

Hodge Decomposition of Information Flow on Complex Networks

Yuuya Fujiki

Graduate School of Science, Kobe University
1-1 Rokkodaicho, Nada
Kobe 657-8501, Japan
107s420s@stu.kobe-u.ac.jp

Taichi Haruna

Graduate School of Science, Kobe University
1-1 Rokkodaicho, Nada
Kobe 657-8501, Japan
tharuna@penguin.kobe-u.ac.jp

ABSTRACT

Decomposition of information flow associated with random threshold network dynamics on random networks with specified degree distributions is studied by numerical simulation. Combinatorial Hodge theory enables us to orthogonally decompose information flow into gradient (unidirectional acyclic flow), harmonic (global circular flow) and curl (local circular flow) components. We show that in-degree distributions have little influence on the relative strength of the circular component (harmonic plus curl) while out-degree distributions with longer tail suppress it. We discuss an implication of this finding on the topology of real-world gene regulatory networks.

Categories and Subject Descriptors

G.2.3 [Discrete Mathematics]: Applications; H.1.1 [Information Systems]: Systems and Information Theory—*general systems theory, information theory*

General Terms

Theory

Keywords

Combinatorial Hodge theory, complex networks, random threshold networks, transfer entropy

1. INTRODUCTION

Science of complex networks reveals that real-world networks found in nature and society have unique structural features [2, 10]. Recently, great attention has been paid to dynamics on complex networks because underlying network structures may have non-trivial influence on dynamics on them [5]. Complex dynamics can give rise to information processing ability to systems represented by networks [9]. We believe that information processing arising from complex dynamics is important to understand function of real-world complex systems such as living cells and brains.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

BICT 2014, December 01-03, Boston, United States

Copyright © 2015 ICST 978-1-63190-053-2

DOI 10.4108/icst.bict.2014.257876

The aim of this paper is to analyze components of information flow associated with dynamics on complex networks in order to reveal intrinsic information processing on them. We employ combinatorial Hodge theory [8] to decompose it. The combinatorial Hodge decomposition theorem decomposes any flow on a network into three components: gradient, harmonic and curl flows. Gradient flows are unidirectional acyclic, while harmonic and curl flows are circular. The difference between the latter two is in that harmonic flows are globally circular on the one hand, curl flows are locally circular on the other hand. As a first step for this direction of study, here we numerically investigate influence of a degree distribution on components of information flow associated with random threshold network dynamics.

Random threshold networks (RTNs) are a simple mathematical model of gene regulatory networks or neural networks [14]. Since their basic properties are well-studied in previous work [13], they are a useful testbed to study information flow in complex systems consisting of many elements with nonlinear interactions. For the quantification of information flow, here we adopt the transfer entropy [16] which has been used to detect causal relationships between two elements in complex systems found in wide range of disciplines [7].

In the literature, information flow on networks has often been studied in relation to criticality of underlying dynamics, namely, whether its total magnitude is maximized at the critical point between ordered and chaotic phases or not [4, 12]. The present work which does not primarily focus on the total magnitude of information flow but analyzes components of information flow will complement the existing approach.

This paper is organized as follows. In Section 2, we review RTNs and describe the parameters of numerical simulation in this paper. In Section 3, we quantify information flow by the transfer entropy. In Section 4, we review combinatorial Hodge theory in the language of linear algebra. In Section 5, the result of numerical simulation is presented. In Section 6, we discuss the degree distribution of real-world gene regulatory networks in the light of the result in Section 5. Finally, in Section 7, conclusions are given.

2. RANDOM THRESHOLD NETWORK DYNAMICS

A random threshold network (RTN) [14] consists of N nodes that are interconnected randomly. Each node can take two states ± 1 . The state of node i at time t is denoted by $x_i(t)$. It may have inputs from other nodes that determine

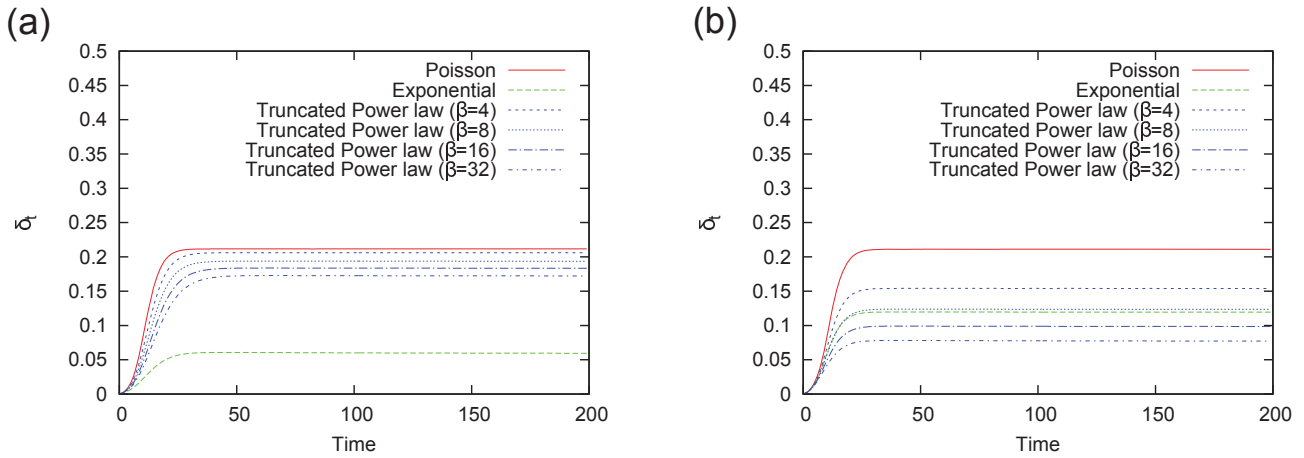


Figure 1: Time evolution of δ_t for RTNs with Poisson, exponential and truncated power law in-degree distributions (a) and out-degree distributions (b).

its future state by the following rule:

$$x_i(t+1) = \text{sgn}(f_i(t)), \quad (1)$$

where

$$f_i(t) = \sum_{j=1}^N w_{ij} x_j(t) + h_i, \quad (2)$$

and if $x \geq 0$, then $\text{sgn}(x) = 1$, otherwise, $\text{sgn}(x) = -1$. w_{ij} is the weight for inputs from node j to i . If node i receives inputs from node j , then $w_{ij} = +1$ or $w_{ij} = -1$ with equal probability. Otherwise, $w_{ij} = 0$. h_i is the threshold for node i . In the following, it is set to 0 for every node. For the time evolution of the whole system, we take the classical updating scheme, namely, all nodes are updated simultaneously.

Let us draw an arc from node j to node i when $w_{ij} \neq 0$. Then, nodes in an RTN form a directed network. The number of incoming arcs (the number of inputs) to node i is called *in-degree* of node i and is denoted by k_i . The number of outgoing arcs from node i is called *out-degree* of node i and is denoted by l_i . Let $P(k, l)$ be the probability that a randomly chosen node has in-degree k and out-degree l . In the following, we only consider the case that the in-degree and the out-degree are independent. Namely, we assume that $P(k, l) = P_{\text{in}}(k)P_{\text{out}}(l)$ where $P_{\text{in}}(k) = \sum_l P(k, l)$ is the in-degree distribution and $P_{\text{out}}(l) = \sum_k P(k, l)$ is the out-degree distribution. The average in-degree (which is equal to the average out-degree) is denoted by z .

In the following numerical simulations, we specify either in-degree or out-degree distributions and let the other half be constructed randomly. For example, when we specify an in-degree distribution, first we generate a random number following the specified in-degree distribution for each node. Second, we attach arcs of that number to each node so that their heads are directed to the node. Finally, the source node of each arc is chosen uniformly at random. In the limit of large N , the out-degree distribution follows a Poisson distribution. In turn, if we specify an out-degree distribution, then we have a Poisson in-degree distribution in the large system size limit.

It is known that RTN dynamics undergo a continuous phase transition from ordered phase to chaotic phase. One can determine the critical condition by considering damage

propagation. Let $p_s(k)$ be the damage spreading probability. It is defined as the probability that a node with in-degree k changes its state when the state of one of the input nodes is changed. The formula for $p_s(k)$ for $k \geq 1$ is known [14]:

$$p_s(k) = 2^{-(k-1)} \binom{k-1}{a_k} \quad (3)$$

where $a_k = (k-1)/2$ if k is an odd number and $a_k = k/2$ otherwise. For a given in-degree distribution $P_{\text{in}}(k)$, the critical condition calculated from a mean-field theory is [13]

$$\sum_{k=1}^{\infty} k P_{\text{in}}(k) p_s(k) = 1. \quad (4)$$

In the following, we consider three discrete probability distributions for both in-degree and out-degree distributions of underlying networks of RTNs. The first one is Poisson distributions $P(k) = e^{-\lambda} \lambda^k / k!$. If the in-degree distribution is Poissonian, then the critical parameter and the critical average in-degree are given by $\lambda_c = z_c \approx 1.85$. The second one is exponential distributions $P(k) = (1 - e^{-1/\kappa}) e^{-k/\kappa}$. If the in-degree distribution is exponential, then we have $\kappa_c \approx 2.89$ and $z_c = e^{-1/\kappa_c} / (1 - e^{-1/\kappa_c}) \approx 2.41$. Finally, the third one is truncated power law distributions $P(k) = C k^{-\alpha} e^{-k/\beta}$ where C is the normalization constant and we fix the value of β . For this, we have $\alpha_c = \infty$ and $z_c = 1$ for all $\beta > 0$. Hence, if an RTN has a truncated power law in-degree distribution, then it is always in the chaotic phase in the limit of large N .

In the following, we study the effect of different in- or out-degree distributions on components of information flow associated with RTN dynamics. In order to compare results between RTNs with different in- or out-degree distributions we fix the average in-degree (or out-degree) z . Since our primary interest in this paper is components of information flow, we choose $z = 4$ to ensure non-trivial information flow with well-defined average magnitude (We checked that we can obtain results similar to those in Section 5 for $z = 3$). RTNs with $z = 4$ are expected to reside in the chaotic phase for all the above three in-degree distributions if N is large. In Figure 1 we numerically check this in RTNs with $N = 400$. Let $\mathbf{x}(0)$ be a randomly chosen initial state of an RTN.

We flip the state of a randomly chosen node and obtain a different initial state $\mathbf{y}(0)$. We measure how this initial difference between the two state by the Hamming distance:

$$d(\mathbf{x}(t), \mathbf{y}(t)) = \frac{1}{N} \sum_{i=1}^N \frac{|x_i(t) - y_i(t)|}{2}. \quad (5)$$

Following [6], we define

$$\delta_t = d(\mathbf{x}(t), \mathbf{y}(t)) - d(\mathbf{x}(0), \mathbf{y}(0)). \quad (6)$$

Note that $d(\mathbf{x}(0), \mathbf{y}(0)) = 1/N$. $\delta^* = \lim_{t \rightarrow \infty} \delta_t < 0$ indicates that the system resides in the ordered phase. On the other hand, $\delta^* > 0$ is an evidence for the chaotic dynamics. In Figure 1 (a), δ_t s for RTNs with specified in-degree distributions are shown. It is averaged over 100 randomly chosen initial states, 100 random weight assignments and 400 random networks in each specified in-degree distribution. In every case, we can see that δ_t converges to a positive value. This suggests that the simulated RTN dynamics are in the chaotic phase. In Figure 2 (b), δ_t s for RTNs with specified out-degree distributions are shown. As in the case for specified in-degree distributions, δ_t converges to a positive value for every out-degree distribution. Note that they all have the same Poissonian in-degree distribution. According to the criticality condition (4) derived from a mean-field theory which involves only an in-degree distribution, all these cases are expected to have the same critical average in-degree $z_c \approx 1.85$. However, the value of δ^* is dependent on out-degree distributions even in the mean-field theory.

3. INFORMATION FLOW

We measure the amount of information transferred from one node to another node in an RTN by the transfer entropy [16].

Suppose that node i receives inputs from node j . Let us consider a virtual agent sitting on node i who are trying to predict the future state of node i from its present state. The amount of average uncertainty associated with his/her prediction is

$$\begin{aligned} & H(X_i(t+1)|X_i(t)) \\ &= - \sum_{x_i(t+1), x_i(t)} p(x_i(t+1), x_i(t)) \log_2 p(x_i(t+1)|x_i(t)), \end{aligned} \quad (7)$$

where $p(x_i(t+1), x_i(t))$ is the joint probability that the agent observes $x_i(t)$ at time t and $x_i(t+1)$ at time $t+1$ and $p(x_i(t+1)|x_i(t))$ is the conditional probability that the agent observes $x_i(t+1)$ at time $t+1$ given he/she observes $x_i(t)$ at time t . Now suppose in addition that the agent can access the present state of node j . He/she can integrate this information into his/her prediction of the future state of node i . The amount of average uncertainty in this case is

$$\begin{aligned} & H(X_i(t+1)|X_i(t), X_j(t)) \\ &= - \sum_{x_i(t+1), x_i(t), x_j(t)} p(x_i(t+1), x_i(t), x_j(t)) \\ & \quad \times \log_2 p(x_i(t+1)|x_i(t), x_j(t)). \end{aligned} \quad (8)$$

One way to define the amount of information transferred from node j to node i is to consider the difference

$$T_{j \rightarrow i} = H(X_i(t+1)|X_i(t)) - H(X_i(t+1)|X_i(t), X_j(t)) \quad (9)$$

which is called the *transfer entropy*. It is the reduction of average uncertainty for the agent to predict the future state of node i from its present state when he/she knows the present state of node j .

In the numerical simulations in this paper, we calculate $T_{j \rightarrow i}$ as follows. Given an RTN and a randomly chosen initial state, the first 100 transient time steps are disregarded and next 1000 time steps are used to estimate probabilities involved in the formula of $T_{j \rightarrow i}$. We averaged $T_{j \rightarrow i}$ over 100 weight assignments while keeping the underlying network topology. Note that δ_t in Figure 1 takes an almost constant value after 100 initial time steps in every case. This indicates that 100 transient time steps taken here are enough to settle the RTN dynamics on the stationary regime.

Let $e_{ij} = T_{i \rightarrow j} - T_{j \rightarrow i}$ where we regard $T_{i \rightarrow j} = 0$ if node j does not receive inputs from node i . We call matrix $e = (e_{ij})$ *information flow*. By the definition of e , it is a skew symmetric matrix, namely, $e_{ij} = -e_{ji}$. This is an instance of edge flow in combinatorial Hodge theory which we review in the next section.

4. COMBINATORIAL HODGE THEORY

We can decompose an edge flow into three components, gradient flow, harmonic flow and curl flow by the combinatorial Hodge decomposition theorem [8]. Here, we review it in the language of linear algebra.

Let $G = (V, E)$ be an undirected graph¹, where V is a finite set of vertices and E is a set of edges. Namely, E is a subset of the set of all unordered pairs of elements from the set V . Let $\binom{V}{k}$ be the set of all k -element subsets of V . Then, $E \subseteq \binom{V}{2}$. Note that $V = \binom{V}{1}$.

$$T = \left\{ \{i, j, k\} \in \binom{V}{3} \mid \{i, j\}, \{j, k\}, \{k, i\} \in E \right\} \quad (10)$$

is the set of all triangles in G .

In order to state the combinatorial Hodge decomposition theorem, we use the following notations: First,

$$C^0 = \{f : V \rightarrow \mathbb{R}\} \quad (11)$$

is a vector space over the real number field \mathbb{R} consisting of all real-valued functions on V . Second,

$$C^1 = \{e : V \times V \rightarrow \mathbb{R} \mid e_{ij} = -e_{ji}, e_{ij} = 0 \text{ for } \{i, j\} \notin E\} \quad (12)$$

is a vector space over \mathbb{R} consisting of all edge flows on G . Finally,

$$\begin{aligned} C^2 = \{t : V \times V \times V \rightarrow \mathbb{R} \mid t_{ijk} = t_{jki} = t_{kij} = \\ -t_{ikj} = -t_{kji} = -t_{jik}, t_{ijk} = 0 \text{ for } \{i, j, k\} \notin T\} \end{aligned} \quad (13)$$

is a vector space over \mathbb{R} consisting of all triangular flows on G . We make C^k ($k = 0, 1, 2$) inner product spaces by introducing the Euclidean inner products $\langle -, - \rangle_{C^k}$. Namely, $\langle f, f' \rangle_{C^0} = \sum_{i \in V} f_i g_i$, $\langle e, e' \rangle_{C^1} = \sum_{\{i, j\} \in E} e_{ij} e'_{ij}$ and $\langle t, t' \rangle_{C^2} = \sum_{\{i, j, k\} \in T} t_{ijk} t'_{ijk}$.

Now, we introduce the combinatorial gradient and curl operators and their duals. The combinatorial gradient operator $\text{grad} : C^0 \rightarrow C^1$ is a linear map defined by

$$\text{grad}(f)_{ij} = f_j - f_i. \quad (14)$$

¹We use the terms *network* and *graph* interchangeably.

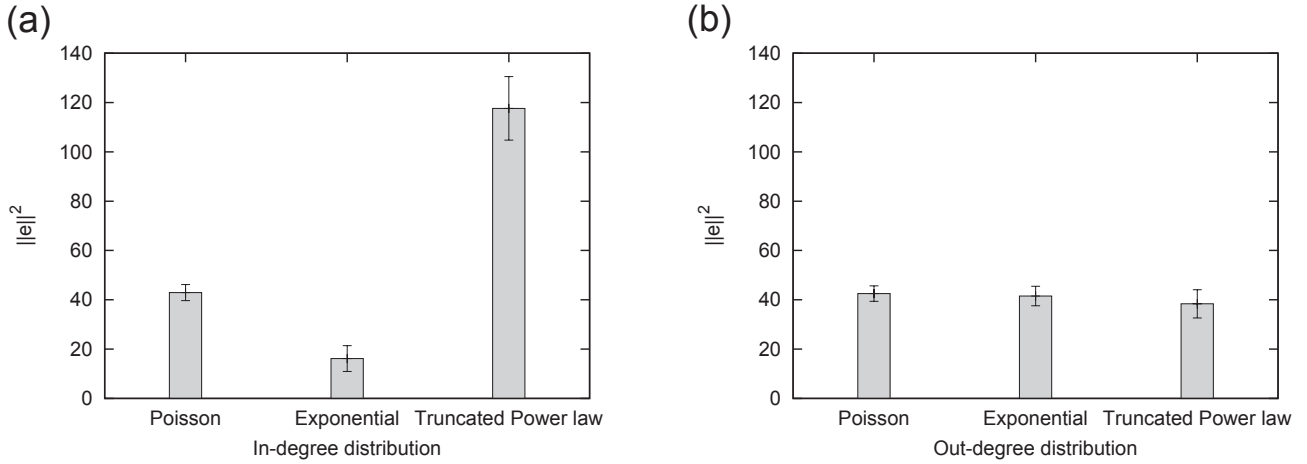


Figure 2: The magnitude of information flow $\|e\|^2$ associated with RTN dynamics with Poisson, exponential and truncated power law ($\beta = 32$) in-degree distributions (a) and out-degree distributions (b).

The combinatorial curl operator $\text{curl} : C^1 \rightarrow C^2$ is a linear map defined by

$$\text{curl}(e)_{ijk} = e_{ij} + e_{jk} + e_{ki}. \quad (15)$$

Their dual operators are denoted by $\text{grad}^* : C^1 \rightarrow C^0$ and $\text{curl}^* : C^2 \rightarrow C^1$, respectively. It turns out that

$$\text{grad}^*(e)_i = - \sum_{j \text{ s.t. } \{i,j\} \in E} e_{ij} \quad (16)$$

from the adjoint relationship $\langle \text{grad}(f), e \rangle_{C^1} = \langle f, \text{grad}^*(e) \rangle_{C^0}$. If we introduce the combinatorial divergence operator $\text{div} : C^1 \rightarrow C^0$ by $\text{div}(e)_i = \sum_{j \text{ s.t. } \{i,j\} \in E} e_{ij}$, then we can write $\text{grad}^* = -\text{div}$. Similarly, for curl^* , we have

$$\text{curl}^*(t)_{ij} = \sum_{k \text{ s.t. } \{i,j,k\} \in T} t_{ijk}. \quad (17)$$

Note that the following closedness property holds: $\text{curl} \circ \text{grad} = 0$ and $\text{div} \circ \text{curl}^* = 0$.

The combinatorial Hodge decomposition theorem asserts that C^1 admits the following orthogonal decomposition:

$$C^1 = \text{Im}(\text{grad}) \oplus \text{Ker}(\Delta_1) \oplus \text{Im}(\text{curl}^*), \quad (18)$$

where $\Delta_1 : C^1 \rightarrow C^1$ defined by $\Delta_1 = \text{curl}^* \circ \text{curl} + \text{grad} \circ \text{grad}^* = \text{curl}^* \circ \text{curl} - \text{grad} \circ \text{div}$ is the one-dimensional *combinatorial Laplacian* (which is called the *graph Helmholtzian* in [8]).² Furthermore, we have

$$\text{Ker}(\text{div}) = \text{Ker}(\Delta_1) \oplus \text{Im}(\text{curl}^*) \quad (19)$$

and

$$\text{Ker}(\Delta_1) = \text{Ker}(\text{div}) \cap \text{Ker}(\text{curl}). \quad (20)$$

$\text{Im}(\text{grad})$ is the subspace of *gradient flows*. Any gradient flow $g \in \text{Im}(\text{grad})$ can be written as a difference of some (negative) potential function $f \in C^0$: $g_{ij} = f_j - f_i$ for all $\{i, j\} \in E$. It represents an *acyclic unidirectional* flow on G .

$\text{Ker}(\text{div})$ is the subspace of *loop flows*. Any loop flow l satisfies $\text{div}(l) = 0$ which means that flow is conserved at any

²Note that the zero-dimensional combinatorial Laplacian $\Delta_0 : C^0 \rightarrow C^0$ defined by $\Delta_0 = \text{grad}^* \circ \text{grad} = -\text{div} \circ \text{grad}$ is so-called the *graph Laplacian*.

vertex i . This implies that any non-zero loop flow l contains a loop such that $i_0 \rightarrow i_1 \rightarrow \dots \rightarrow i_k = i_0$ where arc $i_j \rightarrow i_{j+1}$ indicates $l_{ij}i_{j+1} > 0$ ($j = 0, \dots, k-1$). It represents *circular flows* on G . Equation (19) further decomposes the space of loop flows into two components: One is the space of *harmonic flows* $\text{Ker}(\Delta_1)$ and the other is the space of *curl flows* $\text{Im}(\text{curl}^*)$. By Equation (20), any harmonic flow h has no loop of length 3 (namely, triangles) such that $i \rightarrow j \rightarrow k \rightarrow i$ (namely, $h_{ij}, h_{jk}, h_{ki} > 0$). It represents *global circular flows*. On the other hand, curl flows c may have non-zero curls along triangles. They represent *local circular flows*.

We measure the magnitude of an edge flow e by the square of its l^2 -norm $\|e\|^2 = \langle e, e \rangle_{C^1}$. Let $e = g + l = g + h + c$ be the Hodge decomposition of e where g is the gradient flow component, l is the loop flow component, h is the harmonic flow component and c is the curl flow component. By the orthogonality of the decomposition, we have

$$\|e\|^2 = \|g\|^2 + \|l\|^2 = \|g\|^2 + \|h\|^2 + \|c\|^2. \quad (21)$$

The relative strength of each component can be measured by the following quantities: *gradient ratio* $\gamma = \|g\|^2/\|e\|^2$, *loop ratio* $\lambda = \|l\|^2/\|e\|^2$, *harmonic ratio* $\eta = \|h\|^2/\|e\|^2$ and *curl ratio* $\chi = \|c\|^2/\|e\|^2$. These are characteristics of an individual edge flow.

We also consider the relative size of each subspace: *structural gradient ratio* $\Gamma = \dim(\text{Im}(\text{grad}))/\dim(C^1)$, *structural loop ratio* $\Lambda = \dim(\text{Ker}(\text{div}))/\dim(C^1)$, *structural harmonic ratio* $H = \dim(\text{Ker}(\Delta_1))/\dim(C^1)$ and *structural curl ratio* $X = \dim(\text{Im}(\text{curl}^*))/\dim(C^1)$. Note that these quantities are determined by the structure of the underlying graph G . In particular, Γ and Λ can be calculated immediately from the fact that $\dim(C^1)$ is equal to the number of edges in G and $\dim(\text{Im}(\text{grad}))$ is equal to the number of vertices minus the number of connected components in G . Since the average relative strength of each component of edge flows of a fixed l^2 -norm chosen uniformly at random is equal to corresponding structural ratio, structural ratios provide a reference point for how large each component of an individual edge flow is biased.

5. RESULT

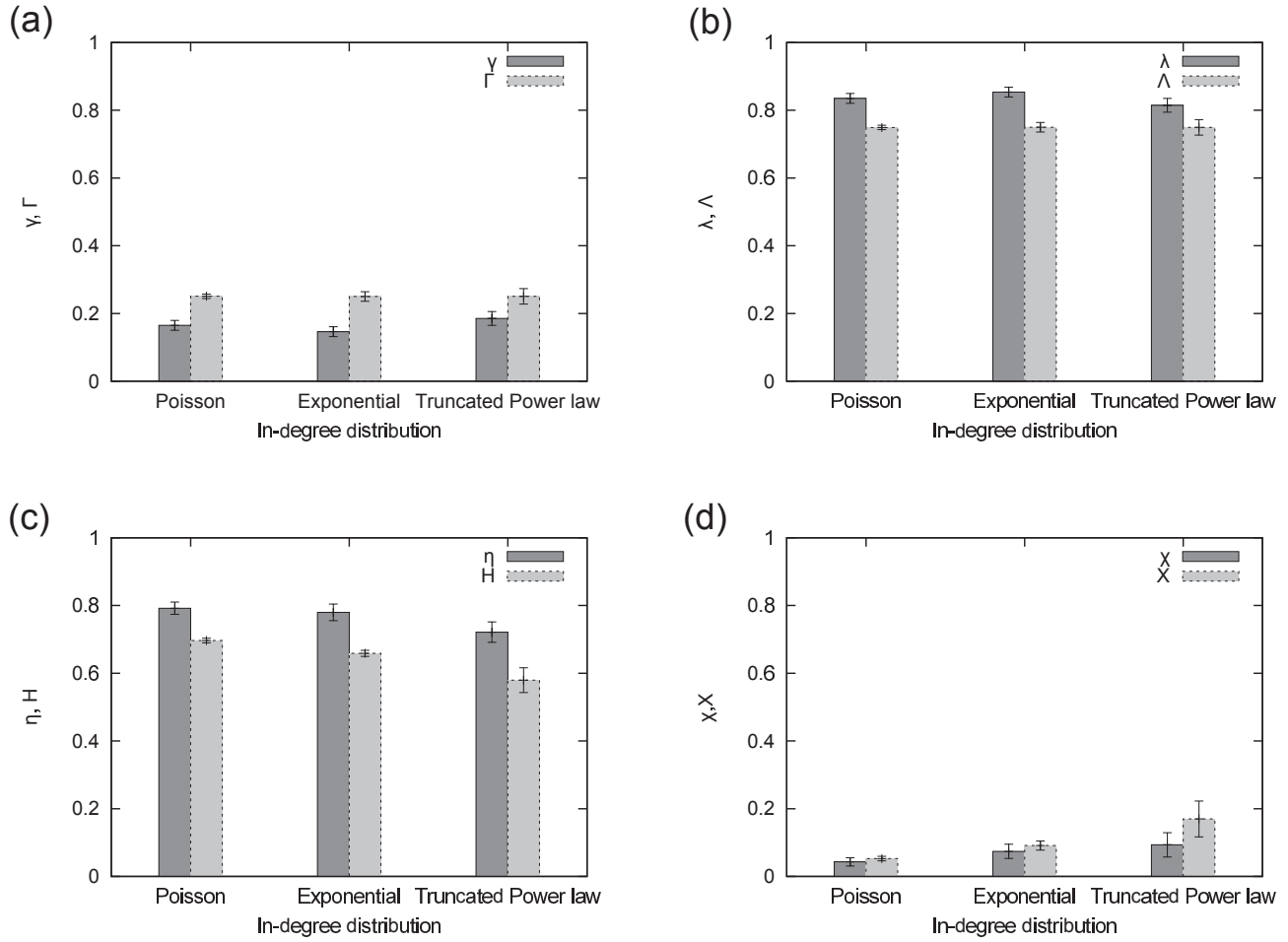


Figure 3: The relative strengths of components of information flow associated with RTN dynamics with Poisson, exponential and truncated power law ($\beta = 32$) in-degree distributions. Corresponding structural ratios are shown together. (a) Gradient ratio γ and structural gradient ratio Γ . (b) Loop ratio λ and structural loop ratio Λ . (c) Harmonic ratio η and structural harmonic ratio H . (d) Curl ratio χ and structural curl ratio X .

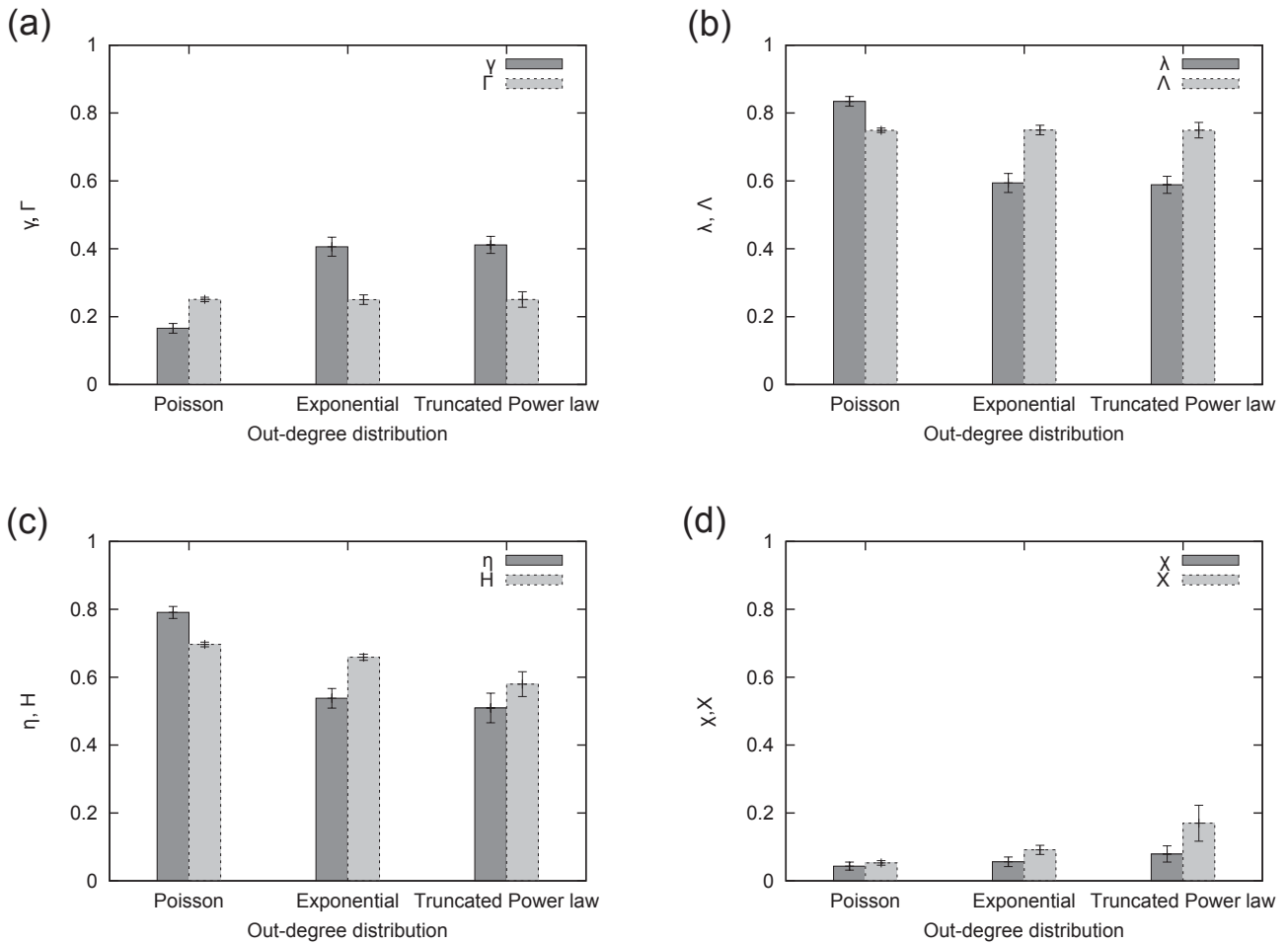


Figure 4: The relative strengths of components of information flow associated with RTN dynamics with Poisson, exponential and truncated power law ($\beta = 32$) out-degree distributions. Corresponding structural ratios are shown together. (a) Gradient ratio γ and structural gradient ratio Γ . (b) Loop ratio λ and structural loop ratio Λ . (c) Harmonic ratio η and structural harmonic ratio H . (d) Curl ratio χ and structural curl ratio X .

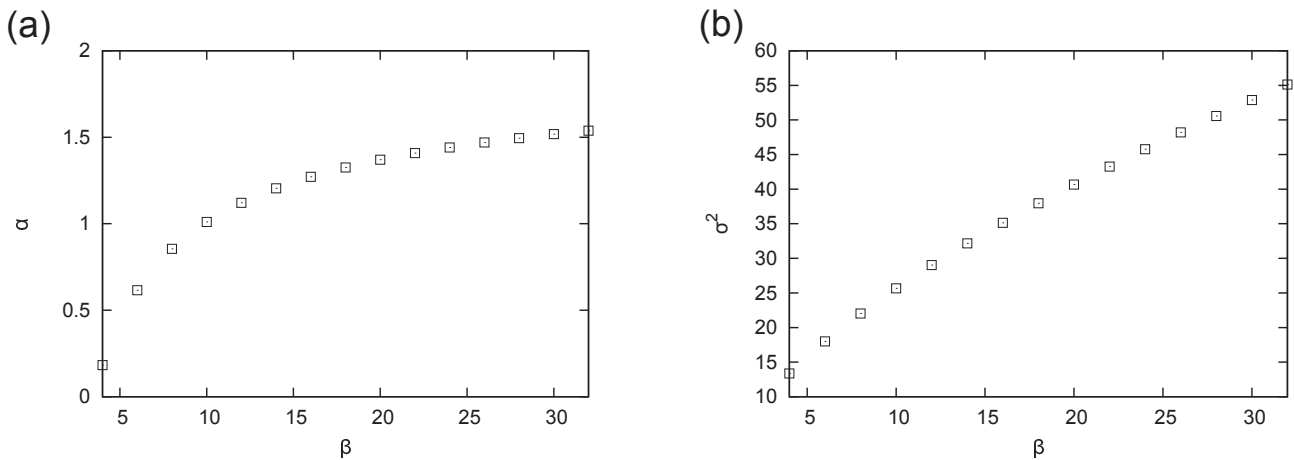


Figure 5: Exponent α of truncated power law distribution with average 4 (a) and its variance (b) for each value of the cutoff parameter β .

We apply the combinatorial Hodge decomposition theorem to information flow on RTNs introduced in Section 3. In the following, all quantities are averaged over 400 RTNs and error bars indicate standard deviations.

First, we compare results between RTNs with different type of in- or out-degree distributions: Poisson, exponential and truncated power law with $\beta = 32$. The average in-degree is fixed at $z = 4$ for all the distributions. The variances are 4, 20 and 55.11, respectively.

Figure 2 (a) shows the magnitude of information flow $\|e\|^2$ when an in-degree distribution is specified. The type of the in-degree distribution greatly affects the value of $\|e\|^2$. The case for the exponential in-degree distribution has the value of $\|e\|^2$ less than half that of the case for the Poisson in-degree distribution. On the other hand, the case for the truncated power law in-degree distribution has the value of $\|e\|^2$ almost three times larger than that of the case for the Poisson in-degree distribution. These result could be explained by different critical average in-degrees. However, the relationship between damage spreading and the magnitude of information flow is not so simple because the case for the Poisson in-degree distribution has the largest value of δ^* .

Figure 2 (b) shows the magnitude of information flow $\|e\|^2$ when an out-degree distribution is specified. In contrast to the cases for specified in-degree distributions, the type of the out-degree distribution has almost no impact on the value of $\|e\|^2$. This may reflect the fact that they all have the same Poisson in-degree distribution.

In Figure 3, the relative strengths of components of information flow are shown together with structural ratios when an in-degree distribution is specified. First note that the structural gradient ratio Γ is approximately 0.25 for every case. This is due to the fact that almost all of nodes are included in the largest connected component in the present numerical simulation condition. If we assume that it has exactly one connected component, then $\Gamma = (N - 1)/(zN) \approx 1/z = 0.25$. This implies that $\Lambda = 1 - \Gamma \approx 0.75$. From Figure 3 (a), we can see that the gradient ratio γ has almost the same value for every case and is significantly less than the structural gradient ratio Γ . Correspondingly, the loop ratio λ has almost the same value for every case and is significantly greater than the structural loop ratio Λ as we can see from Figure 3 (b). The result of further decomposition of loop flows is shown in Figure 3 (c) and (d). We can see that dominant component in loop flows is harmonic flows. We note that the structural curl ratios converge to 0 as N goes to ∞ . This is because the expected number of triangles is constant in the limit of large N for configuration model random networks [11].

In Figure 4, we show the relative strengths of components of information flow together with structural ratios when an out-degree distribution is specified. Note that the structural ratios have almost the same values as those in the case for specified in-degree distributions. This is because the underlying undirected network does not change when the in-degree and the out-degree are exchanged at each node. In contrast to the case for specified in-degree distributions, the gradient ratio γ is significantly larger than the structural gradient ratio Γ for exponential and truncated power law out-degree distributions (Figure 4 (a)) and, correspondingly, the loop ratio λ is significantly smaller than the structural loop ratio

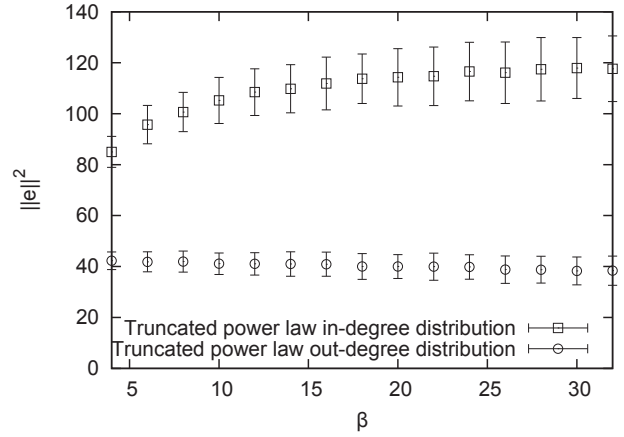


Figure 6: The magnitude of information flow $\|e\|^2$ associated with RTN dynamics with truncated power law in-degree distributions (square) and out-degree distributions (circle).

Λ in these cases (Figure 4 (b)). Loop flows are suppressed by out-degree distributions with longer tails. For components of loop flows, the harmonic counter part is dominant as in the case for specified in-degree distributions.

Second, we show results for truncated power law distributions when the cutoff parameter β is varied. The average in-degree is fixed at $z = 4$. Figure 5 (a) shows the value of exponent α . The variance σ^2 of the distribution is shown in Figure 5 (b). As expected, the larger β is, the larger σ^2 is.

Figure 6 to 8 show the similar tendencies as those found in Figure 2 to 4, respectively. The following four observations are noticeable: First, the magnitude of information flow $\|e\|^2$ becomes larger as β gets larger when an in-degree distribution is specified (Figure 6, squares). Second, in contrast, it is almost constant when an out-degree distribution is specified (Figure 6, circles). Third, the variance of in-degree distributions has little influence on the strength of gradient flows or loop flows (Figure 7). Fourth, in contrast, the larger the variance of the out-degree distribution is, the smaller the loop ratio λ is, when an out-degree distribution is specified (Figure 8).

6. DISCUSSION

In Section 5, we have shown by Hodge decomposing information flow on RTNs that in-degree distributions contribute little to the balance between gradient flow and loop flow, while out-degree distributions have significant impact on the balance. In particular, it is found that having longer-tailed out-degree distributions makes the strength of loop flow smaller.

This finding may give a new insight on the fact that real-world gene regulatory networks have approximately scale-free out-degree distributions, while their in-degree distributions are approximately Poissonian [1]. For example, this holds approximately true of the gene regulatory network of bacterium *E. coli* [3].

In order to numerically investigate the balance between gradient component and loop component of information flow on *E. coli* gene regulatory network, we run RTN dynamics on the directed network reconstructed from real-world

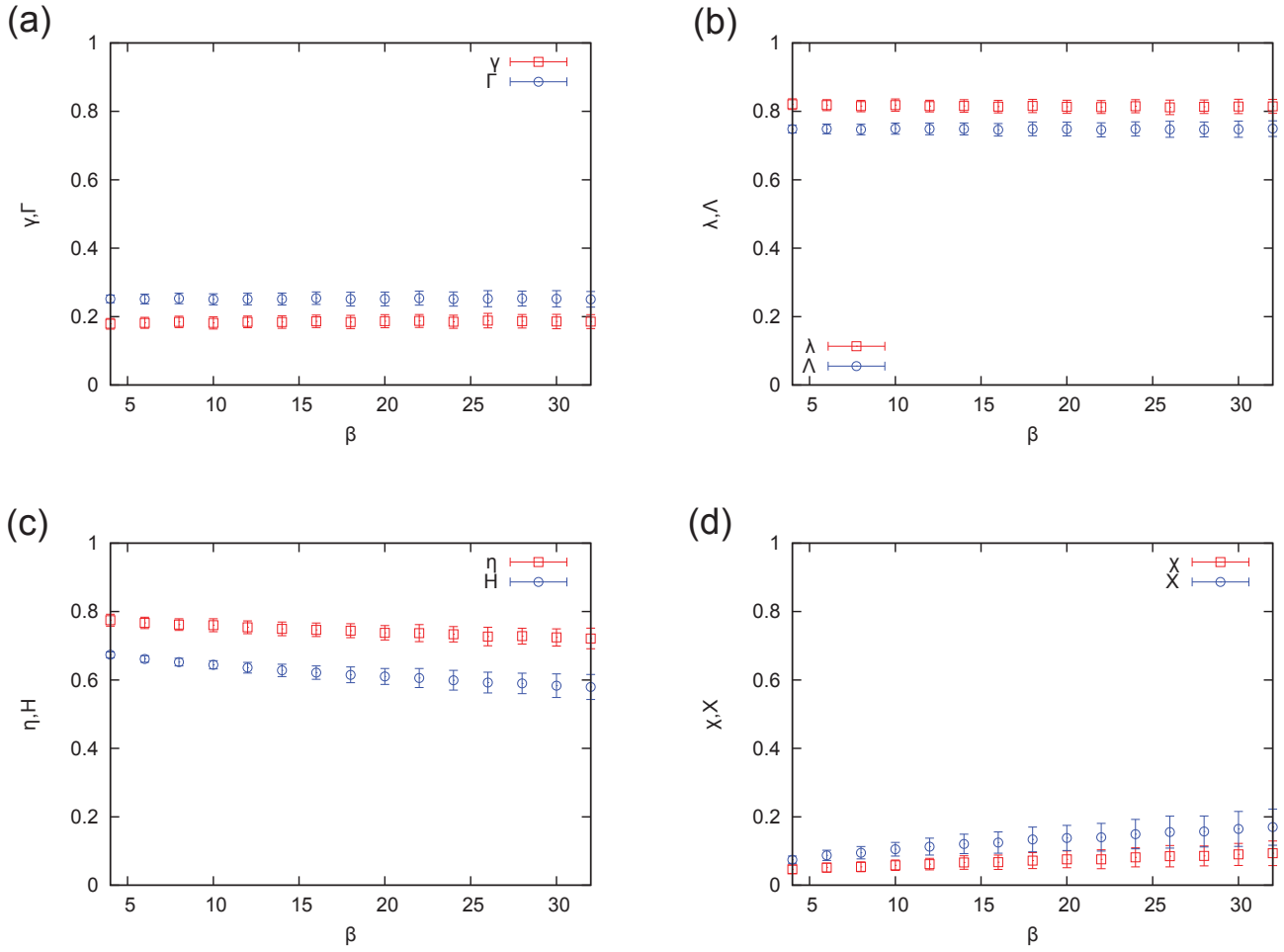


Figure 7: The relative strengths of components of information flow associated with RTN dynamics with truncated power law in-degree distributions. Corresponding structural ratios are shown together. (a) Gradient ratio γ and structural gradient ratio Γ . (b) Loop ratio λ and structural loop ratio Λ . (c) Harmonic ratio η and structural harmonic ratio H . (d) Curl ratio χ and structural curl ratio X .

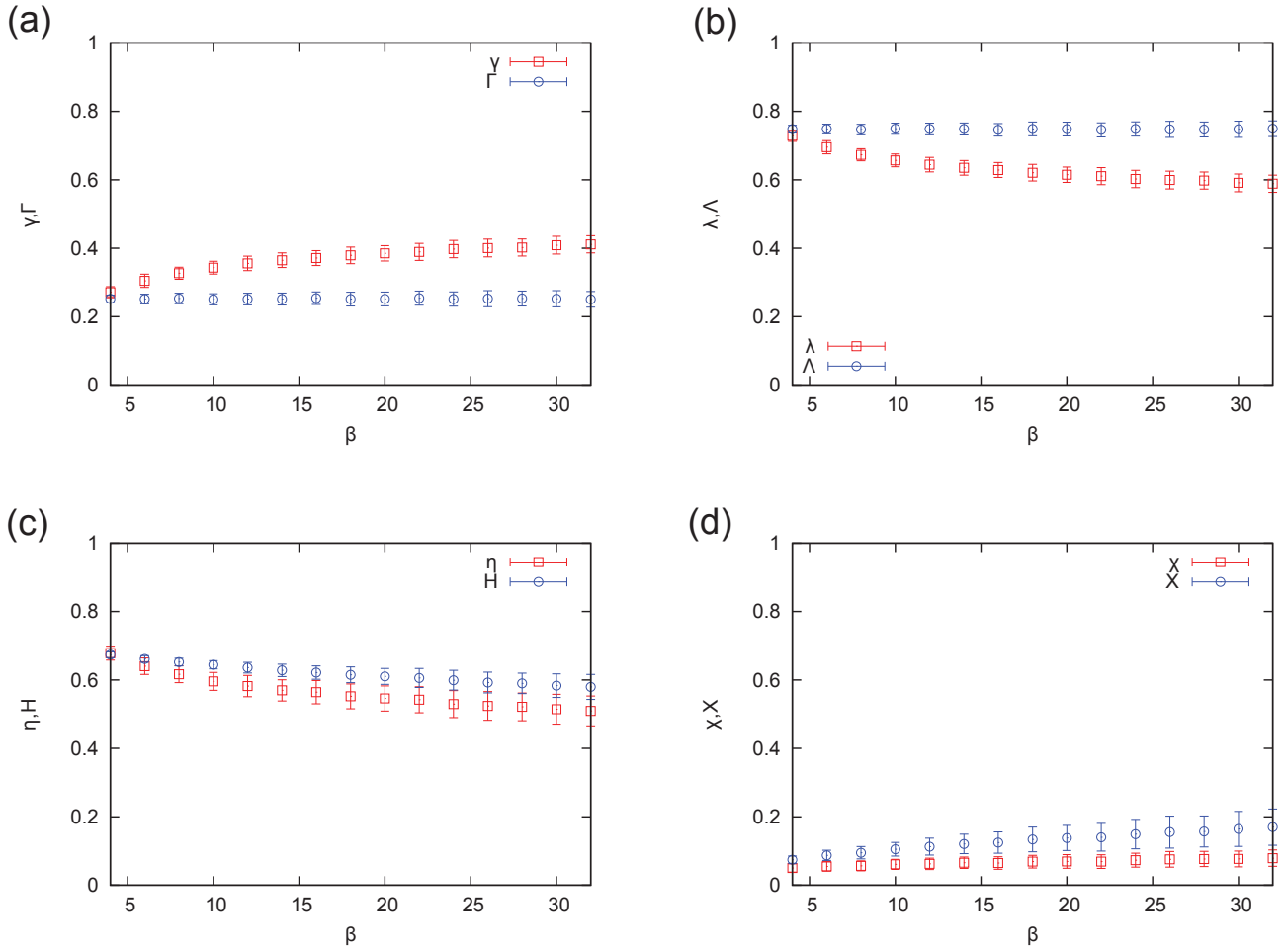


Figure 8: The relative strengths of components of information flow associated with RTN dynamics with truncated power law out-degree distributions. Corresponding structural ratios are shown together. (a) Gradient ratio γ and structural gradient ratio Γ . (b) Loop ratio λ and structural loop ratio Λ . (c) Harmonic ratio η and structural harmonic ratio H . (d) Curl ratio χ and structural curl ratio X .

data [15] and construct information flow as the same way described in Section 3. The *E. coli* gene regulatory network has the following parameter values: the number of node $N = 1763$, the number of arcs $A = 3991$, average in-(or out-) degree $z = 2.26$, variance of in-degree $\text{Var}(k) = 3.00$, variance of out-degree $\text{Var}(l) = 326.66$ and correlation coefficient between in-degree and out-degree $\rho = -0.015$. This relatively small ρ value allows us to assume that $P(k, l) \approx P_{\text{in}}(k)P_{\text{out}}(l)$, the condition which we introduced in Section 2. However, the variance of out-degree is extremely larger compared to the one considered in Section 5. Hence, we expect that the *E. coli* gene regulatory network has quite small value of the loop ratio λ relative to the structural loop ratio Λ . Indeed, we find that it has $\lambda = 0.053$ and $\Lambda = 0.550$. This suggests that information flow on the *E. coli* gene regulatory network is almost specialized for gradient flow. This result is consistent with the view that the primary function of gene regulation networks is to unidirectionally relay environmental signals to transcription responses in feed-forward manner [1] for which circular information flow would be less useful.

7. CONCLUSIONS

In this paper, we have numerically investigated Hodge decomposition of information flow associated with RTN dynamics on complex networks. In particular, we have studied influence of underlying network topology on components of information flow. As a first step for this direction of study, we have controlled in-degree or out-degree distributions and leave other features random. It is found that in-degree distributions have little impact on the balance between gradient component and loop component of information flow, while long-tailed out-degree distributions suppress loop component significantly. We have inferred that this finding shed a new light on the efficacy of scale-free out-degree distributions observed in real-world gene regulatory networks.

Real-world complex networks have structural features other than scale-free degree distribution such as small-world property, degree correlations, network motifs, community structure and so on [11]. Investigation of the influence of these network structures on information flow is left as future work.

8. ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 25280091.

9. REFERENCES

- [1] R. Albert. Scale-free networks in cell biology. *J. Cell Sci.*, 118:4947–4957, 2005.
- [2] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, 2002.
- [3] M. Aldana, E. Balleza, S. Kauffman, and O. Resendiz. Robustness and evolvability in genetic regulatory networks. *J. Theor. Biol.*, 245:433–448, 2007.
- [4] L. Barnett, J. T. Lizier, M. Harré, A. K. Seth, and T. Bossomaier. Information flow in a kinetic ising model peaks in the disordered phase. *Phys. Rev. Lett.*, 111:177203, 2013.
- [5] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Phys. Rep.*, 424:175–308, 2006.
- [6] C. Gershenson. Phase transitions in random boolean networks with different updating schemes. arXiv:nlin/0311008v1, 2003.
- [7] K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka, and J. Bhattacharya. Causality detection based on information-theoretic approaches in time series analysis. *Phys. Rep.*, 441:1–46, 2007.
- [8] X. Jiang, L.-H. Lim, Y. Yao, and Y. Ye. Statistical ranking and combinatorial hodge theory. *Math. Program., Ser. B*, 127:203–244, 2011.
- [9] J. T. Lizier, S. Pritam, and M. Prokopenko. Information dynamics in small-world boolean networks. *Artificial Life*, 17:293–314, 2011.
- [10] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.
- [11] M. E. J. Newman. *Networks: An Introduction*. Oxford Univ. Press Inc., New York, 2010.
- [12] 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.
- [13] T. Rohlf. Critical line in random-threshold networks with inhomogeneous thresholds. *Phys. Rev. E*, 78:066118, 2008.
- [14] T. Rohlf and S. Bornholdt. Criticality in random threshold networks: annealed approximation and beyond. *Physica A*, 310:245–259, 2002.
- [15] H. Salgado, M. Peralta-Gil, S. Gama-Castro, A. Santos-Zavaleta, L. Muniz-Rascado, J. S. Garcia-Sotelo, V. Weiss, H. Solano-Lira, I. Martinez-Flores, A. Medina-Rivera, G. Salgado-Osorio, S. Alquicira-Hernandez, K. Alquicira-Hernandez, A. Lopez-Fuentes, L. Porron-Sotelo, A. M. Huerta, C. Bonavides-Martinez, Y. I. Balderas-Martinez, L. Pannier, M. Olvera, A. Labastida, V. Jimenez-Jacinto, L. Vega-Alvarado, V. del Moral-Chavez, A. Hernandez-Alvarez, E. Morett, and J. Collado-Vides. RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. *Nucl. Acids Res.*, 41:D203–D213, 2013.
- [16] T. Schreiber. Measuring information transfer. *Phys. Rev. Lett.*, 85:461–464, 2000.