE, fixing Γ. In Fig. 1(d), the symbols represent the numerically obtained $T_{s\to N}$ as a function of M under $\Gamma = 0.1$. It can be observed that TE is maximized at approximately $M = 0$ in all the motifs. When the signal is highly biased, TE becomes small because $T_{s\to N} \leqslant H[\boldsymbol{s}^-]$ and the signal ambiguity $H[\boldsymbol{s}^-]$ disappears with the increase in the bias.

We discuss the above numerical results, referring to the previous reports on the biological functions in motifs. The PFFL is called a persistence detector [5,7,8] because it can respond only to a persistent signal and it does not show a clear response to an impulsive one. Conversely, the NFFL behaves as a pulse generator [5,7], where its response to a persistent signal is transient. These functions are consistent with our numerical results, indicating that the PFFL and NFFL function as a low- and high-pass filter, respectively, for information transfer. Regarding feedback loops, it is known that the PFBL performs lock-on switching [5,9,10], whereas the NFBL generates oscillations [5,11–13]. Therefore the ambiguity of the output variable of the NFBL is expected to be larger than that of the PFBL. As shown later, the difference in the ambiguity of the output variable causes a difference in TE of the NFBL and PFBL.

IV. DIAGRAMMATIC EXPANSION

As discussed in the previous section, our numerical results provide supportive evidence that the PFFL, NFFL, PFBL, and NFBL act as a low-pass filter, high-pass filter, lock-on switch, and oscillator, respectively. However, these numerical results were obtained only for specific Boolean functions and error rates. Then, the following questions arise: How robust are these results against the modifications of the model, such as the replacement of the AND function by the OR function, and the heterogenization of the error rates? Can the aforementioned behaviors be explained in terms of the network-motif structures such as the network topology and regulation functions? To address these questions systematically, general formulations for the information transfer are required.

In this section, we extend the diagrammatic expansion method for TE proposed in Ref. [41], i.e., render it applicable to a system with arbitrary input-signal characteristics. This method reveals how the information-transfer characteristics are realized from the topology and regulations of the motifs.

General formula for TE

In preparation for the formulation, certain assumptions about the network topology are made. Because we focus on network systems transmitting information flows, a direct link from the signal source to the output vertex is prohibited. The signal source is independent, i.e., it does not take any input from the other vertices. Moreover, the output variable x_N is in accordance with either one of the following two cases: (i) As in the FFL, it is never employed as an input. (ii) As in the feedback loop (FBL), it is input to a one-variable function of another node.

We derive the TE formula with a sufficiently large l. The details of the derivations are presented in Sec. A of the Supplemental Material (SM) [43]. We first convert the error rate ψ_i into an expansion parameter ϕ_i ($0 \leqslant \phi_i < 1$):

$$\phi_i = 1 - 2\psi_i. \tag{6}$$

Note that when $\phi_i = 0$ for all i, the probability distribution of all the variables except the input signal is uniform, and $P(x_i) = P(\overline{x}_i)$ for any i. This symmetry is the key for considerably reducing the computational complexity. Using ϕ_i, the following conditional probability is expressed as

$$P\left(x_i^t|\boldsymbol{x}_{\boldsymbol{I}_i}^{t-1}\right) = \tfrac{1}{2}\left(1 + \phi_i^t D\left[f_i\left(\boldsymbol{x}_{\boldsymbol{I}_i}^{t-1}\right), x_i^t\right]\right). \tag{7}$$

Here, $D[x, y]$ is the function defined by $D[x, x] = 1$ and $D[x, \overline{x}] = -1$. In Eq. (7), we append t to ϕ_i as a marker to assist in our formulation.

The pathways utilized to express TE are defined in a temporal graph [41]. A similar graph was used in Refs. [44–47]. Figures 4 and 5 display the temporal graphs for the FFL and FBL, respectively. These graphs represent the connections between the stochastic variables and do not specify the input signal, Boolean functions, or error rates. Here, variables s^t and x_i^t are represented by a circle node ($\circ$) at t. Signal nodes $\{s^{t'}|t - l \leqslant t' \leqslant t\}$ are located on the top line. Output nodes $\{x_N^{t'}|t - l \leqslant t' \leqslant t\}$ specially represented by double circles are located on the bottom line. The other nodes $\{x_i^{t'}|i \neq$

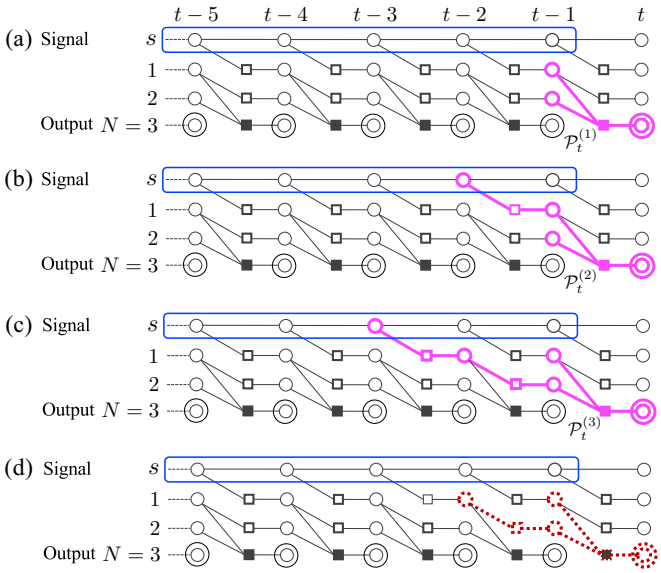

FIG. 4. Temporal graphs for the FFL. (a) Example of a bond: $\phi_N^t D[f_N(x_1^{t-1}, x_2^{t-1}), x_N^t]$ (magenta lines). This is also an existent pathway consisting of a single bond. (b) and (c) Other examples of existent pathways (magenta lines). The abovementioned existent pathways are labeled $\mathcal{P}_t^{(1)}$, $\mathcal{P}_t^{(2)}$, and $\mathcal{P}_t^{(3)}$, respectively, because they contain one, two, and three bonds (squares), respectively, and their rightmost node is the output node at t. (d) Example of a nonexistent pathway (dotted lines). This does not satisfy condition C2 (shown later) because internal node x_1^{t-2} is connected to only one open square within the pathway.

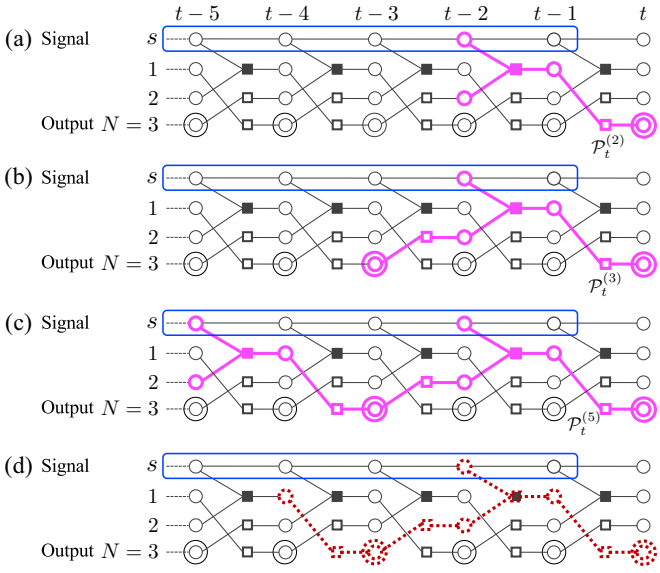

FIG. 5. Temporal graphs for the FBL. (a)–(c) Existent pathways (magenta lines) labeled $\mathcal{P}_t^{(2)}$, $\mathcal{P}_t^{(3)}$, and $\mathcal{P}_t^{(5)}$. (d) Nonexistent pathway (dotted lines). This does not satisfy condition C2 (shown later) because the internal node x_1^{t-4} is connected to only one open square within the pathway.

$N,\ t-l \leqslant t' \leqslant t\}$ are called *internal nodes*. An internal or output node x_i^t is connected to a set of circle nodes $\boldsymbol{x}_{\boldsymbol{I}_i}^{t-1}$ via a square node, which is associated with the function $f_i(\boldsymbol{x}_{\boldsymbol{I}_i}^{t-1})$. If the function is a single-variable function (i.e., $|\boldsymbol{I}_i| = 1$), it is represented by an open square (□); otherwise (i.e., $|\boldsymbol{I}_i| > 1$), it is represented by a solid square (■).

We associate the factor $\phi_i^t D[f_i(\boldsymbol{x}_{\boldsymbol{I}_i}^{t-1}), x_i^t]$ with a subnetwork called a *bond*, which comprises a square, the edges incident on the square, and the circles neighboring the square [Fig. 4(a)]. Moreover, we define the union of bonds (that is not empty) as a *pathway*, $\mathcal{P}_\alpha$. Examples of pathways are depicted in Figs. 4 and 5.

Each pathway has a pathway weight. We use $\mathcal{P}_\alpha$ both for a pathway and for its weight. It is given by

$$\begin{aligned}\mathcal{P}_\alpha &= \frac{1}{2^{J_\alpha}} \sum_{\boldsymbol{x}_\alpha} \prod_{\text{bond}\in\alpha} \phi_i^t\, D\big[f_i\big(\boldsymbol{x}_{\boldsymbol{I}_i}^{t-1}\big), x_i^t\big] \\ &= \frac{1}{2^{J_\alpha}} \sum_{\boldsymbol{x}_\alpha} \prod_{\text{bond}\in\alpha} \phi_i^t\, \sigma\big(f_i\big(\boldsymbol{x}_{\boldsymbol{I}_i}^{t-1}\big)\big)\, \sigma\big(x_i^t\big), \end{aligned} \tag{8}$$

where $\sigma(x) \equiv D[x, 1]$ and $D[x, y] = \sigma(x)\sigma(y)$. $\prod_{\text{bond}\in\alpha}$ represents the product of all the bond factors ϕD's in pathway $\mathcal{P}_\alpha$. $\boldsymbol{x}_\alpha$ denotes a set of internal nodes x_i^t within pathway $\mathcal{P}_\alpha$, and J_α denotes the number of internal nodes. For example, the pathway weight for pathway $\mathcal{P}_t^{(1)}$ in Fig. 4(a) is given by $\mathcal{P}_t^{(1)} = \frac{1}{2^2} \sum_{x_1^{t-1}=0,1} \sum_{x_2^{t-1}=0,1} \phi_N^t D[f_N(x_1^{t-1}, x_2^{t-1}), x_N^t]$.

Using the pathways, the joint probability distribution required for the TE calculation can be expressed as follows (see Sec. A1 of the SM [43]):

$$P\big(x_N^t, \boldsymbol{x}_N^-, \boldsymbol{s}^-\big) = \frac{1}{2^{l+1}} P(\boldsymbol{s}^-) \left[1 + \sum_\alpha \mathcal{P}_\alpha + O(\phi^q)\right], \tag{9}$$

where q is a positive integer, which can be chosen according to the desired calculation precision. It is to be noted that signal nodes are treated separately from the other variables [squares (□) are not introduced between signal transitions]. This enables us to factor out the signal distribution in Eq. (9). Thus the effect of the signal statistics on TE is clarified as shown in Eq. (13).

In Eq. (9), most of the pathways contain $\sum_{x=0,1} \sigma(x) = 0$ due to the symmetry $P(x_i) = P(\overline{x_i})$, and therefore they are nonexistent. The existent pathways satisfy the following graphical conditions.

(C1): In the pathway, circle nodes that do not have nodes on their right side must be one of the output nodes.

(C2): For every internal node in the pathway, if all the squares directly connected to the node and located within the pathway are open, then the number of open squares must be even.

Condition C1 implies that an output node can be reached by starting from any internal node of the pathway and moving forward in time. In other words, an internal node can influence an output node only through a chain of causal interactions. Condition C2 originates from a property of a single-variable Boolean function: $\sigma(f(x))$ is equal to either $+\sigma(x)$ or $-\sigma(x)$. A pathway that does not satisfy condition C2 includes an internal node $x_i^{t'}$ that is surrounded only by an odd number [$2m+1$ ($m \in \mathbb{Z}_{\geqslant 0}$)] of open squares located in the pathway. The pathway contains $\sum_{x_i^{t'}=0,1} \sigma(x_i^{t'})^{2m+1} = \sum_{x_i^{t'}=0,1} \sigma(x_i^{t'}) = 0$. This symmetric property causes the contribution of the pathway to the joint probability distribution, Eq. (9), to vanish. Pathways

that do not satisfy condition C2 are shown in Figs. 4(d) and 5(d) as examples. The proofs of conditions C1 and C2 are presented in Sec. A1 of the SM [43].

Substituting Eq. (9) in the definition of entropy, the two types of conditional entropy are expanded in terms of pathways as follows:

$$H\left[x_N^t \middle| \boldsymbol{x}_N^-, \boldsymbol{s}^-\right] = H\left[x_N^t, \boldsymbol{x}_N^-, \boldsymbol{s}^-\right] - H[\boldsymbol{x}_N^-, \boldsymbol{s}^-] = \ln 2 - \sum_{k=2}^{\infty} \frac{(-1)^k}{k(k-1)} \sum_{\alpha_1} \cdots \sum_{\alpha_k} E\left[\mathcal{P}_{\alpha_1} \cdots \mathcal{P}_{\alpha_k}\right] \quad (10)$$

and

$$H\left[x_N^t \middle| \boldsymbol{x}_N^-\right] = H\left[x_N^t, \boldsymbol{x}_N^-\right] - H[\boldsymbol{x}_N^-] = \ln 2 - \sum_{k=2}^{\infty} \frac{(-1)^k}{k(k-1)} \sum_{\alpha_1} \cdots \sum_{\alpha_k} E\left[\mathcal{P}_{\alpha_1}\right] \cdots E\left[\mathcal{P}_{\alpha_k}\right], \quad (11)$$

where $E[\cdots]$ represents the signal average defined as

$$E[X(\boldsymbol{s}^-)] = \sum_{\boldsymbol{s}^-} X(\boldsymbol{s}^-) P(\boldsymbol{s}^-). \quad (12)$$

For both conditional entropies, the set of pathways $\{\mathcal{P}_{\alpha_1}, \ldots, \mathcal{P}_{\alpha_k}\}$ appearing in each term satisfies the following condition.

(C3): $m(t', \mathcal{P}_\alpha)$ denotes the number of squares directly connected to an output node $x_N^{t'}$ within pathway $\mathcal{P}_\alpha$. For any t', the sum $\sum_{i=1}^{k} m(t', \mathcal{P}_{\alpha_i})$ must be even. In particular, for $t' = t$, the sum must be greater than or equal to 2.

For example, in Fig. 5, $m(t, \mathcal{P}_t^{(2)}) = 1$, $m(t, \mathcal{P}_t^{(3)}) = 1$, $m(t-3, \mathcal{P}_t^{(3)}) = 1$, $m(t, \mathcal{P}_t^{(5)}) = 1$, and $m(t-3, \mathcal{P}_t^{(5)}) = 2$. $\{\mathcal{P}_t^{(2)}, \mathcal{P}_t^{(5)}\}$ satisfies condition C3, whereas $\{\mathcal{P}_t^{(3)}, \mathcal{P}_t^{(5)}\}$ does not; $m(t-3, \mathcal{P}_t^{(3)}) + m(t-3, \mathcal{P}_t^{(5)}) = 3$.

Moreover, the pathway product $E[\mathcal{P}_{\alpha_1} \cdots \mathcal{P}_{\alpha_k}]$ appearing in $H[x_N^t | \boldsymbol{x}_N^-, \boldsymbol{s}^-]$ cancels out with other terms in Eq. (10), unless the set $\{\mathcal{P}_{\alpha_1}, \ldots, \mathcal{P}_{\alpha_k}\}$ satisfies a connectivity condition C4-i. In a similar manner, the pathway product $E[\mathcal{P}_{\alpha_1}] \cdots E[\mathcal{P}_{\alpha_k}]$ appearing in $H[x_N^t | \boldsymbol{x}_N^-]$ cancels out with terms in Eq. (11), unless the set $\{\mathcal{P}_{\alpha_1}, \ldots, \mathcal{P}_{\alpha_k}\}$ satisfies another connectivity condition C4-ii.

(C4-i): $\mathcal{P}_{\alpha_1} \cup \cdots \cup \mathcal{P}_{\alpha_k}$ is connected, and $\mathcal{P}_{\alpha_1} \cup \cdots \cup \mathcal{P}_{\alpha_k} \setminus \mathcal{P}_c$ is connected.

(C4-ii): At least one of the following two conditions must be satisfied: (1) $\mathcal{P}_{\alpha_1} \cup \cdots \cup \mathcal{P}_{\alpha_k}$ is connected, and $\mathcal{P}_{\alpha_1} \cup \cdots \cup \mathcal{P}_{\alpha_k} \setminus \mathcal{P}_c$ is connected. (2) $\mathcal{P}_{\alpha_1} \cup \cdots \cup \mathcal{P}_{\alpha_k} \cup \mathcal{P}_{\boldsymbol{s}^-}$ is connected, and $\mathcal{P}_{\alpha_1} \cup \cdots \cup \mathcal{P}_{\alpha_k} \cup \mathcal{P}_{\boldsymbol{s}^-} \setminus \mathcal{P}_c$ is connected.

Here, the symbol $\cup$ denotes the union of pathways in the temporal graph, $\mathcal{P}_c$ corresponds to any of the k pathways, $\setminus \mathcal{P}_c$ represents the subtraction of $\mathcal{P}_c$ from a union, and $\mathcal{P}_{\boldsymbol{s}^-}$ corresponds to a subgraph consisting of signal nodes $\boldsymbol{s}^-$ and the edges between them in the temporal graph. In Fig. 6, an example of a pathway product that satisfies neither condition C4-i nor condition C4-ii is illustrated, where $\mathcal{P}_{\boldsymbol{s}^-}$ is indicated by a blue rectangle.

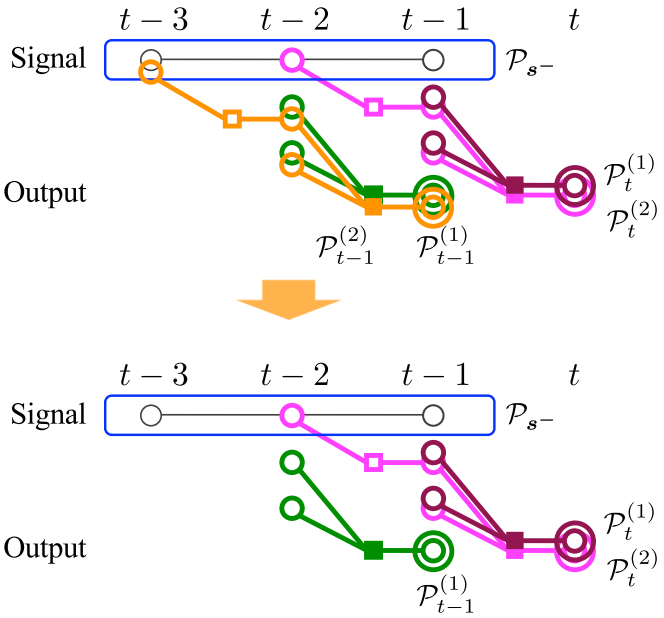

FIG. 6. Nonexistent pathway product $\mathcal{P}_t^{(1)}\mathcal{P}_t^{(2)}\mathcal{P}_{t-1}^{(1)}\mathcal{P}_{t-1}^{(2)}$ in the FFL. Because the union of pathways $\mathcal{P}_t^{(1)} \cup \mathcal{P}_t^{(2)} \cup \mathcal{P}_{t-1}^{(1)} \cup \mathcal{P}_{t-1}^{(2)}$ is disconnected, condition C4-i is violated. Thus $E[\mathcal{P}_t^{(1)}\mathcal{P}_t^{(2)}\mathcal{P}_{t-1}^{(1)}\mathcal{P}_{t-1}^{(2)}]$ is not required for calculating $H[x_N^t | \boldsymbol{x}_N^-, \boldsymbol{s}^-]$. Adding $\mathcal{P}_{\boldsymbol{s}^-}$ to it, the union $\mathcal{P}_t^{(1)} \cup \mathcal{P}_t^{(2)} \cup \mathcal{P}_{t-1}^{(1)} \cup \mathcal{P}_{t-1}^{(2)} \cup \mathcal{P}_{\boldsymbol{s}^-}$ is connected. However, if $\mathcal{P}_{t-1}^{(2)}$ is removed, the union $\mathcal{P}_t^{(1)} \cup \mathcal{P}_t^{(2)} \cup \mathcal{P}_{t-1}^{(1)} \cup \mathcal{P}_{\boldsymbol{s}^-}$ is disconnected, i.e., condition C4-ii is violated. Therefore $E[\mathcal{P}_t^{(1)}]E[\mathcal{P}_t^{(2)}]E[\mathcal{P}_{t-1}^{(1)}]E[\mathcal{P}_{t-1}^{(2)}]$ is not required for calculating $H[x_N^t | \boldsymbol{x}_N^-]$.

Taking the difference between the conditional entropies in Eqs. (10) and (11), we finally obtain the general TE formula:

$$T_{s \to N} = \sum_{k=2}^{\infty} \frac{(-1)^k}{k(k-1)} \sum_{\alpha_1} \cdots \sum_{\alpha_k} \left(E\left[\mathcal{P}_{\alpha_1} \cdots \mathcal{P}_{\alpha_k}\right] - E\left[\mathcal{P}_{\alpha_1}\right] \cdots E\left[\mathcal{P}_{\alpha_k}\right]\right), \quad (13)$$

where pathway products satisfying condition C4-i and those satisfying condition C4-ii contribute to $E[\mathcal{P}_{\alpha_1} \cdots \mathcal{P}_{\alpha_k}]$ and $E[\mathcal{P}_{\alpha_1}] \cdots E[\mathcal{P}_{\alpha_k}]$, respectively. The factors that characterize the system appear implicitly in this formula. The signal statistics are employed for averaging the pathways and pathway products, and the Boolean functions and error rates are included in the pathways. Note that Eq. (13) is not restricted to systems with a Markov information source.

In Eq. (13), for the set $\{\mathcal{P}_{\alpha_1}, \ldots, \mathcal{P}_{\alpha_k}\}$ that satisfies both condition C4-i and condition C4-ii, $E[\mathcal{P}_{\alpha_1} \cdots \mathcal{P}_{\alpha_k}]$ and $E[\mathcal{P}_{\alpha_1}] \cdots E[\mathcal{P}_{\alpha_k}]$ can be canceled out by each other in certain cases. If all the pathways in set $\{\mathcal{P}_{\alpha_1}, \ldots, \mathcal{P}_{\alpha_k}\}$ have no signal nodes, $E[\mathcal{P}_{\alpha_1} \cdots \mathcal{P}_{\alpha_k}] = E[\mathcal{P}_{\alpha_1}] \cdots E[\mathcal{P}_{\alpha_k}]$ holds. Moreover, if only one pathway in set $\{\mathcal{P}_{\alpha_1}, \ldots, \mathcal{P}_{\alpha_k}\}$ has signal nodes, the same equation holds. Therefore the set $\{\mathcal{P}_{\alpha_1}, \ldots, \mathcal{P}_{\alpha_k}\}$ that actually contributes to TE satisfies the following condition.

(C5): At least two pathways in k pathways $\{\mathcal{P}_{\alpha_1}, \ldots, \mathcal{P}_{\alpha_k}\}$ must have signal nodes.

Conditions C3, C4-i, C4-ii, and C5 have intuitive interpretations. Condition C3 indicates that any existent pathway product has the output node x_N^t. Conditions C4-i and C4-ii

indicate that pathways in any existent pathway product are connected. Condition C5 indicates that any existent pathway product contains signal nodes. Therefore their combination indicates that only the pathway products that bridge the input and output achieve information transfer. Thus the essential structures for information transfer are extracted by the graphical conditions. The proofs of conditions C3, C4-i, C4-ii, and C5 are presented in Secs. A2, A3, and A4, respectively, of the SM [43].

In practical applications, even a pathway product that satisfies all the graphical conditions does not need to be considered if its order of ϕ is high. The reason is that the contribution of a high-order pathway product to TE is less. Such products are those containing (i) many pathways and (ii) long-past nodes. The contribution from a pathway with long-past nodes decays on the signal timescale Γ. Thus pathway products with long-past nodes can be neglected. This implies that TE is independent of l for sufficiently large l in the definition Eq. (5).

The graphical restrictions [conditions (C1)–(C5)] that remove numerous irrelevant pathways and their products enable efficient computation of TE. Applying Eq. (13) to a concrete system, TE is further decomposed into its individual elements as illustrated in Fig. 2.

V. APPLICATION TO NETWORK MOTIFS

We apply the general formula in Eq. (13) to the motifs. One may suspect that Boolean functions would need to be specified before starting the computation of TE because the pathway weights appearing in Eq. (13) depend on the choice of the Boolean functions. However, Fourier transformation of the Boolean functions [48,49] makes it possible to defer the substitution of a specific set of Boolean functions into the pathway weight Eq. (8); thus it provides the TE formula for a certain network with arbitrary Boolean functions. In this section, we first introduce the Fourier transformation and then present the TE formulas for the FFL and FBL. Finally, we discuss how the length difference of the respective routes from the signal source to the output vertex in a motif generates the signal-timescale dependence of TE.

A. Fourier transformation of Boolean functions

An n-variable function $\sigma(f(\boldsymbol{x}))$ can be expressed as the following Fourier series [48,49]:

$$\sigma(f(\boldsymbol{x})) = \sum_{\boldsymbol{k}} \hat{f}_{(\boldsymbol{k})} \left\{ \prod_{j\in\boldsymbol{k}} \sigma(x_j) \right\}, \tag{14}$$

$$\hat{f}_{(\boldsymbol{k})} = \frac{1}{2^n} \sum_{\boldsymbol{x}} \sigma(f(\boldsymbol{x})) \left\{ \prod_{j\in\boldsymbol{k}} \sigma(x_j) \right\}, \tag{15}$$

where $\boldsymbol{k}$ is a subset of input variables $\boldsymbol{x}$ and $\hat{f}_{\boldsymbol{k}}$ is the Fourier coefficient of f at $\boldsymbol{k}$. We illustrate the Fourier transformation for one- and two-variable functions. For $f(x_i)$ (namely, $n = 1$), $\boldsymbol{k}$ is either empty $\emptyset$ or x_i:

$$\sigma(f(x_i)) = \hat{f}_{(\emptyset)} + \hat{f}_{(i)}\sigma(x_i),$$
$$\hat{f}_{(\emptyset)} = \frac{1}{2} \sum_{x_i=0,1} \sigma(f(x_i)),$$
$$\hat{f}_{(i)} = \frac{1}{2} \sum_{x_i=0,1} \sigma(f(x_i))\,\sigma(x_i).$$

Because $\hat{f}_{(\emptyset)} = \frac{1}{2}[\sigma(f(0)) + \sigma(f(1))] = 0$, $\sigma(f(x_i)) = \hat{f}_{(i)}\sigma(x_i)$. When f is the identity (negation) function, $\hat{f}_{(i)} = +1$ (-1).

For $n = 2$, $\boldsymbol{k} \in \{\emptyset, x_j, x_k, (x_j, x_k)\}$, and we denote the Fourier series as follows:

$$\sigma(f(x_j, x_k)) = \hat{f}_{(\emptyset)} + \hat{f}_{(j)}\sigma(x_j) + \hat{f}_{(k)}\sigma(x_k) + \hat{f}_{(j,k)}\sigma(x_j)\,\sigma(x_k),$$
$$\hat{f}_{(\emptyset)} = \frac{1}{2^2} \sum_{x_j=0,1} \sum_{x_k=0,1} \sigma(f(x_j, x_k)),$$
$$\hat{f}_{(j)} = \frac{1}{2^2} \sum_{x_j=0,1} \sum_{x_k=0,1} \sigma(f(x_j, x_k))\,\sigma(x_j),$$
$$\hat{f}_{(k)} = \frac{1}{2^2} \sum_{x_j=0,1} \sum_{x_k=0,1} \sigma(f(x_j, x_k))\,\sigma(x_k),$$
$$\hat{f}_{(j,k)} = \frac{1}{2^2} \sum_{x_j=0,1} \sum_{x_k=0,1} \sigma(f(x_j, x_k))\,\sigma(x_j)\,\sigma(x_k).$$

Each Fourier coefficient measures the correlation between the output variable of the function and the respective input variable or a combination. For example, $\hat{f}_{(j)}$ measures the correlation between $\sigma(f(x_j, x_k))$ and $\sigma(x_j)$. When $f(x_j, x_k)$ is the AND function, $\hat{f}_{(\emptyset)} = -\frac{1}{2}$ and $\hat{f}_{(j)} = \hat{f}_{(k)} = \hat{f}_{(j,k)} = +\frac{1}{2}$. When it is the OR function, $\hat{f}_{(\emptyset)} = \hat{f}_{(j)} = \hat{f}_{(k)} = +\frac{1}{2}$ and $\hat{f}_{(j,k)} = -\frac{1}{2}$. When it is the XOR function, $\hat{f}_{(\emptyset)} = \hat{f}_{(j)} = \hat{f}_{(k)} = 0$ and $\hat{f}_{(j,k)} = -1$, i.e., the XOR function correlates only with a combination of two input variables.

This transformation simplifies a polynomial expression of a pathway weight, Eq. (8), by performing the summation over internal nodes $\boldsymbol{x}_\alpha$. As shown later, for any pathway, the Fourier components of the Boolean functions contributing to the pathway weight are those projected along the pathway.

B. TE of the FFL and FBL

The steps for calculating the TE of the motifs are summarized below: (i) The temporal graph is drawn for a motif (Figs. 4 and 5). (ii) The existent pathways satisfying conditions C1 and C2 are listed. (iii) The pathway products satisfying conditions C3–C5 are constructed up to a desired order of ϕ_i. (iv) The pathway weights appearing in the pathway products constructed in step (iii) are simplified using Fourier transformation. (v) The constructed pathway products are averaged over the signal distribution according to Eq. (13).

In Secs. B1 and B2 of the SM [43], we indicate the existent pathways required for calculating TE of the FFL and FBL, respectively. As an example of a Fourier-transformed pathway

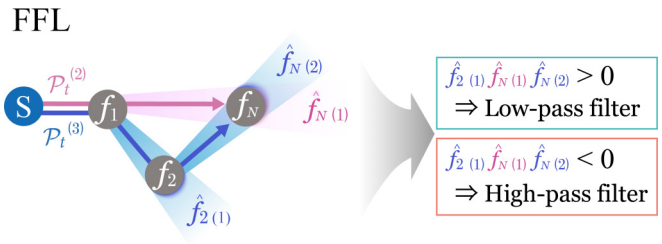

FIG. 7. Schematic of the origin of the TE signal-timescale dependence in the FFL. Here, for example, $\hat{f}_{N(1)}$ is represented as the "shadow" of f_N cast by the "incident light" $1 \to N$. The filtering property of the motif is determined by the combination of the projected Boolean functions along pathways $\mathcal{P}_t^{(2)}$ and $\mathcal{P}_t^{(3)}$.

weight, we present $\mathcal{P}_t^{(3)}$ for the FFL drawn in Fig. 4(c):

$$\mathcal{P}_t^{(3)} = \hat{f}_{1(s)}\hat{f}_{2(1)}\hat{f}_{N(2)}\phi_1\phi_2\phi_N\sigma(s^{t-3})\sigma\left(x_N^t\right), \tag{16}$$

where the Boolean functions are not yet specified. Although this pathway contains $f_N(x_1, x_2)$, only the Fourier component of the direction along the pathway, $\hat{f}_{N(2)}$, contributes to the pathway weight.

The pathway products required for the FFL and FBL are also shown in Secs. B1 and B2, respectively, of the SM [43]. The formulas for the FFL and FBL up to the order of ϕ^7 comprise ten and five pathway products, respectively. These formulas precisely reproduce the numerically obtained TE values, as shown in Figs. 1(c) and 1(d), where $\psi_i = 0.3$ $(i = 1, 2, N)$ are assumed. As the error rates are increased $(\psi_i \to \frac{1}{2})$, the accuracy of our perturbative method improves. We point out that all the pathways used for obtaining the formulas contain no nodes $s^{t'}$ or $x_i^{t'}$ $(t' < t - 5)$. This indicates that, up to $O(\phi^7)$, the TE of the motifs is independent of l when $l \geqslant 5$.

The first and second leading terms in the expansion series for the FFL are sufficient to identify the difference between the PFFL and NFFL. Assuming $\phi_i = \phi$ for simplicity, the TE of the FFL is given by

$$T_{s\to N}^{\mathrm{FFL}} = \tfrac{1}{2}\hat{f}_{N(1)}^{\,2}(1 - M^2)\phi^4 + \hat{f}_{2(1)}\hat{f}_{N(1)}\hat{f}_{N(2)}\,(G_1 - M^2)\phi^5 + O(\phi^6), \tag{17}$$

where G_L is the autocorrelation function of the signal defined as

$$G_L = E[\sigma(s^{t-L})\sigma(s^t)]. \tag{18}$$

The correlation G_1 can be simply described in terms of the harmonic mean Γ:

$$G_1 = 1 - 2\Gamma, \tag{19}$$

i.e., G_1 increases with the increase in the average signal timescale $1/\Gamma$. Thus we find that the timescale dependence of TE is caused by the presence of G_1 in Eq. (17).

Equation (17) further indicates that, for $M \simeq 0$, an increase in G_1, i.e., an increase in the signal timescale, leads to an increase in $T_{s\to N}^{\mathrm{FFL}}$, if a motif satisfies

$$\hat{f}_{2(1)}\hat{f}_{N(1)}\hat{f}_{N(2)} > 0. \tag{20}$$

Note that the left-hand side consists of the Fourier components along the pathways (Fig. 7). In contrast, a smaller G_1, i.e., a shorter signal timescale, leads to an increase in $T_{s\to N}^{\mathrm{FFL}}$ for any motif that satisfies $\hat{f}_{2(1)}\hat{f}_{N(1)}\hat{f}_{N(2)} < 0$. Hence the emergence of either the low- or high-pass filtering property for a time-varying signal is determined only by the sign of $\hat{f}_{2(1)}\hat{f}_{N(1)}\hat{f}_{N(2)}$. This reveals that such filtering properties are robust with respect to various modifications, such as the replacement of the AND function by the OR function, position change of an inhibition interaction or negative link from $1 \to 2$ to $2 \to N$, and shuffling of noise parameters ψ_1, ψ_2, and ψ_N. Thus the inequalities in Eq. (20) and $\hat{f}_{2(1)}\hat{f}_{N(1)}\hat{f}_{N(2)} < 0$ can be employed for defining the PFFL and NFFL, respectively.

The leading term of Eq. (17) includes a square of the bias M. Therefore $T_{s\to N}^{\mathrm{FFL}}$ is maximized when $M \simeq 0$, independent of the signal timescale, as observed in Fig. 1(d).

The TE formula for the FBL up to ϕ^7 is presented in Sec. B2 of the SM [43]. It turns out that the FBL has the same leading term as the FFL. Therefore TE in the FBL is maximized when $M \simeq 0$ [Fig. 1(d)]. Here, we show the TE of the FBL with an unbiased signal:

$$T_{s\to N}^{\mathrm{FBL}}(M = 0) = \tfrac{1}{2}\hat{f}_{1(s)}^{\,2}\phi^4 + \tfrac{1}{2}\hat{f}_{1(s,2)}^{\,2}\phi^6 - \hat{f}_{N(1)}\hat{f}_{2(N)}\hat{f}_{1(\emptyset)}^{\,2}\hat{f}_{1(2)}\phi^7 + O(\phi^8). \tag{21}$$

We can see that Eq. (21) is independent of the signal timescale. For $M \neq 0$, TE depends on the signal timescale Γ. However, this dependence is weak because it first appears in the order of ϕ^7.

Noting that $\hat{f}_{2(N)} = \frac{1}{2}$ and $\hat{f}_{2(N)} = -\frac{1}{2}$ for the PFBL and NFBL that we employed, respectively, the difference in TE between the PFBL and NFBL [Figs. 1(c) and 1(d)] is due to the ϕ^7 term in Eq. (21). More specifically, this term originates from $H[x_N^t|\boldsymbol{x}_N^-]$, as shown in Sec. B2 of the SM [43]. Thus our analytical method verifies the intuition mentioned in Sec. III that the TE difference among FBLs is generated by the difference in the ambiguities of the output variables.

As shown above, the TE of the FFL depends on the signal timescale due to the existence of G_1 in the second leading term, whereas that of the FBL hardly depends on the signal timescale because G_L is absent in both the first and second leading terms. In the following section, we discuss the network structure that causes G_L to appear in the low-order terms of TE.

C. Signal-timescale independence in a biparallel motif

For FFL motifs, the TE dependence on the signal timescale is caused by G_1. This originates from the pathway product $\mathcal{P}_t^{(2)}\mathcal{P}_t^{(3)}$, as indicated by the derivation process for Eq. (17). The pathway product corresponds to the interaction between the shortest route $s \to 1 \to N$ and second shortest route $s \to 1 \to 2 \to N$ in the regulation network (Fig. 7). In contrast, for FBL motifs, which do not exhibit signal-timescale dependence, only route $s \to 1 \to N$ connects the signal source to the output vertex. If a regulation network has only multiple information routes of the same length, does TE depend on the signal timescale? To answer this question, we introduce a motif called "biparallel," which only has routes of the same length, as shown in Fig. 8(a).

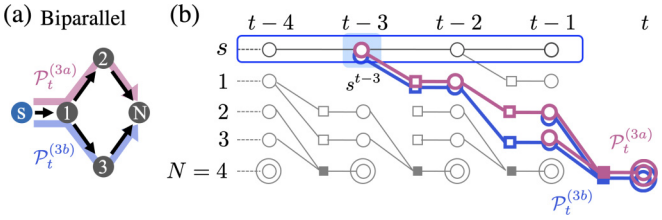

FIG. 8. (a) Biparallel network motif. Two existent pathways $\mathcal{P}_t^{(3a)}$ and $\mathcal{P}_t^{(3b)}$ are drawn, which are defined in (b). (b) Existent pathways $\mathcal{P}_t^{(3a)}$ and $\mathcal{P}_t^{(3b)}$ for the biparallel network motif, where superscripts a and b are introduced to distinguish two pathways of the same order of ϕ^3. We emphasize that they share an identical signal node s^{t-3}. Consequently, the pathway product $\mathcal{P}_t^{(3a)}\mathcal{P}_t^{(3b)}$ does not provide G_L. Therefore TE does not depend on the signal timescale.

In the biparallel motif, TE hardly shows timescale dependence (the derivation is presented in Sec. B4 of the SM [43]):

$$T_{s\to N}^{\text{biparallel}} = \tfrac{1}{2}\left[\hat{f}_{N(2)}^{\,2} + \hat{f}_{N(3)}^{\,2} + 2\hat{f}_{2(1)}\hat{f}_{3(1)}\hat{f}_{N(2)}\hat{f}_{N(3)}\right] \times(1-M^2)\phi^6 + O(\phi^8). \tag{22}$$

This is because the two pathways, corresponding to the two shortest routes from the signal source to the output vertex, share the same signal node [Fig. 8(b)] and, consequently, their interaction does not provide G_L. Conversely, for the signal-timescale dependence of TE, i.e., the filtering property for information flow, at least two routes with different lengths should exist in a regulation network.

VI. DISCUSSION AND CONCLUSION

Utilizing the abstraction of a Boolean network model, we revealed the information-transfer characteristics of an individual motif, which is the most basic functional unit in a complex network, including biological, technological, and social systems. The PFFL and NFFL have low- and high-pass filtering properties, respectively, for fluctuating signals with various timescales. Such properties originate from the interaction between information pathways with different lengths. The emergence of a low- or high-pass filter is solely determined by the positive or negative sign of the product of the Fourier components along the pathways. This indicates that emergence is possible even with various structural modifications. In contrast, the PFBL and NFBL can transmit information stably, independently of the signal timescale. This is because they do not have multiple information routes with different lengths.

Owing to the systematic nature of the TE formula, the aforementioned information-transfer characteristics in each motif were clarified to be independent of the details of its regulation rules and determined only by topological aspects of pathway combinations. In other words, the informational function indicated by a motif originates from the connection pattern of the motif, that is, what defines the motif or the defining property of the motif. Discovering such general properties solely by using numerical calculations and experiments would be difficult.

Our graphical expansion method avoids all the obstacles for the analytical computation of TE. Expansion around the uniform probability distribution with symmetrical property $P(x_i) = P(\bar{x}_i)$ realizes significant reduction in the computational complexity. Consequently, the few pathway products that satisfy all the graphical restriction conditions express the essential features of the interactions that give rise to information flows. This reduction is also found in the literature [41]. There are two aspects to this study that we wish to highlight. First, the special treatment of the signal nodes in constructing the joint probability distribution Eq. (9) enables the postponement of the consideration of signals to the final operation, that is, averaging the pathways, for obtaining TE. This technique successfully separates TE into the parts derived from the internal structure and those derived from the signal characteristics. Second, Fourier transformation of the Boolean functions enables TE calculation without requiring the determination of specific functional forms for the regulations. Consequently, a general condition on the regulatory relations for the emergence of the filtering property is derived for a motif. Because our method can be extended to other entropic quantities, such as mutual information, dissipation rate, and learning rate, it would be employed to perturbatively examine relations between information-theoretic measures [50–59].

Our theory can provide design principles for regulation networks that transfer information for different purposes. Although only one signal source was considered in this paper, biological systems often receive signals from various sources simultaneously and process them. For example, the neural network of *Caenorhabditis elegans* can integrate multiple sensory inputs, such as O_2, CO_2, and the temperature [60]. ERBB signal transduction can respond to a variety of ligand molecules [61]. To understand such information processing, it is desirable to include multiple information sources in our formulation.

Our methods can be applied to Boolean networks with a general topology. Although we focused on motifs in this paper, our methods can be beneficial for understanding the effect of other topological properties, such as the redundancy, centrality, hierarchy, clustering, and degree distributions [62], on the information-transfer characteristics. Moreover, application of the methods to real biological networks can enable the identification of the dominantly contributing pairs or combinations of pathways for transmitting information. The identification of such pairs or combinations will contribute to bridging the gap between observable microscopic processes, such as firing of neurons or expression of genes, and the macroscopic emergence of biological functions.

ACKNOWLEDGMENTS

We thank Hiroshi Ito for providing helpful comments. This work was supported by JSPS KAKENHI Grants No. JP19K03663, No. JP22K03453, and No. JP22K06347 and the RIKEN iTHEMS program.

[1] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, Network motifs: Simple building blocks of complex networks, Science **298**, 824 (2002).
[2] M. Reigl, U. Alon, and D. B. Chklovskii, Search for computational modules in the *C. elegans* brain, BMC Biol. **2**, 25 (2004).
[3] U. Alon, Network motifs: Theory and experimental approaches, Nat. Rev. Genet. **8**, 450 (2007).
[4] M. Adler and R. Medzhitov, Emergence of dynamic properties in network hypermotifs, Proc. Natl. Acad. Sci. USA **119**, e2204967119 (2022).
[5] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (CRC, Boca Raton, FL, 2019).
[6] S. Chakravarty and A. Csikász-Nagy, Systematic analysis of noise reduction properties of coupled and isolated feedforward loops, PLoS Comput. Biol. **17**, e1009622 (2021).
[7] S. Mangan and U. Alon, Structure and function of the feed-forward loop network motif, Proc. Natl. Acad. Sci. USA **100**, 11980 (2003).
[8] S. Mangan, A. Zaslaver, and U. Alon, The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks, J. Mol. Biol. **334**, 197 (2003).
[9] H. Bolouri and E. H. Davidson, Modeling transcriptional regulatory networks, BioEssays **24**, 1118 (2002).
[10] E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. Van Oudenaarden, Multistability in the lactose utilization network of *Escherichia coli*, Nature (London) **427**, 737 (2004).
[11] J. Blau and M. W. Young, Cycling *vrille* expression is required for a functional *Drosophila* clock, Cell **99**, 661 (1999).
[12] M. B. Elowitz and S. Leibler, A synthetic oscillatory network of transcriptional regulators, Nature (London) **403**, 335 (2000).
[13] R. Avraham and Y. Yarden, Feedback regulation of EGFR signalling: Decision making by early and delayed loops, Nat. Rev. Mol. Cell Biol. **12**, 104 (2011).
[14] J. J. Tyson and B. Novák, Functional motifs in biochemical reaction networks, Annu. Rev. Phys. Chem. **61**, 219 (2010).
[15] W. Li, S. Krishna, S. Pigolotti, N. Mitarai, and M. H. Jensen, Switching between oscillations and homeostasis in competing negative and positive feedback motifs, J. Theor. Biol. **307**, 205 (2012).
[16] R. Albert, B. R. Acharya, B. W. Jeon, J. G. Zañudo, M. Zhu, K. Osman, and S. M. Assmann, A new discrete dynamic model of ABA-induced stomatal closure predicts key feedback loops, PLoS Biol. **15**, e2003451 (2017).
[17] L. Stone, D. Simberloff, and Y. Artzy-Randrup, Network motifs and their origins, PLoS Comput. Biol. **15**, e1006749 (2019).
[18] A. K. Dey, Y. R. Gel, and H. V. Poor, What network motifs tell us about resilience and reliability of complex networks, Proc. Natl. Acad. Sci. USA **116**, 19368 (2019).
[19] D. S. Glass, X. Jin, and I. H. Riedel-Kruse, Nonlinear delay differential equations and their application to modeling biological network motifs, Nat. Commun. **12**, 1788 (2021).
[20] W. Zheng, C. Gu, H. Yang, and J. H. T. Rohling, Motif structure for the four subgroups within the suprachiasmatic nuclei affects its entrainment ability, Phys. Rev. E **105**, 014314 (2022).
[21] S. A. Kauffman, Metabolic stability and epigenesis in randomly constructed genetic nets, J. Theor. Biol. **22**, 437 (1969).
[22] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, The information dynamics of phase transitions in random Boolean networks, in *Artificial Life XI: Proceedings of the 11th International Conference on the Simulation and Synthesis of Living Systems* (MIT Press, Cambridge, MA, 2008), pp. 374–381.
[23] R. Heckel, S. Schober, and M. Bossert, Harmonic analysis of Boolean networks: Determinative power and perturbations, EURASIP J. Bioinf. Syst. Biol. **2013**, 6 (2013).
[24] J. G. Klotz, D. Kracht, M. Bossert, and S. Schober, Canalizing Boolean functions maximize mutual information, IEEE Trans. Inf. Theor. **60**, 2139 (2014).
[25] M. T. Matache and V. Matache, Logical reduction of biological networks to their most determinative components, Bull. Math. Biol. **78**, 1520 (2016).
[26] T. Haruna and K. Nakajima, Maximizing local information transfer in Boolean networks, New J. Phys. **20**, 083046 (2018).
[27] F. M. Weidner, J. D. Schwab, S. D. Werle, N. Ikonomi, L. Lausser, and H. A. Kestler, Capturing dynamic relevance in Boolean networks using graph theoretical measures, Bioinformatics **37**, 3530 (2021).
[28] T. Pentzien, B. L. Puniya, T. Helikar, and M. T. Matache, Identification of biologically essential nodes via determinative power in logical models of cellular processes, Front. Physiol. **9**, 1185 (2018).
[29] M. Porfiri and M. R. Marín, Information flow in a model of policy diffusion: An analytical study, IEEE Trans. Network Sci. Eng. **5**, 42 (2017).
[30] R. H Goodman and M. Porfiri, Topological features determining the error in the inference of networks using transfer entropy, Math. Eng. **2**, 34 (2020).
[31] L. Novelli, F. M. Atay, J. Jost, and J. T. Lizier, Deriving pairwise transfer entropy from network structure and motifs, Proc. R. Soc. London A **476**, 20190779 (2020).
[32] S. Otsubo and T. Sagawa, Information-thermodynamic characterization of stochastic Boolean networks, arXiv:1803.04217.
[33] J. Kesseli, P. Rämö, and O. Yli-Harja, On spectral techniques in analysis of Boolean networks, Phys. D (Amsterdam) **206**, 49 (2005).
[34] J. Kesseli, P. Rämö, and O. Yli-Harja, Tracking perturbations in Boolean networks with spectral methods, Phys. Rev. E **72**, 026137 (2005).
[35] P. Rämö, S. Kauffman, J. Kesseli, and O. Yli-Harja, Measures for information propagation in Boolean networks, Phys. D (Amsterdam) **227**, 100 (2007).
[36] A. S. Ribeiro, S. A. Kauffman, J. Lloyd-Price, B. Samuelsson, and J. E. S. Socolar, Mutual information in random Boolean models of regulatory networks, Phys. Rev. E **77**, 011901 (2008).
[37] T. Parmer, L. M. Rocha, and F. Radicchi, Influence maximization in Boolean networks, Nat. Commun. **13**, 3457 (2022).
[38] E. Schneidman, M. J. Berry, R. Segev, and W. Bialek, Weak pairwise correlations imply strongly correlated network states in a neural population, Nature (London) **440**, 1007 (2006).
[39] S. Cocco, S. Leibler, and R. Monasson, Neuronal couplings between retinal ganglion cells inferred by efficient inverse statistical physics methods, Proc. Natl. Acad. Sci. USA **106**, 14058 (2009).
[40] Y. Chen, B. Q. Rosen, and T. J. Sejnowski, Dynamical differential covariance recovers directional network structure in

multiscale neural systems, Proc. Natl. Acad. Sci. USA **119**, e2117234119 (2022).

[41] F. Mori and T. Okada, Diagrammatic expansion of information flows in stochastic Boolean networks, Phys. Rev. Res. **2**, 043432 (2020).

[42] T. Schreiber, Measuring Information Transfer, Phys. Rev. Lett. **85**, 461 (2000).

[43] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.5.013037 for derivations and proofs.

[44] J. Runge, J. Heitzig, V. Petoukhov, and J. Kurths, Escaping the Curse of Dimensionality in Estimating Multivariate Transfer Entropy, Phys. Rev. Lett. **108**, 258701 (2012).

[45] S. Ito and T. Sagawa, Information Thermodynamics on Causal Networks, Phys. Rev. Lett. **111**, 180603 (2013).

[46] J. Runge, Quantifying information transfer and mediation along causal pathways in complex systems, Phys. Rev. E **92**, 062829 (2015).

[47] P. Jiang and P. Kumar, Interactions of information transfer along separable causal paths, Phys. Rev. E **97**, 042310 (2018).

[48] S. Jukna, *Boolean Function Complexity: Advances and Frontiers*, Algorithms and Combinatorics Vol. 27 (Springer, New York, 2012).

[49] R. O'Donnell, *Analysis of Boolean Functions* (Cambridge University Press, Cambridge, 2014).

[50] C. Jarzynski, Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale, Annu. Rev. Condens. Matter Phys. **2**, 329 (2011).

[51] A. C. Barato and U. Seifert, Unifying Three Perspectives on Information Processing in Stochastic Thermodynamics, Phys. Rev. Lett. **112**, 090601 (2014).

[52] J. M. Horowitz and M. Esposito, Thermodynamics with Continuous Information Flow, Phys. Rev. X **4**, 031015 (2014).

[53] J. M. Horowitz, Multipartite information flow for multiple Maxwell demons, J. Stat. Mech. (2015) P03006.

[54] D. Loutchko, M. Eisbach, and A. S. Mikhailov, Stochastic thermodynamics of a chemical nanomachine: The channeling enzyme tryptophan synthase, J. Chem. Phys. **146**, 025101 (2017).

[55] R. A. Brittain, N. S. Jones, and T. E. Ouldridge, What we learn from the learning rate, J. Stat. Mech. (2017) 063502.

[56] R. Chetrite, M. Rosinberg, T. Sagawa, and G. Tarjus, Information thermodynamics for interacting stochastic systems without bipartite structure, J. Stat. Mech. (2019) 114002.

[57] A. B. Boyd, A. Patra, C. Jarzynski, and J. P. Crutchfield, Shortcuts to thermodynamic computing: The cost of fast and faithful information processing, J. Stat. Phys. **187**, 17 (2022).

[58] H. Kiwata, Relationship between Schreiber's transfer entropy and Liang-Kleeman information flow from the perspective of stochastic thermodynamics, Phys. Rev. E **105**, 044130 (2022).

[59] S. Yoshida, Y. Okada, E. Muneyuki, and S. Ito, Thermodynamic role of main reaction pathway and multi-body information flow in membrane transport, Phys. Rev. Res. **4**, 023229 (2022).

[60] D. D. Ghosh, M. N. Nitabach, Y. Zhang, and G. Harris, Multisensory integration in *C. elegans*, Curr. Opin. Neurobiol. **43**, 110 (2017).

[61] Y. Yarden and M. X. Sliwkowski, Untangling the ErbB signalling network, Nat. Rev. Mol. Cell Biol. **2**, 127 (2001).

[62] F. Menczer, S. Fortunato, and C. A. Davis, *A First Course in Network Science* (Cambridge University Press, Cambridge, 2020).