Entropy production and fluctuation theorems on complex networks

Cite as: Chaos **30**, 053125 (2020); https://doi.org/10.1063/1.5143031
Submitted: 20 December 2019 . Accepted: 22 April 2020 . Published Online: 12 May 2020

Jaewoo Jung, Jaegon Um, Deokjae Lee, Yong W. Kim, D. Y. Lee, H. K. Pak, and B. Kahng

Entropy production and fluctuation theorems on complex networks

Cite as: Chaos **30**, 053125 (2020); doi: 10.1063/1.5143031
Submitted: 20 December 2019 · Accepted: 22 April 2020 ·
Published Online: 12 May 2020

Jaewoo Jung,[1] Jaegon Um,[1,2] Deokjae Lee,[1] Yong W. Kim,[3] D. Y. Lee,[4] H. K. Pak,[4,5] and B. Kahng[1,a)]

AFFILIATIONS

[1]CCSS, CTP and Department of Physics and Astronomy, Seoul National University, Seoul 08826, South Korea
[2]BK21PLUS Physics Division, Pohang University of Science and Technology, Pohang 37673, South Korea
[3]Department of Physics, Lehigh University, Bethlehem, Pennsylvania 18015, USA
[4]Center for Soft and Living Matter, Institute for Basic Science, Ulsan 44919, South Korea
[5]Department of Physics, Ulsan National Institute of Science and Technology, Ulsan 44949, South Korea

[a)]**Author to whom correspondence should be addressed:** bkahng@snu.ac.kr

ABSTRACT

Entropy production (EP) is a fundamental quantity useful for understanding irreversible process. In stochastic thermodynamics, EP is more evident in probability density functions of trajectories of a particle in the state space. Here, inspired by a previous result that complex networks can serve as state spaces, we consider a data packet transport problem on complex networks. EP is generated owing to the complexity of pathways as the packet travels back and forth between two nodes along the same pathway. The total EPs are exactly enumerated along all possible shortest paths between every pair of nodes, and the functional form of the EP distribution is proposed based on our numerical results. We confirm that the EP distribution satisfies the detailed and integral fluctuation theorems. Our results should be pedagogically helpful for understanding trajectory-dependent EP in stochastic processes and exploring nonequilibrium fluctuations associated with the entanglement of dividing and merging among the shortest pathways in complex networks.

Entropy production (EP) is essential for a deeper understanding of stochastic dynamics in nonequilibrium systems. Recently, in stochastic thermodynamics, EP is determined based on trajectories of a particle in the state space. However, because the trajectories are virtual, one can hardly imagine the origin of the EP and understand the related fluctuation theorems. Here, we model the state space as scale-free networks and consider a transport problem based on them similar to a packet transport on the Internet. A packet travels back and forth along all possible shortest pathways between every pair of nodes. Owing to the complexity of pathways of scale-free networks, the EP distribution is asymmetric and satisfies the detailed and integral fluctuation theorems. This result suggests that the scale-free networks may serve as a platform for further research on stochastic dynamics.

I. INTRODUCTION

Entropy production (EP) plays a key role in quantifying dissipative work, such as heat, in the second law of thermodynamics. Further, in nonequilibrium systems, EP serves as a key quantity to measure the degree of irreversibility of a stochastic process.[1–4] Recently, EP is discussed in terms of trajectories in phase space of a single particle;[5,6] it is defined as the logarithm of the ratio of probabilities that a dynamic proceeds in the forward and corresponding reverse directions along a single pathway in phase space between two states.[6,7] The ensemble of these EPs over all possible trajectories between every pair of states has an EP distribution (EPD).

EPD satisfies fluctuation theorems (FTs), viz., the integral FT and the detailed FT, developed by Crooks,[8] Jarzynski,[9] and others for the dissipated work in nonequilibrium systems.[10–13] The relation between the EPD and FTs has been experimentally tested in the work distribution for various experimental setups,[14,15] including RNA folding,[16] colloidal suspensions,[17] and electric circuit.[18] However, it is extremely difficult to experimentally find the probability density function of each trajectory along which a particle proceeds in the forward and reverse directions.[15] This is because the trajectories are virtual and the number of trajectories increases exponentially as the number of steps is increased.

Here, we aim to obtain a complete set of the probabilities of dynamic trails for a given stochastic process between every two

states. Then we obtain an EPD and justify the FTs. Hence, we consider a biased random walk along the shortest pathways between every pair of nodes on complex networks, which represent the state spaces of the stochastic process.

Complex networks may be regarded as state spaces. For instance, a node in the protein folding network represents a protein conformation, and a link between two nodes is connected when a protein conformation is changed to another in consecutive steps.[19] Recent studies using molecular dynamics simulations revealed that the protein folding network is not a random network that has degrees following a Poisson distribution. Instead, it is a scale-free (SF) network that has a power-law degree distribution,[20] where a few nodes are connected to many nodes and the remaining many nodes are connected to a few nodes.

The protein folding dynamics from denatured states to the natural state[19,20] proceed along a few major pathways on the protein folding network, rather than along the pathways randomly selected at each node as in a diffusion process.[21] Thus, the folding dynamics may be regarded as biased random walks along the shortest pathway on the conformation network. This view may be drastic in certain systems. However, for the system where the stochastic process along the shortest pathways is so dominant that the contributions by other pathways are negligible, this view may be justified. Considering this simplification, we are able to identify an essential factor that produces nontrivial EPs. We will further discuss the asymmetric probabilities (transition rate) between two nodes (states) in the forward and reverse directions naturally arising from the asymmetric entanglement of the shortest pathways: if the shortest pathway between a pair of nodes is one channel, then the EP would be zero. However, the shortest pathways between two randomly selected nodes could have more than one channel in complex networks. In such a case, the shortest pathways are entangled, divided, and/or merged. Hence, the probability to pass through a link in one direction can be different from that in a reverse direction. The EP then is nonzero. We enumerate all EPs, composed of $N_{\rm EP} \equiv n_{\rm sp} N(N-1)$ elements, where $n_{\rm sp}$ is the mean number of shortest pathways between a pair of nodes. The EPD is completely composited.

The biased random walks also occur in a data packet transport on the Internet. Assuming that a data packet is sent from one node to a destination on the Internet at each time step, the packet is transmitted to a neighbor according to the router protocol toward the final destination. Unless the traffic is congested, packets travel along the shortest path between the two nodes,[22] thus considering biased random walks on the shortest pathways is natural for packet dynamics on the Internet. In social networks, the flow along the shortest pathways on complex networks was used to quantify a person's influence in society.[23] An efficient algorithm for identifying every possible shortest pathway between every pair of nodes costs computational complexity $O(N^2 \log N)$, where N is the network size.[23–25] Thus, obtaining the EPs and their distribution can be implemented within reasonable system sizes.

Further, we consider a packet transport along the shortest pathway on random and scale-free networks and calculate the probabilities of passing across each link on the shortest pathways when each pair of nodes sends and receives a packet. These probabilities generate EPs and, thereby, the EPD. Using this complete EPD, we verify analytically and numerically the detailed and integral FTs. We show that the cumulative EPD exhibits a double exponential form.

This paper is organized as follows: In Sec. II, we introduce an EP induced by the topological complexity of the shortest pathways on a network. In Sec. III, we verify that the total EP obtained from every shortest pathway satisfies the integral FT and the detailed FT. In Sec. IV, we present numerical results of the EPD for several model networks. In Sec. V, we propose a functional form of the cumulative EPD based on numerical results and determine the EPD. A summary is presented in Sec. VI.

II. EP ON NETWORKS

We consider data packet transport from a source node i to a target node j along a shortest pathway $\vec{\alpha}$ of length d_{st} on a given network, where $\vec{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_{d_{st}-1}, \alpha_{d_{st}})$. In the sequence $\vec{\alpha}$, each element stands for the node on the shortest pathway with the boundary condition, $\alpha_0 = i$ and $\alpha_{d_{st}} = j$. Then, the probability $P[\vec{\alpha}]$ that transport occurs along the pathway $\vec{\alpha}$ is given as $P[\vec{\alpha}] = \rho(i)\rho(j|i)\,\Pi[\vec{\alpha}; i, j]$, where $\rho(i)$ denotes the probability that node i is selected as a source and $\rho(j|i)$ is the conditional probability that node j is chosen as a target. Given the pair of source and target, (i, j), the transition probability $\Pi[\vec{\alpha}; i, j]$ along $\vec{\alpha}$ is determined by the topology of the shortest pathways from i to j. The shortest pathway is either single or multiple, and multiple pathways are either in parallel or entangled, as shown in Fig. 1. As a packet travels along the shortest pathways, it can reach a branching node. Then, the packet chooses one branch among all the branches with probability $1/N_b$, where N_b is the number of branches on the shortest pathway. This random choice is repeated as the packet reaches a branching node. When the packet reaches a target, the probability of taking that shortest pathway can be calculated as the product

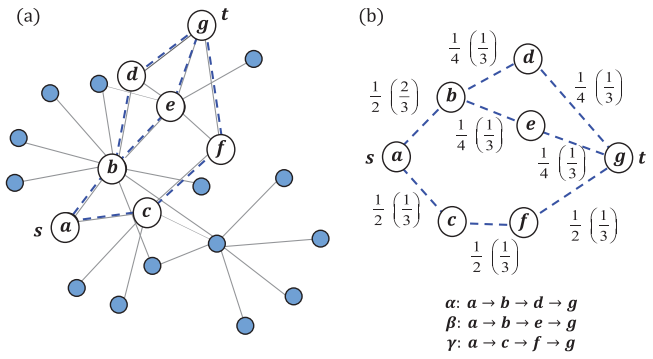

FIG. 1. (a) Sample network to illustrate the EPs along each shortest path from *a* to *g* and shortest return path from *g* to *a*. For pathway α, a data packet travels along the pathway $a \to b \to d \to g$ and returns in the reverse direction $g \to d \to b \to a$. (b) At node *a*, there are two ways to move toward node *g* with equal probability. The packet takes the link $a \to b$ with conditional probability 1/2. Next, it takes the link $b \to d$ with conditional probability 1/2. The link $d \to g$ is taken with conditional probability one. Accordingly, the transition probability along the pathway $\vec{\alpha}$, denoted as $\Pi[\vec{\alpha}; a, g]$, is found to be 1/4. In the reverse trajectory $\vec{\alpha}'$, the transition probability $\Pi[\vec{\alpha}'; g, a]$ is found to be 1/3. Table I shows the transition probabilities along each shortest pathway in the forward and corresponding reverse directions.

of these probabilities, as illustrated in Fig. 1. The effect of this random choice mimics the stochastic noise in the dynamic process. Let us consider the reverse process where the packet returns along the corresponding reverse path $\vec{\alpha}'$ of $\vec{\alpha}$ from node j to node i, where $\vec{\alpha}' = (\alpha_{d_{st}}, \alpha_{d_{st}-1}, \ldots, \alpha_1, \alpha_0)$. Since $\vec{\alpha}'$ is also one of the shortest pathways from source j to target i, one can define the probability $P[\vec{\alpha}']$ in the same manner as $P[\vec{\alpha}]$. In general, the probabilities for the path $\vec{\alpha}$ and its reverse one $\vec{\alpha}'$ may be different. The discrepancy can be regarded as the irreversibility for the transport along $\vec{\alpha}$, and thus the corresponding EP is defined as follows:

$$\Delta S[\vec{\alpha}] = \ln\left[\frac{P[\vec{\alpha}]}{P[\vec{\alpha}']}\right] = \ln\left[\frac{\rho(i)\rho(j|i)\,\Pi[\vec{\alpha};i,j]}{\rho(j)\rho(i|j)\,\Pi[\vec{\alpha}';j,i]}\right]. \quad (1)$$

In this problem, $\rho(i) = \rho(j) = 1/N$, because the node is selected randomly from among N nodes. The conditional probability is also given by $\rho(j|i) = \rho(i|j) = 1/(N-1)$ because node j (i) is randomly selected from $N-1$ nodes excluding node i (j). Therefore, the nonzero EP in Eq. (1) is caused by only the difference between $\Pi[\vec{\alpha};i,j]$ and $\Pi[\vec{\alpha}';j,i]$. We will show that $\Pi[\vec{\alpha};i,j]$ can differ from $\Pi[\vec{\alpha}';j,i]$ owing to the entanglement of dividing and merging among the shortest pathways on complex networks.

We consider a simple example to explain how to calculate the transition probabilities on the shortest pathways. Figure 1 is a subgraph of a network showing the shortest pathways between two nodes, a and g as the source (s) and target (t), respectively. There exist three shortest pathways, which are denoted as $\vec{\alpha}$, $\vec{\beta}$, and $\vec{\gamma}$, with the length $d_{st} = 3$. Let us first consider packet transport along the pathway $\vec{\alpha} = (a, b, d, g)$, from a toward node g. At node a, the packet needs to choose either node b or node c, which we assume are chosen with equal probability, as the site of the next step. Thus, hopping from a to b occurs with probability 1/2, as does hopping from a to c. Next, it chooses node d with conditional probability 1/2, because the pathway is divided into two possibilities. Thus, the packet arrives at node d with probability 1/4. Then, it travels to the target $t = g$ without any branching, i.e., with conditional probability one. Accordingly, the transition probability is given as $\Pi[\vec{\alpha};a,g] = 1/4$ along the pathway $a \xrightarrow{1/2} b \xrightarrow{1/2} d \xrightarrow{1} g$, where the number of each arrow represents conditional probability. On the other hand, when it returns from node g to a along the reverse trajectory $\vec{\alpha}' = (g, d, b, a)$, one can see $\Pi[\vec{\alpha}';g,a] = 1/3$ along the pathway $g \xrightarrow{1/3} d \xrightarrow{1} b \xrightarrow{1} a$. Thus, the two transition probabilities are not the same: $\Pi[\vec{\alpha};a,g] \neq \Pi[\vec{\alpha}';g,a]$. Further, $\rho(a) = \rho(g) = 1/N$, and $\rho(j|i) = \rho(i|j) = 1/(N-1)$, which yield $\Delta S[\vec{\alpha}] = \ln(3/4)$ by the definition Eq. (1). The EPs along the pathways $\vec{\beta}$ and $\vec{\gamma}$ can be similarly calculated and are listed in Table I.

TABLE I. Probability that a packet takes each shortest pathway and corresponding entropy production. Π denotes the transition probability.

Pathway	Pathway from s to t	Π	Pathway from t to s	Π	ΔS
$\vec{\alpha}$	$a \xrightarrow{1/2} b \xrightarrow{1/2} d \xrightarrow{1} g$	$\frac{1}{4}$	$g \xrightarrow{1/3} d \xrightarrow{1} b \xrightarrow{1} a$	$\frac{1}{3}$	$\ln\frac{3}{4}$
$\vec{\beta}$	$a \xrightarrow{1/2} b \xrightarrow{1/2} e \xrightarrow{1} g$	$\frac{1}{4}$	$g \xrightarrow{1/3} e \xrightarrow{1} b \xrightarrow{1} a$	$\frac{1}{3}$	$\ln\frac{3}{4}$
$\vec{\gamma}$	$a \xrightarrow{1/2} c \xrightarrow{1} f \xrightarrow{1} g$	$\frac{1}{2}$	$g \xrightarrow{1/3} f \xrightarrow{1} c \xrightarrow{1} a$	$\frac{1}{3}$	$\ln\frac{3}{2}$

One can easily find that the transition probability is normalized as $\Pi[\vec{\alpha};a,g] + \Pi[\vec{\beta};a,g] + \Pi[\vec{\gamma};a,g] = 1$. Therefore, for all possible shortest pathways of $N(N-1)$ pairs on the complex network, the probability $P[\vec{\alpha}]$ is also normalized, $\sum_{\vec{\alpha}} P[\vec{\alpha}] = 1$.

III. FLUCTUATION THEOREMS

Here, we obtain the EPD over all possible shortest pathways between every pair of nodes. The EPD $P(\Delta S)$ is given by

$$\begin{aligned} P(\Delta S) &= \sum_{\vec{\alpha}}\sum_{i}\sum_{j\neq i} \delta(\Delta S - \Delta S[\vec{\alpha}])\rho(i)\rho(j|i)\,\Pi[\vec{\alpha};i,j] \\ &= \sum_{\vec{\alpha}'}\sum_{j}\sum_{i\neq j} \delta(\Delta S - \Delta S[\vec{\alpha}])\rho(j)\rho(i|j)\,\Pi[\vec{\alpha}';j,i]e^{\Delta S[\vec{\alpha}]} \\ &= \sum_{\vec{\alpha}'}\sum_{j}\sum_{i\neq j} \delta(\Delta S + \Delta S[\vec{\alpha}'])\rho(j)\rho(i|j)\,\Pi[\vec{\alpha}';j,i]e^{-\Delta S[\vec{\alpha}']} \\ &= P(-\Delta S)e^{\Delta S}, \end{aligned} \quad (2)$$

where we have used the fact that $\Delta S[\vec{\alpha}'] = -\Delta S[\vec{\alpha}]$ and $\sum_{\vec{\alpha}}$ can be replaced by $\sum_{\vec{\alpha}'}$ because the Jacobian is 1. We note that even though $\Delta S[\vec{\alpha}'] = -\Delta S[\vec{\alpha}]$, the EPD is asymmetric. This asymmetry is caused by the definition of the EPD itself. For instance, when a data packet travels from node i to j along the path $\vec{\alpha}$ and then travels along the corresponding reverse path $\vec{\alpha}'$, $\Delta S[\vec{\alpha}]$ is generated for the path $\vec{\alpha}$, whereas $\Delta S[\vec{\alpha}'] = -\Delta S[\vec{\alpha}]$ for the reverse path $\vec{\alpha}'$. These two EPs contribute to the EPD by their weights $P[\vec{\alpha}]$ and $P[\vec{\alpha}']$, respectively. Accordingly, even though $\Delta S[\vec{\alpha}] = -\Delta S[\vec{\alpha}']$, because $\Pi[\vec{\alpha};i,j] \neq \Pi[\vec{\alpha}';j,i]$, $P(\Delta S) \neq P(-\Delta S)$, but $P(\Delta S) = P(-\Delta S)e^{\Delta S}$.

The relation $P(\Delta S) = P(-\Delta S)e^{\Delta S}$ is known as the detailed FT and is an instance of the Gallavotti–Cohen symmetry of the

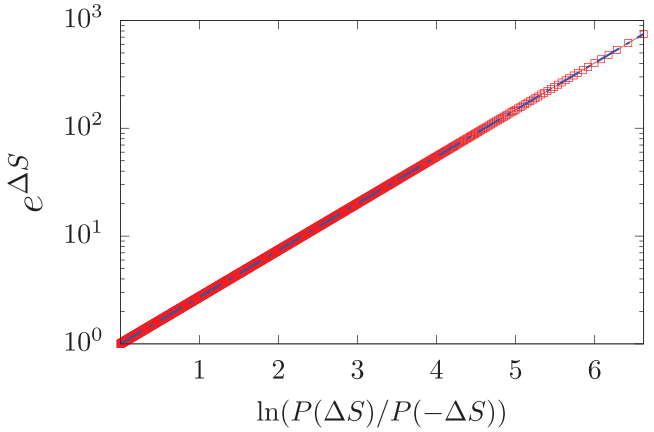

FIG. 2. Plot of the detailed FT. The ratio of the EPDs in the forward and backward directions is equal to $e^{\Delta S}$. Numerical data are obtained from the Barabási–Albert (BA) model networks introduced in Appendix on which a data packet is sent and returned between every pair of nodes along the shortest pathways. Data points lie exactly on the straight line with slope one.

probability density function.[12] We confirm the detailed FT numerically in Fig. 2. From Eq. (2), the integral FT is derived as $\langle e^{-\Delta S}\rangle = \sum_{\Delta S} e^{-\Delta S} P(\Delta S) = 1$.

IV. SCALING OF THE EP DISTRIBUTION

We perform numerical simulations to obtain EPs based on transport along every shortest pathway between all possible pairs of nodes on several networks: the BA model,[26] the ER model,[27] and the Chung–Lu (CL) model[28] with the degree exponents $\gamma = 2.2$ and $\gamma = 2.5$. We will explain these three model networks in Appendix. The EPDs $P(\Delta S)$ on these networks are shown in Fig. 3(a). All these networks were constructed with the same mean degree $\langle k\rangle = 8$ and system size $N = 2^{11} \times 10$. We obtain these EPDs on the giant component of each network. The EPDs have different shapes. The width of the EPD on the BA model is generally wide, whereas that on the ER model is generally narrow. This result arises from the extent of the entanglement of the shortest pathways for the ER and SF networks.

In statistical mechanics, the entropy is an extensive quantity with respect to the system size N. However, in this problem, the length d_{st} of each pathway plays a role similar to that of N in Euclidean space. Thus, we rescale the EP ΔS by the path length and define $\Delta S/d_{st}$. The EPDs obtained for different network sizes N collapse onto a single curve, as shown in Figs. 3(b) and 3(c).

V. ASYMMETRIC EP DISTRIBUTION AND TESTING DFT

It is interesting to find the functional form of the asymptotic EPD, called asymptote, because the mean entropy production is related to the large deviation function.[29] We consider a dataset composed of $N_{\rm EP}$ EPs obtained from all possible shortest pathways between every pair of nodes of a given network such as the BA model, for instance. Next, we consider the cumulative distribution of $P(\Delta S)$, i.e., $F(x) = \text{Prob}\{\Delta S \le x\}$, that is, $F(x) = \int_{-\infty}^{x} d\Delta S P(\Delta S)$. In general, to find functional form of a cumulative distribution in asymptotic limit, one may often use the extreme value statistics (EVS), which is valid for the distribution composed of independent and identically distributed (i.i.d.) random variables. It is known that there exist three types of functional forms: (i) Gumbel type, (ii) Fréchet type, and (iii) Weibull type. On the other hand, the EPs we consider here are not i.i.d., but correlated as they satisfy the FTs. Thus, the EVS method is not justified. Nevertheless, we check if the EPD may fit to one of these EVS categories to obtain some hint for the functional form of the EPD.

We plot the cumulative distribution $F(\Delta S)$ in the forms of $\ln(-\ln F(\Delta S))$ for $\Delta S > 0$ and $\ln(-\ln(\tilde{F}(\Delta S)))$ for $\Delta S < 0$, where $\tilde{F}(\Delta S) = 1 - F(\Delta S)$, as a function of ΔS in Fig. 4, motivated by the Gumbel-type distribution. In these figures, we used the dataset of EPs collected from 300 different network configurations for each model network of the same network size N to reduce the fluctuations arising from different network topologies. Figures 4(b), 4(d), 4(f), and 4(h) show that the data of $\ln(-\ln F(\Delta S))$ excluding the tail part decrease with increasing ΔS along a straight line for the case $\Delta S > 0$ with some wiggling behavior, but the data in the tail part decay more rapidly. We remark that the region of the straight line becomes wider as the system size is larger. On the other hand, the rapidly-decay

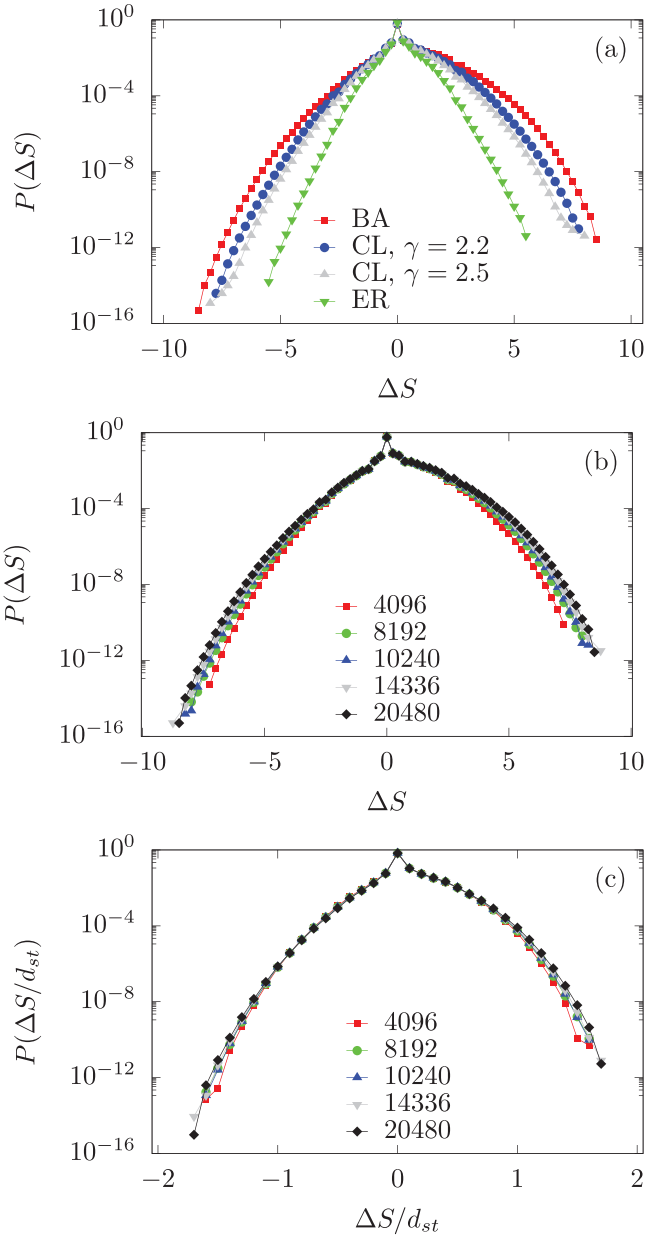

FIG. 3. (a) EPDs on the four model networks: BA model, scale-free CL model with degree exponent $\gamma = 2.2$, scale-free CL model with degree exponent $\gamma = 2.5$, and ER model, from top to bottom. Data are obtained from the giant component of each model network of system size $N = 2^{11} \times 10 \approx 2 \times 10^4$ and mean degree $\langle k\rangle = 8$. They are averaged over 300 network configurations. All EPDs exhibit peaks at $\Delta S = 0$, which are attributed to transport along untangled pathways. (b) EPDs on BA networks of different system sizes, $N = 4096$, 8192, 10 240, 14 336, and 20 480. As N is increased, the EP curves tend to converge to the asymptotic one. (c) Distributions of EPs divided by the Hamming distance d_{st} between a source (s) and a target (t) for each pathway, that is, $\Delta S/d_{st}$. The system sizes are the same as those in (b). The data for the different system sizes collapse onto a single curve.

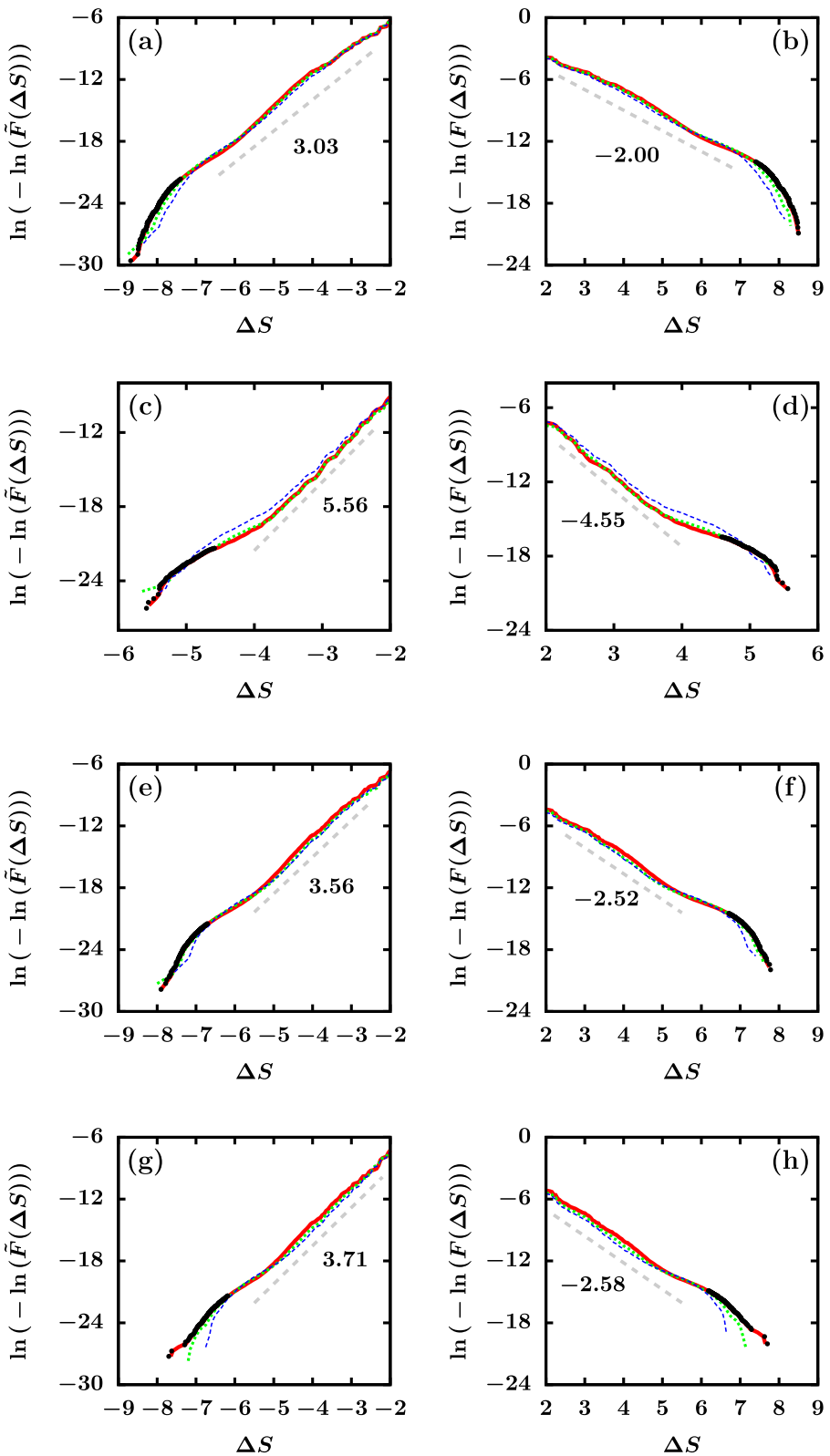

FIG. 4. Plots of the cumulative EPDs in the forms of $\ln(-\ln\tilde{F}(\Delta S))$ for $\Delta S < 0$ and $\ln(-\ln F(\Delta S))$ for $\Delta S > 0$ as a function of ΔS for several model networks: (a) and (b) BA model; (c) and (d) ER model; (e) and (f) CL scale-free network model with degree exponent $\gamma = 2.2$; and (g) and (h) $\gamma = 2.5$. The data points are obtained from different system sizes $N = 10240$ (blue), 14336 (green), and 20480 (red) for each model network. For each given N, the data points are obtained from 300 different network configurations. The data points denoted by symbol (•) in the tail part are collected from the extrema ΔS of each network configuration for $N = 20\,480$.

region becomes wider as the number of network configurations is increased. To unveil the origin of the rapid decay for $\Delta S > 0$, we pick up the maxima of ΔS from each network configuration for each network model of the same size, and they are identified in each plot of Fig. 4 and denoted by symbol (•). We find that these maxima cover the tail part, suggesting that the tail part is mainly composed of the maxima of ΔS of each network configuration. Because the fraction of those maxima is $O(1/N_{\rm EP})$ for each network configuration, we may ignore this contribution in our discussion about checking the DFT using the EPD.

The cumulative distribution may be regarded as

$$\ln(-\ln F(\Delta S)) \approx -a\Delta S + b \quad \text{for} \quad \Delta S > 0, \tag{3}$$

$$\ln(-\ln \tilde{F}(\Delta S)) \approx a'\Delta S + b' \quad \text{for} \quad \Delta S < 0, \tag{4}$$

where a and a' are positive constants and b and b' are constants. The constants (a, a') are estimated to be $(2.00, 3.04)$ for the BA model, $(4.54, 5.56)$ for the ER model, $(2.52, 3.56)$ for the CL model with degree exponent 2.2, and $(2.58, 3.71)$ for the CL model with degree exponent 2.5 in Fig. 4.

Next, we derive the functional form of $P(\Delta S)$. From $F(\Delta S) \approx \exp(-e^{-a\Delta S+b})$ for $\Delta S > 0$, the EPD is derived as

$$P(\Delta S) = \left.\frac{dF(x)}{dx}\right|_{x=\Delta S} \approx a\, \exp(-e^{-ax+b})e^{-ax+b}|_{x=\Delta S} \approx a\exp(-a\Delta S + b) \tag{5}$$

for sufficiently large $y \equiv a\Delta S - b \gg e^{-y}$. On the other hand, for $\Delta S < 0$, $1 - F(\Delta S) = \exp(-e^{a'\Delta S+b'})$. Then, similarly,

$$P(\Delta S) = -\left.\frac{dF(x)}{dx}\right|_{x=\Delta S} \approx a'\exp(-e^{a'x+b'})e^{a'x+b'}\Big|_{x=\Delta S} \approx a'\exp(a'\Delta S + b') \tag{6}$$

for sufficiently negatively large ΔS. Therefore,

$$\frac{P(\Delta S)}{P(-\Delta S)} \approx \frac{ae^{b}}{a'e^{b'}}\exp(-(a-a')\Delta S). \tag{7}$$

Here, we recall that the difference $a - a'$ of the two constants is very roughly close to minus one regardless of the model networks. Thus, we may say that formula (7) supports the DFT.

VI. SUMMARY AND DISCUSSION

In this paper, we considered the EPD arising from the complexity of the shortest pathways from one node to another on complex networks. We verified explicitly that this EPD satisfies well-known FTs, i.e., the integral FT and the detailed FT. To obtain the result, we considered a data packet transport problem in which a packet travels back and forth between every pair of nodes along each of the shortest pathways. At a branching node along the way, a packet chooses one branch randomly. The effect of this random choice reflects the stochastic process in dynamics in nonequilibrium systems. Owing to the complexity of the shortest pathways, the probabilities of taking a shortest pathway in one direction and the corresponding reverse direction can be different, resulting in a nonzero EP. We calculated this difference explicitly and found the functional form of the cumulative EPD in the large-EP limit in positive and negative regions using direct measurement. The cumulative EPD behaves as $F(\Delta S) \approx \exp(-e^{-a\Delta S+b})$ in the positive region and $1 - F(\Delta S) \sim \exp(-e^{a'\Delta S+b'})$ in the negative region, where a and a' are positive constants and b and b' are constants. The constants depend on networks and differ from each other. The ratio $P(\Delta S)/P(-\Delta S)$ scales as $\exp(-(a - a')\Delta S)$ asymptotically. We find that the numerical difference $a - a'$ is roughly -1 regardless of the network models, which supports the detailed FT.

In the stochastic thermodynamics, the fluctuation theorems were derived from the total EP that is based on the trajectory-dependent EP in the stochastic process. However, because the trajectories of the stochastic dynamic process are rather virtual, one can hardly imagine the origin of the total EP and understand the origin of the fluctuation theorems. In this study, one can identify the reverse trajectory easily and can calculate the total EP explicitly. Thus, our result would be pedagogically helpful not only for understanding the concept of trajectory-dependent EP in stochastic processes, but also for exploring nonequilibrium fluctuations related to the entanglement of dividing and merging among the shortest pathways of complex networks.

Note added in Proof. We became aware of a paper[30] from an anonymous referee, which considered the average EP of stochastic processes between two states on random networks. Although the paper[30] and ours might start from a similar motivation, they differ from each other in the aspects of aims and detailed dynamics.

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea under Grant Nos. 2014R1A3A2069005 (B.K.), 2017R1D1A1B03030872 (J.U.), and 2017R1A6A3A11033971 (D.L.); the National Research Foundation of Korea; the Brain Pool program under Grant No. 171S-1-2-1882 (Y.W.K.); and the IBS under Grant No. IBS-R020-D1 (H.K.P.). B.K. thanks C. Kwon and J. D. Noh for useful discussion.

APPENDIX: NETWORK MODELS

(i) A BA network is constructed as follows: At the beginning, there exist m_0 nodes in the system. At each time step, a node is added with m links in the system, where $m \le m_0$. Each link is connected to a node i with degree k_i with the probability $p(k_i) = k_i/\sum_j k_j$. This process is repeated until the total number of nodes in the system becomes N. This model generates a scale-free network with the degree distribution $P_d(k) \sim k^{-\gamma}$ with $\gamma = 3$.[26]

(ii) A CL network is constructed as follows: At the beginning, there exist a fixed number of N nodes indexed $i = 1, \ldots, N$ in the system. Then, a node i is assigned a weight of $w_i = (i + i_0 - 1)^{-\mu}$, where $\mu \in [0, 1)$ is a control parameter and $i_0 \propto N^{1-1/2\mu}$ for $1/2 < \mu < 1$ and $i_0 = 1$ for $\mu < 1/2$. Then, two different nodes (i, j) are selected with their probabilities equal to the normalized weights, $w_i/\sum_k w_k$ and $w_j/\sum_k w_k$, respectively, and a link is added between them unless one already exists. This process is repeated until pN links are created in the system, where p is a control parameter. There exists a percolation

threshold p_c, above which a macroscopic-scale large cluster is generated. We considered the data packet transport problem on such large networks. The obtained network is scale-free in degree distribution with the exponent $\gamma = 1 + 1/\mu$.

(iii) An Erdős–Rényi network is constructed as follows: At the beginning, there exist a fixed number of N vertices in the system. At each time step, two nodes are selected randomly. They are connected with a link unless they are already connected. This process is repeated until pN links are created in the system, where p is a control parameter. It is known that $p_c = 1/2$ is the percolation threshold. Thus, for $p > p_c$, a macroscopic-scale large cluster is generated. We considered the data packet transport problem on such large networks. The obtained network has the degree distribution following a Poisson distribution.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).
[2] P. Strasberg, G. Schaller, T. Brandes, and M. Esposito, Phys. Rev. X **7**, 021003 (2017).
[3] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, Nat. Phys. **11**, 131 (2015).
[4] H. Park, J. Korean Phys. Soc. **72**, 1413 (2018).
[5] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).
[6] M. Esposito and C. Van den Broeck, Phys. Rev. Lett. **104**, 090601 (2010).
[7] J. Kurchan, J. Phys. A **31**, 3719 (1998).
[8] G. E. Crooks, J. Stat. Phys. **90**, 1481 (1998).
[9] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[10] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).
[11] D. J. Evans and D. J. Searles, Phys. Rev. E **50**, 1645 (1994).
[12] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995).
[13] J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95**, 333 (1999).
[14] N. Garnier and S. Ciliberto, Phys. Rev. E **71**, 060101 (2005).
[15] S. Ciliberto, S. Joubaud, and A. Petrosyan, J. Stat. Mech. **2010**, P12003 (2010).
[16] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, Nature **437**, 231 (2005).
[17] S. Schuler, T. Speck, C. Tietz, J. Wrachtrup, and U. Seifert, Phys. Rev. Lett. **94**, 180602 (2005).
[18] N. Garnier and S. Ciliberto, Phys. Rev. E **71**, 060101(R) (2005).
[19] F. Rao and A. Caflisch, J. Mol. Biol. **342**(1), 299 (2004).
[20] D. Gfeller, P. De Los Rios, A. Caflisch, and F. Rao, Proc. Natl. Acad. Sci. U.S.A. **104**, 1817 (2007).
[21] C. Levinthal, in *Mossbauer Spectroscopy in Biological Systems*, edited by J. T. P. De Brunner and E. Munck (University of Illinois Press, Monticello, 1969).
[22] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. **87**, 278701 (2001).
[23] M. E. J. Newman, Phys. Rev. E **64**, 016131 (2001).
[24] M. E. J. Newman, Phys. Rev. E **64**, 016132 (2001).
[25] E. W. Dijkstra, Num. Math. **1**, 269 (1959).
[26] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[27] P. Erdős and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).
[28] F. Chung and L. Lu, Ann. Combin. **6**, 125 (2002).
[29] J. Mehl, T. Speck, and U. Seifert, Phys. Rev. E **78**, 011123 (2008).
[30] D. M. Busiello, J. Hidalgo, and A. Maritan, Phys. Rev. E **96**, 06210 (2017).