Analysis of the convergence of the degree distribution of contracting random networks towards a Poisson distribution using the relative entropy

Ido Tishby, Ofer Biham, and Eytan Katzav
Racah Institute of Physics, The Hebrew University, Jerusalem 9190401, Israel

We present analytical results for the structural evolution of random networks undergoing contraction processes via generic node deletion scenarios, namely, random deletion, preferential deletion, and propagating deletion. Focusing on configuration model networks, which exhibit a given degree distribution $P_0(k)$ and no correlations, we show using a rigorous argument that upon contraction the degree distributions of these networks converge towards a Poisson distribution. To this end, we use the relative entropy $S_t = S[P_t(k)||\pi(k|\langle K \rangle_t)]$ of the degree distribution $P_t(k)$ of the contracting network at time t with respect to the corresponding Poisson distribution $\pi(k|\langle K \rangle_t)$ with the same mean degree $\langle K \rangle_t$ as a distance measure between $P_t(k)$ and Poisson. The relative entropy is suitable as a distance measure since it satisfies $S_t \geq 0$ for any degree distribution $P_t(k)$, while equality is obtained only for $P_t(k) = \pi(k|\langle K \rangle_t)$. We derive an equation for the time derivative dS_t/dt during network contraction and show that the relative entropy decreases monotonically to zero during the contraction process. We thus conclude that the degree distributions of contracting configuration model networks converge towards a Poisson distribution. Since the contracting networks remain uncorrelated, this means that their structures converge towards an Erdős-Rényi (ER) graph structure, substantiating earlier results obtained using direct integration of the master equation and computer simulations [Tishby *et al.*, Phys. Rev. E **100**, 032314 (2019)]. We demonstrate the convergence for configuration model networks with degenerate degree distributions (random regular graphs), exponential degree distributions, and power-law degree distributions (scale-free networks).

I. INTRODUCTION

Complex network architectures and dynamical processes taking place on them play a central role in current research [1–3]. Since the 1960s, mathematical studies of networks were focused on model systems such as the Erdős-Rényi (ER) network [4–6], which exhibits a Poisson degree distribution of the form $\pi(k|c) = e^{-c}c^k/k!$, where c is the mean degree [7]. In an ER network of N nodes, each pair of nodes is connected with probability p, where $p = c/(N - 1)$. In fact, ER networks form a maximum entropy ensemble under the constraint that the mean degree is fixed [8–11]. In the 1990s, the growing availability of data on large biological, social, and technological networks revolutionized the field. Motivated by the observation that the World Wide Web [12] and scientific citation networks [13] exhibit power-law degree distributions, Barabási and Albert (BA) introduced a simple model that captures the essential growth dynamics of such networks [14,15]. A key feature of the BA model is the preferential attachment mechanism, namely, the tendency of new nodes to attach preferentially to high degree nodes. Using mean-field equations and computer simulations it was shown that the combination of growth and preferential attachment leads to the emergence of scale-free networks with power-law degree distributions [15]. This result was later confirmed and generalized using a more rigorous formulation based on the master equation [16,17]. It was subsequently found that a large variety of empirical networks exhibit such scale-free

structures, which are remarkably different from ER networks [14,18].

In many of these networks the growth phase is not likely to proceed indefinitely. Moreover, networks may be exposed to node deletion processes due to node failures, attacks, and epidemics, which may eventually halt the expansion phase and induce the contraction and eventual collapse of the network. Since network growth is a kinetic nonequilibrium process, it is not a reversible process, namely, the contraction process is not the same as the growth process when played backwards in time. A particularly interesting example of the contraction phase can be seen in the field of social networks. Such networks may lose users due to loss of interest, concerns about privacy, or their migration to other social networks [19,20]. Another example of great practical importance is the cascading failure of power grids [21,22]. Infectious processes such as epidemics that spread in a network [23,24] lead to the contraction of the subnetwork of uninfected nodes and may thus be considered as network contraction processes. Similarly, network immunization schemes [25] also belong to the class of network contraction processes because they induce the contraction of the subnetwork of susceptible nodes.

Three generic scenarios of network contraction were identified: the scenario of random node deletion that describes the random, inadvertent failure of nodes, the scenario of preferential node deletion that describes intentional attacks that are more likely to focus on highly connected nodes, and the scenario of propagating node deletion that describes

viral and infectious processes that spread like epidemics. It was found that scale-free networks are resilient to attacks targeting random nodes, but are vulnerable to attacks that target high-degree nodes or hubs. Using the framework of percolation theory, it was shown that when the number of deleted nodes exceeds some threshold, the network breaks down into disconnected components [26–30]. However, the evolution of the network structure throughout the contraction phase was not addressed.

In a recent paper we analyzed the structural evolution of networks during the contraction process [31]. To this end we derived a master equation for the time dependence of the degree distribution during network contraction via the random deletion, preferential deletion, and propagating deletion scenarios. Using the relative entropy and the degree-degree correlation function we showed that the ER graph structure, which exhibits a Poisson degree distribution, is an asymptotic structure for these network collapse scenarios, in analogy to the way in which the scale-free structure is an asymptotic solution for the preferential attachment growth scenario.

In this paper we use the relative entropy to provide a rigorous proof that the ER structure is an attractive solution for the three contraction scenarios. This means that the ER structure is a universal asymptotic structure for contracting networks. For simplicity, we consider initial networks drawn from configuration model network ensembles that exhibit a desired degree distribution $P_0(k)$ and no degree-degree correlations. During the contraction process the degree distribution of the network evolves. We denote the degree distribution at time t by $P_t(k)$ and its mean degree by $\langle K \rangle_t$. We use the relative entropy $S_t = S[P_t(k)||\pi(k|\langle K \rangle_t)]$ as a distance measure between the degree distribution $P_t(k)$ of the contracting network and the corresponding Poisson distribution $\pi(k|\langle K \rangle_t)$ with the same mean degree $\langle K \rangle_t$. Using this measure we obtain rigorous results for the convergence of the degree distribution of contracting networks towards a Poisson distribution. To this end, we derive an equation for the time derivative dS_t/dt of the relative entropy during network contraction. This equation can be expressed in the form $dS_t/dt = \Delta_A(t) + \Delta_B(t)$. We show that $\Delta_A(t) < 0$ for any degree distribution. We also show that $\Delta_B(t) < 0$ for degree distributions whose tails decay more slowly than the tail of the Poisson distribution with the same mean degree. This condition is generically satisfied by the heavy-tail distributions that emerge from network growth processes. In contrast, in networks that exhibit narrow degree distributions the $\Delta_B(t)$ term turns out to be small and has little effect on the convergence, which is dominated by $\Delta_A(t)$. This implies that the relative entropy decreases monotonically during the contraction process. Since the relative entropy satisfies $S_t \geq 0$ for any degree distribution $P_t(k)$, while equality is obtained only for $P_t(k) = \pi(k|\langle K \rangle_t)$ we conclude that the degree distributions of contracting networks converge towards a Poisson distribution. This conclusion is corroborated by the fact that the relative entropy provides an upper bound for the total variation distance, which is a standard measure of the difference between probability distributions. We demonstrate the convergence for configuration model networks with a degenerate degree distribution (random regular graphs), exponential degree distribution, and power-law degree distribution (scale-free networks).

The paper is organized as follows. In Sec. II we present the three generic network contraction scenarios studied in this paper. In Sec. III we present the master equation and show that the Poisson distribution is a solution of the master equation for the three contraction scenarios. In Sec. IV we present the relative entropy and express it in terms of the Shannon entropy and the cross-entropy. In Sec. V we present rigorous results showing that the relative entropy decays to zero in any of the three contraction scenarios. In Sec. VI we present analytical results and computer simulations for the contraction of configuration model networks with a degenerate degree distribution (random regular graphs), an exponential degree distribution, and a power-law degree distribution (scale-free networks). The results are discussed in Sec. VII and summarized in Sec. VIII.

II. NETWORK CONTRACTION PROCESSES

We consider network contraction processes in which at each time step a single node is deleted together with its links. The initial network consists of N_0 nodes, so at time t the network size is reduced to $N_t = N_0 - t$ nodes. The deletion of a node of degree k, whose neighbors are of degrees k_i', $i = 1, 2, \ldots, k$, eliminates the deleted node from the degree sequence and reduces the degrees of its neighbors to $k_i' - 1$, $i = 1, 2, \ldots, k$. The node deleted at each time step is selected randomly. However, the probability of a node to be selected for deletion may depend on its degree, according to the specific network contraction scenario. Here we focus on three generic scenarios of network contraction: the scenario of random node deletion that describes the random, inadvertent failure of nodes, the scenario of preferential node deletion that describes intentional attacks that are more likely to focus on highly connected nodes, and the scenario of propagating node deletion that describes cascading failures and infectious processes that spread throughout the network.

In the random deletion scenario, at each time step a random node is selected for deletion. In this scenario each one of the nodes in the network at time t has the same probability to be selected for deletion, regardless of its degree. Since at time t there are N_t nodes in the network, the probability of each one of them to be selected for deletion is $1/N_t$. In the preferential deletion scenario the probability of a node to be selected for deletion at time t is proportional to its degree at that specific time. This means that the probability of a given node of degree k to be deleted at time t is $k/[N_t\langle K \rangle_t]$. This is equivalent to selecting a random edge in the network and randomly choosing for deletion one of the two nodes at its ends. In the propagating deletion scenario at each time step the node to be deleted is randomly selected among the neighbors of the node deleted in the previous time step. In the case that the node deleted in the previous time step does not have any yet-undeleted neighbor we pick a random node, randomly select one of its neighbors for deletion, and continue the process from there.

Here we focus on the contraction of undirected networks of initial size N, which are drawn from a configuration model network ensemble with a given initial degree distribution $P_0(k)$ and no degree-degree correlations. The degree distribution is bounded from above and below such that

$k_{\min} \leqslant k \leqslant k_{\max}$. For example, the commonly used choice of $k_{\min} = 1$ eliminates the possibility of isolated nodes in the network. Choosing $k_{\min} = 2$ also eliminates the leaf nodes. Controlling the upper bound is important in the case of fat-tail degree distributions such as power-law degree distributions. The configuration model network ensemble is a maximum entropy ensemble under the condition that the degree distribution $P(k)$ is imposed [32–37]. In such uncorrelated networks the deletion of a node at time t does not induce correlations between the remaining $N_t - 1$ nodes. Thus, upon deletion of a node from a configuration model network of size N_t, the resulting network remains a configuration model network with a suitably adjusted degree distribution $P_{t+1}(k)$.

III. THE MASTER EQUATION AND ITS POISSON SOLUTION

Consider an ensemble of networks of size N_0 and degree distribution $P_0(k)$, with mean degree $\langle K \rangle_0$. At each time step a single node is deleted from the network. In addition to the primary effect of the loss of the deleted node, the damage to the network also includes a secondary effect as each neighbor of the deleted node loses one link. An intrinsic property of the secondary effect is that it is always of a preferential nature. This is due to the fact that the probability of a node of degree k' to be a neighbor of the deleted node is proportional to k'. The number of nodes in the network at time t is $N_t = N_0 - t$. The number of nodes of degree k at time t is denoted by $N_t(k)$, where $\sum_k N_t(k) = N_t$. The time-dependent degree distribution is given by

$$P_t(k) = \frac{N_t(k)}{N_t}. \tag{1}$$

The mean degree and the second moment of the degree distribution at time t are denoted by $\langle K^n \rangle_t$ where $n = 1$ and 2, respectively.

The master equation [38,39] for the temporal evolution of the degree distribution $P_t(k)$ during network contraction processes was derived in Ref. [31]. To demonstrate the derivation of the master equation we consider below the relatively simple case of random node deletion. The time dependence of $N_t(k)$ depends on the primary effect, given by the probability that the node selected for deletion is of degree k as well as on the secondary effect of node deletion on neighboring nodes of degrees k and $k + 1$. In random node deletion the probability that the node selected for deletion at time t is of degree k is given by $N_t(k)/N_t$. Thus, the rate at which $N_t(k)$ decreases due to the primary effect of the deletion of nodes of degree k is given by

$$R_t(k \rightarrow \varnothing) = \frac{N_t(k)}{N_t}, \tag{2}$$

where $\varnothing$ represents the empty set. In case that the node deleted at time t is of degree k', it affects k' adjacent nodes, which lose one link each. The probability of each one of these k' nodes to be of degree k is given by $kN_t(k)/[N_t\langle K \rangle_t]$. We denote by $W_t(k \rightarrow k - 1)$ the expectation value of the number of nodes of degree k that lose a link at time t and are reduced to degree $k - 1$. Summing up over all possible values of k', we find that the secondary effect of random node deletion on nodes

of degree k amounts to

$$W_t(k \rightarrow k - 1) = \frac{kN_t(k)}{N_t}. \tag{3}$$

Similarly, the secondary effect on nodes of degree $k + 1$ amounts to

$$W_t(k + 1 \rightarrow k) = \frac{(k + 1)N_t(k + 1)}{N_t}. \tag{4}$$

The time evolution of $N_t(k)$ can be expressed in terms of the forward difference

$$\Delta_t N_t(k) = N_{t+1}(k) - N_t(k). \tag{5}$$

Combining the primary and the secondary effects on the time dependence of $N_t(k)$ we obtain

$$\Delta_t N_t(k) = -R_t(k \rightarrow \varnothing) + [W_t(k + 1 \rightarrow k) - W_t(k \rightarrow k - 1)]. \tag{6}$$

Since nodes are discrete entities the process of node deletion is intrinsically discrete. Therefore, the replacement of the forward difference $\Delta_t N_t(k)$ by a time derivative of the form $dN_t(k)/dt$ involves an approximation. The error associated with this approximation was evaluated in Ref. [31]. It was shown that except for the limit of extremely narrow degree distributions the error is of order $1/N_t^2$, which quickly vanishes in the large network limit. This means that the replacement of the forward difference by a time derivative has little effect on the results and a clear technical advantage.

Inserting the expressions for $R_t(k \rightarrow \varnothing)$, $W_t(k \rightarrow k - 1)$, and $W_t(k + 1 \rightarrow k)$ from Eqs. (2), (3), and (4), respectively, into Eq. (6) and replacing $\Delta_t N_t(k)$ by $dN_t(k)/dt$ we obtain

$$\frac{d}{dt}N_t(k) = \frac{(k + 1)[N_t(k + 1) - N_t(k)]}{N_t}. \tag{7}$$

The derivation of the master equation is completed by taking the time derivative of Eq. (1), which is given by

$$\frac{d}{dt}P_t(k) = \frac{1}{N_t}\frac{d}{dt}N_t(k) - \frac{N_t(k)}{N_t^2}\frac{d}{dt}N_t. \tag{8}$$

Inserting the time derivative of $N_t(k)$ from Eq. (7) into Eq. (8) and using the fact that $dN_t/dt = -1$, we obtain the master equation for the random deletion scenario, which is given by

$$\frac{d}{dt}P_t(k) = \frac{1}{N_t}[(k + 1)P_t(k + 1) - kP_t(k)]. \tag{9}$$

The derivation of the master equations for the preferential deletion and the propagating deletion scenarios can be performed along similar lines [31]. Interestingly, the resulting master equations for these three network contraction scenarios can be written in a unified manner, in the form

$$\frac{d}{dt}P_t(k) = F_A(t) + F_B(t), \tag{10}$$

where

$$F_A(t) = \frac{A_t}{N_t}[(k + 1)P_t(k + 1) - kP_t(k)] \tag{11}$$

accounts for the secondary effect on the neighbors of the deleted node, which lose one link each, while

$$F_{\mathrm{B}}(t) = -\frac{B_t(k)}{N_t} P_t(k) \qquad (12)$$

accounts for the primary effect, namely, the loss of the deleted node [31]. The coefficients A_t and $B_t(k)$ are given by

$$A_t = \begin{cases} 1 & \text{random deletion} \\ \frac{\langle K^2 \rangle_t}{\langle K \rangle_t^2} & \text{preferential deletion} \\ \frac{\langle K^2 \rangle_t - 2\langle K \rangle_t}{\langle K \rangle_t^2} & \text{propagating deletion} \end{cases} \qquad (13)$$

and

$$B_t(k) = \begin{cases} 0 & \text{random deletion} \\ \frac{k - \langle K \rangle_t}{\langle K \rangle_t} & \text{preferential deletion} \\ \frac{k - \langle K \rangle_t}{\langle K \rangle_t} & \text{propagating deletion} \end{cases} . \qquad (14)$$

The master equation consists of a set of coupled ordinary differential equations for $P_t(k)$, $k = 0, 1, 2, \dots, k_{\max}$, or in other words it is a partial difference-differential equation. In order to calculate the time evolution of the degree distribution $P_t(k)$ during the contraction process one solves the master equation using direct numerical integration [40], starting from the initial network that consists of N_0 nodes whose degree distribution is $P_0(k)$. For any finite network the degree distribution is bounded from above by an upper bound denoted by $k_{\max}$, which satisfies the condition $k_{\max} \leqslant N_0 - 1$. Since the contraction process can delete edges only from the remaining nodes and cannot increase the degree of any node, the upper cutoff $k_{\max}$ is maintained throughout the contraction process.

The $F_{\mathrm{A}}(t)$ term of the master equation, given by Eq. (11), is referred to as the trickle-down term [41]. This term represents the step by step downwards flow of probability from high to low degrees. This process is illustrated in Fig. 1(a). The coefficient A_t of the trickle-down term depends on the network contraction scenario according to Eq. (13). In the case of random node deletion $A_t = 1$, because the probability of a node to be selected for deletion does not depend on its degree. In the case of preferential node deletion A_t is proportional to $\langle K^2 \rangle_t$ because the probability of a node to be deleted is proportional to its degree k while the magnitude of the secondary effect is also proportional to k.

The $F_{\mathrm{B}}(t)$ term of the master equation, given by Eq. (12), is referred to as the redistribution term. As can be seen in Eq. (14), this term vanishes in the random deletion scenario. However, in the preferential and propagating deletion scenarios the redistribution term is negative for $k > \langle K \rangle_t$ and positive for $k < \langle K \rangle_t$. Thus the redistribution term decreases the probabilities $P_t(k)$ for values of k that are above the mean degree and increases them for values of k that are below the mean degree, as illustrated in Fig. 1(b). The size of the redistribution term is proportional to the absolute value $|k - \langle K \rangle_t|$, which means that nodes of degrees that are much higher or much lower than $\langle K \rangle_t$ are most strongly affected by this term.

Consider an ER network of N_t nodes with mean degree c_t. Its degree distribution follows a Poisson distribution of the
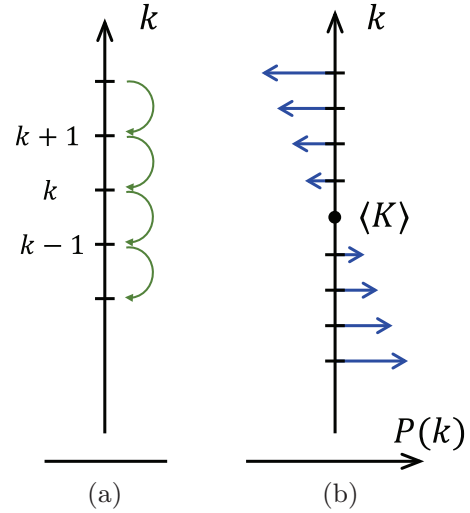

FIG. 1. Illustration of the time dependence of the degree distribution $P_t(k)$ during network contraction processes, described by the master equation (10). (a) In the trickle-down term $F_{\mathrm{A}}(t)$, given by Eq. (11), the probability flows downwards step by step from degree $k + 1$ to k and from k to $k - 1$. This way high-degree nodes become less probable and low-degree nodes become more probable as the contraction process evolves. (b) In the redistribution term $F_{\mathrm{B}}(t)$, given by Eq. (12), for values of k above the mean degree $\langle K \rangle_t$ the probability $P_t(k)$ decreases at a rate proportional to $k - \langle K \rangle_t$, while for values of k below $\langle K \rangle_t$ the probability $P_t(k)$ increases at a rate proportional to $\langle K \rangle_t - k$. Here the flow of probability is nonlocal in the k axis, namely, probability is lost at high degrees and instantaneously emerges at low degrees.

form

$$\pi(k|c_t) = \frac{e^{-c_t} c_t^k}{k!}. \qquad (15)$$

The second moment of this degree distribution is equal to $c_t(c_t + 1)$. To examine the contraction process of ER networks we start from an initial network of N_0 nodes whose degree distribution follows a Poisson distribution $\pi(k|c_0)$, where c_0 is the mean degree of the initial network. Inserting $\pi(k|c_t)$ into the master equation (10) we find that the time derivative on the left-hand side is given by

$$\frac{d}{dt} \pi(k|c_t) = -\frac{dc_t}{dt} \left(1 - \frac{k}{c_t} \right) \pi(k|c_t). \qquad (16)$$

On the other hand, inserting $\pi(k|c_t)$ on the right-hand side of Eq. (10), we obtain

$$\frac{d}{dt} \pi(k|c_t) = \frac{A_t}{N_t} (c_t - k) \pi(k|c_t) - \frac{B_t(k)}{N_t} \pi(k|c_t). \qquad (17)$$

In order that $\pi(k|c_t)$ will be a solution of Eq. (10), the right-hand sides of Eqs. (16) and (17) must coincide. In the case of random deletion this implies that

$$\frac{1}{c_t} \frac{dc_t}{dt} = -\frac{1}{N_t}. \qquad (18)$$

Integrating both sides for $t' = 0$ to t, we obtain the solution $c_t = c_0 N_t / N_0$. Repeating the analysis presented above for the cases of preferential deletion and propagating deletion it is

found that $\pi(k|c_t)$ solves the master equation (10) for the three network contraction scenarios, while the mean degree, c_t decreases linearly in time according to

$$c_t = c_0 - Rt, \tag{19}$$

where the rate R depends on the network contraction scenario and is given by

$$R = \begin{cases} \frac{c_0}{N_0} & \text{random deletion} \\ \frac{c_0+2}{N_0} & \text{preferential deletion} \\ \frac{c_0}{N_0} & \text{propagating deletion} \end{cases} \tag{20}$$

This means that an ER network exposed to any one of the three contraction scenarios remains an ER network at all times, with a mean degree that decreases according to Eq. (19).

IV. THE RELATIVE ENTROPY

In order to establish that networks exposed to these contraction scenarios actually converge towards the ER structure, it remains to show that the Poisson solution is attractive. To quantify the convergence of $P_t(k)$, whose mean degree is $\langle K \rangle_t$, towards a Poisson distribution, we use the relative entropy (also referred to as the Kullback-Leibler divergence), defined by [42]

$$S_t = S[P_t(k)||\pi(k|\langle K \rangle_t)] = \sum_{k=0}^{\infty} P_t(k) \ln\left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)}\right], \tag{21}$$

where $\pi(k|\langle K \rangle_t)$ is the Poisson distribution, given by Eq. (15), with the same mean degree as $P_t(k)$, namely, $\langle K \rangle_t$. The relative entropy S_t is a distance measure between the whole degree distribution $P_t(k)$ and the reference distribution $\pi(k|\langle K \rangle_t)$. It also quantifies the added information associated with constraining the degree distribution $P_t(k)$ rather than only the mean degree $\langle K \rangle_t$, as nicely shown in Refs. [35–37]. The Poisson distribution is a proper reference distribution for the relative entropy because it satisfies $\pi(k|\langle K \rangle_t) > 0$ for all the non-negative integer values of k. Using the log-sum inequality [43], one can show that the relative entropy is always nonnegative and satisfies $S_t = 0$ if and only if $P_t(k) = \pi(k|\langle K \rangle_t)$ [44,45]. Therefore, S_t can be used as a measure of the distance between a given network and the corresponding ER network with the same mean degree.

The relative entropy $S[P(k)||\pi(k|c)]$ of a degree distribution $P(k)$ with mean degree $\langle K \rangle$ with respect to a Poisson distribution $\pi(k|c)$ with mean degree c can be decomposed in the form

$$S[P(k)||\pi(k|c)] = -S[P(k)] + C[P(k)||\pi(k|c)], \tag{22}$$

where

$$S[P(k)] = -\sum_{k=0}^{\infty} P(k) \ln[P(k)] \tag{23}$$

is the Shannon entropy [46] of $P(k)$, while

$$C[P(k)||\pi(k|c)] = -\sum_{k=0}^{\infty} P(k) \ln[\pi(k|c)] \tag{24}$$

is the cross-entropy [47] between $P(k)$ and $\pi(k|c)$. The Poisson distribution $\pi(k|c)$ satisfies

$$\ln[\pi(k|c)] = -c + k \ln(c) - \ln(k!). \tag{25}$$

Inserting $\ln[\pi(k|c)]$ from Eq. (25) into Eq. (24), we obtain

$$S[P(k)||\pi(k|c)] = \sum_{k=0}^{\infty} P(k) \ln[P(k)] + c - \langle K \rangle \ln(c) + \sum_{k=0}^{\infty} \ln(k!) P(k). \tag{26}$$

Equation (26) provides the relative entropy of any degree distribution $P(k)$ whose mean degree is $\langle K \rangle$, with respect to a Poisson distribution with mean degree c. In order to find the value of c for which $S[P(k)||\pi(k|c)]$ is minimal we differentiate $S[P(k)||\pi(k|c)]$ with respect to c and solve the equation

$$\frac{d}{dc} S[P(k)||\pi(k|c)] = 1 - \frac{\langle K \rangle}{c} = 0. \tag{27}$$

We find that $S[P(k)||\pi(k|c)]$ is minimized when the condition $c = \langle K \rangle$ is satisfied. This implies that for any degree distribution $P(k)$ with mean degree $\langle K \rangle$, the closest Poisson distribution $\pi(k|c)$, in terms of the relative entropy, is the Poisson distribution with mean degree $c = \langle K \rangle$.

Using the result discussed above, one can express the relative entropy $S[P(k)||\pi(k|c)]$ in the form

$$S[P(k)||\pi(k|c)] = S[P(k)||\pi(k|\langle K \rangle)] + \delta S(c, \langle K \rangle), \tag{28}$$

where $S[P(k)||\pi(k|\langle K \rangle)]$ is the relative entropy of $P(k)$ with respect to a Poisson distribution whose mean is $\langle K \rangle$, and

$$\delta S(c, \langle K \rangle) = \langle K \rangle \left[\left(\frac{c}{\langle K \rangle} - 1\right) - \ln\left(\frac{c}{\langle K \rangle}\right) \right] \tag{29}$$

is the added entropy due to the difference between c and $\langle K \rangle$. Note that $\delta S(c, \langle K \rangle) \geqslant 0$ for any choice of $\langle K \rangle > 0$ and $c > 0$, while $\delta S(c, \langle K \rangle) = 0$ only in the case that $c = \langle K \rangle$.

Going back to Eq. (22), the relative entropy $S[P(k)||\pi(k|\langle K \rangle)]$ can be expressed in the form

$$S[P(k)||\pi(k|\langle K \rangle)] = -S[P(k)] + C[P(k)||\pi(k|\langle K \rangle)], \tag{30}$$

where $S[P(k)]$ is given by Eq. (23) and

$$C[P(k)||\pi(k|\langle K \rangle)] = \langle K \rangle - \langle K \rangle \ln(\langle K \rangle) + \sum_{k=0}^{\infty} \ln(k!) P(k). \tag{31}$$

To evaluate the last term in Eq. (31) we recall that $\ln(0!) = \ln(1!) = 0$, while the $k = 2$ term is $\ln(2)P(2)$. For $k \geqslant 3$ we use the Stirling approximation [48]

$$\ln(k!) = \left(k + \frac{1}{2}\right) \ln(k) - k + \frac{1}{2} \ln(2\pi). \tag{32}$$

Inserting $\ln(k!)$ for $k \geqslant 3$ from Eq. (32) into Eq. (31) and rearranging terms, we obtain

$$
\begin{aligned}
C[P(k)||\pi(k|c)] = {} & -\langle K \rangle \ln(\langle K \rangle) + \sum_{k=2}^{\infty} \left(k + \frac{1}{2} \right) \ln(k) P(k) \\
& + \frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(2\pi) P(0) \\
& + \left[1 - \frac{1}{2} \ln(2\pi) \right] P(1) \\
& + \left[2 - \frac{3}{2} \ln(2) - \frac{1}{2} \ln(2\pi) \right] P(2), \quad (33)
\end{aligned}
$$

where the terms involving $P(0)$, $P(1)$, and $P(2)$ result from the adjustment of the summation due to the fact that Eq. (32) is used only for $k \geqslant 3$. Note that in the case of distributions in which $k_{\min} \geqslant 1$, one assigns $P(k) = 0$ for $0 \leqslant k \leqslant k_{\min} - 1$. Using Eq. (33), the relative entropy of the degree distribution $P_t(k)$ of a contracting network with respect to the corresponding Poisson distribution $\pi_t(k|\langle K \rangle_t)$ with the same mean degree $\langle K \rangle_t$ is given by

$$
\begin{aligned}
S_t = {} & \sum_{k=0}^{\infty} P_t(k) \ln[P_t(k)] - \langle K \rangle_t \ln(\langle K \rangle_t) \\
& + \sum_{k=2}^{\infty} \left(k + \frac{1}{2} \right) \ln(k) P_t(k) \\
& + \frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(2\pi) P_t(0) + \left[1 - \frac{1}{2} \ln(2\pi) \right] P_t(1) \\
& + \left[2 - \frac{3}{2} \ln(2) - \frac{1}{2} \ln(2\pi) \right] P_t(2). \quad (34)
\end{aligned}
$$

Equation (34) is used in order to evaluate the relative entropy during the contraction process, where $P_t(k)$ is obtained either from numerical integration of the master equation or from computer simulations.

V. CONVERGENCE OF THE RELATIVE ENTROPY

In each of the network contraction scenarios, the degree distribution $P_t(k)$ evolves in time according to the master equation [Eq. (10)]. As a result, the relative entropy S_t of the network also evolves as the network contracts. The time derivative of S_t is given by

$$
\frac{d}{dt} S_t = \sum_{k=0}^{\infty} \ln \left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \right] \frac{d}{dt} P_t(k) + \sum_{k=0}^{\infty} \frac{d}{dt} P_t(k)
$$
$$
- \sum_{k=0}^{\infty} \frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \frac{d}{dt} \pi(k|\langle K \rangle_t). \quad (35)
$$

Replacing the order of the summation and the derivative in the second term on the right-hand side of Eq. (35), we obtain

$$
\sum_{k=0}^{\infty} \frac{d}{dt} P_t(k) = \frac{d}{dt} \left[\sum_{k=0}^{\infty} P_t(k) \right] = 0. \quad (36)
$$

Inserting the derivative $d\pi(k|\langle K \rangle_t)/dt$ from Eq. (16) into the third term on the right-hand side of Eq. (35), we obtain

$$
\sum_{k=0}^{\infty} \frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \frac{d}{dt} \pi(k|\langle K \rangle_t)
$$
$$
= -\frac{d\langle K \rangle_t}{dt} \sum_{k=0}^{\infty} \left(1 - \frac{k}{\langle K \rangle_t} \right) P_t(k) = 0. \quad (37)
$$

Since the second and third terms in Eq. (35) vanish, the time derivative of the relative entropy is simply given by

$$
\frac{d}{dt} S_t = \sum_{k=0}^{\infty} \ln \left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \right] \frac{d}{dt} P_t(k). \quad (38)
$$

This is a general equation that applies to any network contraction scenario in which the Poisson distribution $\pi(k|\langle K \rangle_t)$ is a solution. The relative entropy satisfies $S_t \geqslant 0$ for any degree distribution $P_t(k)$. It vanishes if and only if $P_t(k) = \pi(k|\langle K \rangle_t)$. Therefore, in order to prove the convergence of the degree distribution $P_t(k)$ towards a Poisson distribution in a given network contraction scenario, one needs to show that for this scenario $dS_t/dt < 0$. To this end, we use Eq. (38), where we replace the derivative dP_t/dt by the right-hand side of the master equation (10).

For the analysis below it is convenient to express the time evolution of the relative entropy, given by Eq. (38), in the form

$$
\frac{d}{dt} S_t = \Delta_{\mathrm{A}}(t) + \Delta_{\mathrm{B}}(t), \quad (39)
$$

where $\Delta_{\mathrm{A}}(t)$ emanates from the $F_{\mathrm{A}}(t)$ term (trickle-down term) of the master equation and $\Delta_{\mathrm{B}}(t)$ emanates from the $F_{\mathrm{B}}(t)$ term (redistribution term). The contribution of the trickle-down term to dS_t/dt is given by

$$
\Delta_{\mathrm{A}}(t) = \frac{A_t}{N_t} \sum_{k=0}^{\infty} \ln \left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \right] [(k+1) P_t(k+1) - k P_t(k)],
$$
$$
(40)
$$

where A_t is given by Eq. (13), and the contribution of the redistribution term is given by

$$
\Delta_{\mathrm{B}}(t) = -\frac{B}{N_t} \sum_{k=0}^{\infty} \ln \left[\frac{P_t(k)}{\pi_t(k|\langle K \rangle_t)} \right] \left(\frac{k}{\langle K \rangle_t} - 1 \right) P_t(k), \quad (41)
$$

where

$$
B = \begin{cases} 0 & \text{random deletion} \\ 1 & \text{preferential deletion} \\ 1 & \text{propagating deletion} \end{cases} \quad (42)
$$

In order to show that the degree distribution of the contracting network converges towards a Poisson distribution, one needs to show that during the contraction process $\Delta_{\mathrm{A}}(t) + \Delta_{\mathrm{B}}(t) < 0$. Below we consider each one of these terms separately. We show that in all three network contraction scenarios and for any initial degree distribution $P_0(k)$, the trickle-down term satisfies $\Delta_{\mathrm{A}}(t) < 0$ at all times during the contraction process. For the redistribution term $\Delta_{\mathrm{B}}(t)$ we obtain a necessary and sufficient condition on the instantaneous degree distribution $P_t(k)$ under which $\Delta_{\mathrm{B}}(t) < 0$. The condition essentially states that $\Delta_{\mathrm{B}}(t) < 0$ for any degree distribution whose tail decays more slowly than the tail of the Poisson

distribution, which decays superexponentially. This condition is generically satisfied by empirical networks, which are formed via growth processes. The degree distributions of such networks typically exhibit fat tails, which decay much more slowly than Poisson.

A. Convergence due to the trickle-down term

To gain more insight on the structure of the $\Delta_A(t)$ term, given by Eq. (40), it is useful to express it in the form

$$\Delta_A(t) = \frac{A_t}{N_t} \left\{ \sum_{k=0}^{\infty} \ln \left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \right] (k+1) P_t(k+1) \right.$$
$$\left. - \sum_{k=1}^{\infty} \ln \left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \right] k P_t(k) \right\}. \quad (43)$$

Taking a factor of $\langle K \rangle_t$ out of the curly brackets and multiplying the numerators and denominators in the arguments of the logarithmic functions by $k/\langle K \rangle_t$ (for $k \geq 1$), we obtain

$$\Delta_A(t) = \frac{A_t}{N_t} \{ \langle K \rangle_t + \ln[P_t(0)] \} P_t(1)$$
$$+ \frac{A_t \langle K \rangle_t}{N_t} \left\{ \sum_{k=1}^{\infty} \ln \left[\frac{\widetilde{P}_t(k)}{\pi(k-1|\langle K \rangle_t)} \right] \widetilde{P}_t(k+1) \right.$$
$$\left. - \sum_{k=1}^{\infty} \ln \left[\frac{\widetilde{P}_t(k)}{\pi(k-1|\langle K \rangle_t)} \right] \widetilde{P}_t(k) \right\}, \quad (44)$$

where

$$\widetilde{P}_t(k) = \frac{k}{\langle K \rangle_t} P_t(k) \quad (45)$$

is the degree distribution of nodes selected via a random edge in a random network with degree distribution $P_t(k)$. Similarly, the distribution

$$\pi(k-1|\langle K \rangle_t) = \frac{k}{\langle K \rangle_t} \pi(k|\langle K \rangle_t) \quad (46)$$

can be interpreted as the degree distribution of nodes selected via a random edge in an ER network with a Poisson degree distribution of the form $\pi(k|\langle K \rangle_t)$.

Rewriting $\widetilde{P}_t(k+1)$ in the form $[\widetilde{P}_t(k+1)/\widetilde{P}_t(k)]\widetilde{P}_t(k)$, one can express the $\Delta_A(t)$ term as a covariance of the form

$$\Delta_A(t) = \frac{A_t}{N_t} (\langle K \rangle_t P_t(1) + \ln[P_t(0)] P_t(1)$$
$$- \frac{P_t(1)}{\langle K \rangle_t} S[\widetilde{P}_t(k) || \pi(k-1|\langle K \rangle_t)]$$
$$+ \widetilde{\mathbb{E}}_t \left\{ \frac{\widetilde{P}_t(k+1)}{\widetilde{P}_t(k)} \ln \left[\frac{\widetilde{P}_t(k)}{\pi(k-1|\langle K \rangle_t)} \right] \right\}$$
$$- \widetilde{\mathbb{E}}_t \left[\frac{\widetilde{P}_t(k+1)}{\widetilde{P}_t(k)} \right] \widetilde{\mathbb{E}}_t \left\{ \ln \left[\frac{\widetilde{P}_t(k)}{\pi(k-1|\langle K \rangle_t)} \right] \right\} \right), \quad (47)$$

where $\widetilde{\mathbb{E}}_t[f(k)] = \sum_k f(k) \widetilde{P}_t(k)$. In particular,

$$\widetilde{\mathbb{E}}_t \left[\frac{\widetilde{P}_t(k+1)}{\widetilde{P}_t(k)} \right] = \sum_{k=1}^{\infty} \left[\frac{\widetilde{P}_t(k+1)}{\widetilde{P}_t(k)} \right] \widetilde{P}_t(k) = 1 - \frac{P_t(1)}{\langle K \rangle_t}. \quad (48)$$

In order that the covariance will be negative, in domains in which $\widetilde{P}_t(k)$ is an increasing function [namely, $\widetilde{P}_t(k+1) > \widetilde{P}_t(k)$], it should be lower than the corresponding Poisson distribution [namely, $\widetilde{P}_t(k) < \pi(k-1|\langle K \rangle_t)$], while in domains in which $\widetilde{P}_t(k)$ is a decreasing function it should be higher than the corresponding Poisson distribution.

In order to prove that $\Delta_A(t) < 0$ for any degree distribution $P_t(k)$ at all stages of the contraction process we rewrite Eq. (40) in the form

$$\Delta_A(t) = \Delta_A^P(t) - \Delta_A^\pi(t), \quad (49)$$

where

$$\Delta_A^P(t) = \frac{A_t}{N_t} \sum_{k=0}^{\infty} \ln[P_t(k)][(k+1)P_t(k+1) - kP_t(k)] \quad (50)$$

and

$$\Delta_A^\pi(t) = \frac{A_t}{N_t} \sum_{k=0}^{\infty} \ln[\pi(k|\langle K \rangle_t)][(k+1)P_t(k+1) - kP_t(k)]. \quad (51)$$

Separating the sum in Eq. (50) into two sums and replacing $k + 1$ by k in the first sum, we obtain

$$\Delta_A^P(t) = \frac{A_t}{N_t} \left\{ \sum_{k=1}^{\infty} \ln[P_t(k-1)] k P_t(k) - \sum_{k=1}^{\infty} \ln[P_t(k)] k P_t(k) \right\}. \quad (52)$$

Expressing the degree distribution $P_t(k)$ in terms of $\widetilde{P}_t(k)$,

$$\Delta_A^P(t) = \frac{A_t \langle K \rangle_t}{N_t} \left\{ \sum_{k=1}^{\infty} \ln[P_t(k-1)] \widetilde{P}_t(k) \right.$$
$$\left. - \sum_{k=1}^{\infty} \ln[\widetilde{P}_t(k)] \widetilde{P}_t(k) \right\}$$
$$+ \frac{A_t}{N_t} \sum_{k=1}^{\infty} \ln \left(\frac{k}{\langle K \rangle_t} \right) k P_t(k). \quad (53)$$

Combining the first two terms in Eq. (53) and splitting the last term, we obtain

$$\Delta_A^P(t) = -\frac{A_t \langle K \rangle_t}{N_t} \sum_{k=1}^{\infty} \widetilde{P}_t(k) \ln \left[\frac{\widetilde{P}_t(k)}{P_t(k-1)} \right] + \frac{A_t}{N_t} \langle K \ln(K) \rangle_t$$
$$- \frac{A_t}{N_t} \langle K \rangle_t \ln(\langle K \rangle_t). \quad (54)$$

In order to evaluate Δ_A^π we insert

$$\ln[\pi(k|\langle K \rangle_t)] = -\langle K \rangle_t + k \ln(\langle K \rangle_t) - \ln(k!) \quad (55)$$

into Eq. (51) and obtain

$$\Delta_A^\pi(t) = \frac{A_t}{N_t} \sum_{k=0}^{\infty} [-\langle K \rangle_t + k \ln(\langle K \rangle_t)$$
$$- \ln(k!)][(k+1)P_t(k+1) - kP_t(k)]. \quad (56)$$

Carrying out the summation and using the identity

$$\ln(k!) = \ln[(k+1)!] - \ln(k+1), \quad (57)$$

we obtain

$$\Delta_{\mathrm{A}}^{\pi}(t) = \frac{A_t}{N_t} \langle K \ln(K) \rangle_t - \frac{A_t}{N_t} \langle K \rangle_t \ln(\langle K \rangle_t). \tag{58}$$

Inserting the results for $\Delta_{\mathrm{A}}^{\mathrm{P}}$ and $\Delta_{\mathrm{A}}^{\pi}$, from Eqs. (54) and (58), respectively, into Eq. (49), we obtain

$$\Delta_{\mathrm{A}}(t) = -\frac{A_t \langle K \rangle_t}{N_t} S[\widetilde{P}_t(k) || P_t(k-1)], \tag{59}$$

where

$$S[\widetilde{P}_t(k) || P_t(k-1)] = \sum_{k=1}^{\infty} \widetilde{P}_t(k) \ln \left[\frac{\widetilde{P}_t(k)}{P_t(k-1)} \right] \tag{60}$$

is the relative entropy of $\widetilde{P}_t(k)$ with respect to $P_t(k-1)$. Note that Eq. (60) is valid only if $P_t(k-1) > 0$ for all values of k for which $\widetilde{P}_t(k) > 0$. This means that the degree distribution should not have any gaps, namely, values of k' for which $P_t(k') = 0$ while $P_t(k) > 0$ for any $k > k'$. In practice, even if there are such gaps in the initial degree distribution $P_0(k)$, they are quickly filled up due to the trickle-down term $F_{\mathrm{A}}(t)$ of the master equation, given by Eq. (11).

Since the relative entropy must be positive, we find that $\Delta_A(t) < 0$ for any degree distribution $P_t(k)$ that differs from $\pi(k|\langle K \rangle_t)$. Actually, since the only distribution for which $S[\widetilde{P}_t(k) || P_t(k-1)] = 0$ is the Poisson distribution, this process can converge only to the Poisson distribution. In the random deletion scenario, only the $\Delta_A(t)$ term contributes to the time evolution of S_t, while the $\Delta_B(t)$ term vanishes. This means that in the random deletion scenario the distance between $P_t(k)$ and the corresponding Poisson distribution $\pi(k|\langle K \rangle_t)$ with the same mean degree $\langle K \rangle_t$ decreases monotonically at any stage during the contraction process. In the preferential deletion and the propagating deletion scenarios the convergence also depends on the $\Delta_B(t)$ term, which is considered below.

B. Convergence due to the redistribution term

In order to gain insight on the $\Delta_B(t)$ term, we rewrite Eq. (41) in the form

$$\Delta_{\mathrm{B}}(t) = -\frac{B}{N_t} \left\{ \sum_{k=1}^{\infty} \ln \left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \right] \frac{k}{\langle K \rangle_t} P_t(k) \right. $$
$$\left. - \sum_{k=0}^{\infty} \ln \left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \right] P_t(k) \right\}. \tag{61}$$

Taking the factor of $1/\langle K \rangle_t$ out of the curly brackets, we obtain

$$\Delta_{\mathrm{B}}(t) = -\frac{B}{\langle K \rangle_t N_t} \left\{ \sum_{k=1}^{\infty} k \ln \left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \right] P_t(k) \right.$$
$$\left. - \langle K \rangle_t \sum_{k=0}^{\infty} \ln \left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \right] P_t(k) \right\}. \tag{62}$$

The expression in the curly brackets is, in fact, equal to the covariance between k and $\ln[P_t(k)/\pi(k|\langle K \rangle_t)]$ under the

distribution $P_t(k)$, namely,

$$\Delta_{\mathrm{B}}(t) = -\frac{B}{\langle K \rangle_t N_t} \left\{ \left\langle k \ln \left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \right] \right\rangle \right.$$
$$\left. - \langle K \rangle_t \left\langle \ln \left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \right] \right\rangle \right\}. \tag{63}$$

Therefore, in the case of distributions for which the correlation between k and $\ln[P_t(k)/\pi(k|\langle K \rangle_t)]$ is positive, the term in the curly brackets is positive and $\Delta_B(t) < 0$. In this case the $\Delta_B(t)$ term contributes to the convergence of $P_t(k)$ towards a Poisson distribution. Such positive correlation essentially implies that for large values of k, $P_t(k)$ tends to be larger than $\pi(k|\langle K \rangle_t)$, namely, it has a heavier tail than the Poisson distribution with the same mean value. Since network growth processes generically lead to fat tail distributions such as the power-law distributions of scale-free networks, it is expected that most empirical networks will exhibit a positive correlation between k and $\ln[P_t(k)/\pi(k|\langle K \rangle_t)]$.

In those cases in which the correlation between k and $\ln[P_t(k)/\pi(k|\langle K \rangle_t)]$ is negative, the term in the curly brackets is negative and $\Delta_B(t) > 0$. In this case the $\Delta_B(t)$ term works against the convergence of $P_t(k)$ towards a Poisson distribution. However, comparing the coefficients of $\Delta_A(t)$ and $\Delta_B(t)$ one finds that the coefficient of $\Delta_A(t)$ is effectively larger by a factor of $\langle K^2 \rangle / \langle K \rangle$ than the coefficient of $\Delta_B(t)$. Therefore, it is expected that the $\Delta_A(t)$ term will be dominant and induce the convergence of $P_t(k)$ towards Poisson even in those cases in which $\Delta_B(t) > 0$.

To gain more insight into the sign of $\Delta_B(t)$ from a different perspective, we use Eqs. (45) and (46) to express $\Delta_B(t)$ of Eq. (61) in the form

$$\Delta_{\mathrm{B}}(t) = -\frac{B}{N_t} \sum_{k=1}^{\infty} \ln \left[\frac{\widetilde{P}_t(k)}{\pi(k-1|\langle K \rangle_t)} \right] \widetilde{P}_t(k)$$
$$+ \frac{B}{N_t} \sum_{k=0}^{\infty} \ln \left[\frac{P_t(k)}{\pi(k|\langle K \rangle_t)} \right] P_t(k). \tag{64}$$

The first sum in Eq. (64) is the relative entropy of the degree distribution $\widetilde{P}_t(k)$ with respect to the shifted Poisson distribution $\pi(k-1|\langle K \rangle_t)$. This is essentially a distance measure between the degree distribution of nodes selected preferentially in a network whose degree distribution is $P_t(k)$ and the degree distribution of nodes selected preferentially from the corresponding Poisson distribution with the same mean degree. The second term in Eq. (64) is the relative entropy of the degree distribution $P_t(k)$ with respect to the Poisson distribution $\pi(k|\langle K \rangle_t)$, which is essentially a distance measure between $P_t(k)$ and $\pi(k|\langle K \rangle_t)$. Thus, Eq. (64) can be written in the form

$$\Delta_{\mathrm{B}}(t) = -\frac{B}{N_t} \{ S[\widetilde{P}_t(k) || \pi(k-1|\langle K \rangle_t)]$$
$$- S[P_t(k) || \pi(k|\langle K \rangle_t)] \}. \tag{65}$$

In the case that the degree distributions obtained for the preferential selection are farther apart than the degree distributions obtained for random selection, $\Delta_B(t) < 0$, while in the opposite case $\Delta_B(t) > 0$.

There is an important distinction between the two terms in Eq. (65). The second term is the relative entropy of $P_t(k)$ with respect to the Poisson distribution $\pi(k|\langle K \rangle_t)$ with the same mean degree $\langle K \rangle_t$. In contrast, the first term is the relative entropy of $\widetilde{P}_t(k)$ with respect to the Poisson distribution $\pi(k-1|\langle K \rangle_t)$. The mean degree of $\widetilde{P}_t(k)$ is

$$\langle \widetilde{K} \rangle_t = \frac{\langle K^2 \rangle_t}{\langle K \rangle_t}, \tag{66}$$

while the mean degree of $\pi(k-1|\langle K \rangle_t)$ is $\langle K \rangle_t + 1$. Therefore, Eq. (65) can be written in the form

$$\Delta_{\mathrm{B}}(t) = -\frac{B}{N_t} \{ S[\widetilde{P}_t(k) || \pi(k-1|\langle \widetilde{K} \rangle_t - 1)]$$
$$- S[P_t(k) || \pi(k|\langle K \rangle_t)] \}$$
$$- \frac{B}{N_t} \delta S(\langle \widetilde{K} \rangle_t, \langle K \rangle_t + 1), \tag{67}$$

where $\delta S(\langle \widetilde{K} \rangle_t, \langle K \rangle_t + 1)$ is given by Eq. (29). This implies that $\Delta_{\mathrm{B}}(t) < 0$ as long as

$$S[\widetilde{P}_t(k) || \pi(k-1|\langle \widetilde{K} \rangle_t - 1)]$$
$$> S[P_t(k) || \pi(k|\langle K \rangle_t)] - \delta S(\langle \widetilde{K} \rangle_t, \langle K \rangle_t + 1). \tag{68}$$

Since $\delta S(\langle \widetilde{K} \rangle_t, \langle K \rangle_t + 1)$ is always positive and its value increases as $P(k)$ becomes broader, this condition is expected to be satisfied for any degree distribution that exhibits a heavy tail. From our experience, degree distributions for which $\Delta_{\mathrm{B}} > 0$ are very special, usually hand-crafted for the mission. In those cases, Δ_{A}, which is always negative, as proven above, is much larger in absolute value than Δ_{B}.

VI. CONTRACTION OF NETWORKS WITH GIVEN INITIAL DEGREE DISTRIBUTIONS

Here we apply the framework presented above to three examples of configuration model networks, with a degenerate degree distribution (also known as random regular graphs), an exponential degree distribution and a power-law degree distribution (scale-free networks).

A. Random regular graphs

A random regular graph (RRG) is a configuration model network in which all the nodes are of the same degree, $k = c_0$, namely,

$$P_0(k) = \delta_{k,c_0}, \tag{69}$$

where c_0 is an integer. Here we consider the case of $c_0 \geqslant 3$, in which the giant component encompasses the whole network. In order to leave room for contraction into a nontrivial degree distribution, we choose RRGs with $c_0 \gg 1$. Since in node deletion processes the degrees of nodes in the network are only reduced and never increase it is clear that the range of degrees of the contracted network will be limited to $0 \leqslant k \leqslant c_0$. This means that in the case that the initial network is an RRG the tail of the degree distribution of the contracted network will be truncated above $k = c_0$. Thus, the convergence towards Poisson is expected to be relatively slow.

To evaluate the relative entropy of the initial RRG network with respect to the corresponding Poisson distribution we

insert the degenerate distribution of Eq. (69) into Eq. (21). We obtain the initial relative entropy

$$S_0 = \ln \left[\frac{1}{\pi(c_0|c_0)} \right]. \tag{70}$$

Inserting the Poisson degree distribution into Eq. (70) we obtain

$$S_0 = c_0 - c_0 \ln(c_0) + \ln(c_0!). \tag{71}$$

Using the Stirling approximation to evaluate $\ln(c_0!)$, we obtain

$$S_0 = \frac{1}{2} \ln(c_0) + \frac{1}{2} \ln(2\pi). \tag{72}$$

Below we analyze the convergence of a configuration model network with a degenerate degree distribution towards an ER graph structure upon contraction. In particular, we calculate the time-dependent degree distribution $P_t(k)$ during contraction and examine its convergence towards $\pi(k|\langle K \rangle_t)$. To this end we perform direct numerical integration of the master equation (10) and computer simulations, starting from a configuration model network with a degree distribution given by Eq. (69), and evaluate the time-dependent relative entropy S_t.

In Fig. 2 we present the relative entropy S_t as a function of time (represented by $N_t/N_0 = 1 - t/N_0$) for a random regular graph of size $N_0 = 10^4$ with a degenerate degree distribution in which all the nodes are of degree $c_0 = 10$, that contracts via (a) random node deletion, (b) preferential node deletion, and (c) propagating node deletion. The results obtained from numerical integration of the master equation (solid lines) are in excellent agreement with the results obtained from computer simulations, namely, direct simulations of contracting networks (circles). In all three cases the relative entropy quickly decays, which implies that the degree distribution $P_t(k)$ of the contracting network converges towards a Poisson distribution. The decay rate of S_t is comparable in all the three scenarios. This implies that for extremely narrow degree distributions such as the degenerate distribution the preferential and the propagating deletion scenarios do not exhibit faster convergence than the random deletion scenario.

In Fig. 3(a) we present the degree distribution $P_0(k)$ of a random regular graph (solid line) of size $N_0 = 10^4$ with a degenerate degree distribution in which all the nodes are of degree $c_0 = 10$. The corresponding Poisson distribution with the same mean degree $\langle K \rangle_0 = c_0$ is also shown (dashed line). Clearly, it is highly dissimilar to the degenerate distribution. The random regular graph undergoes a network contraction process via the random node deletion scenario. In Fig. 3(b) we present the degree distribution $P_t(k)$ of the contracted network at time $t = 8000$, where the contracted network size is $N_t = 2000$. The results obtained from the numerical integration of the master equation (solid line) are in excellent agreement with the results of computer simulations (circles). They are very well converged towards the corresponding Poisson distribution $\pi(k|\langle K \rangle_t)$ with the same mean degree $\langle K \rangle_t$ (dashed line).
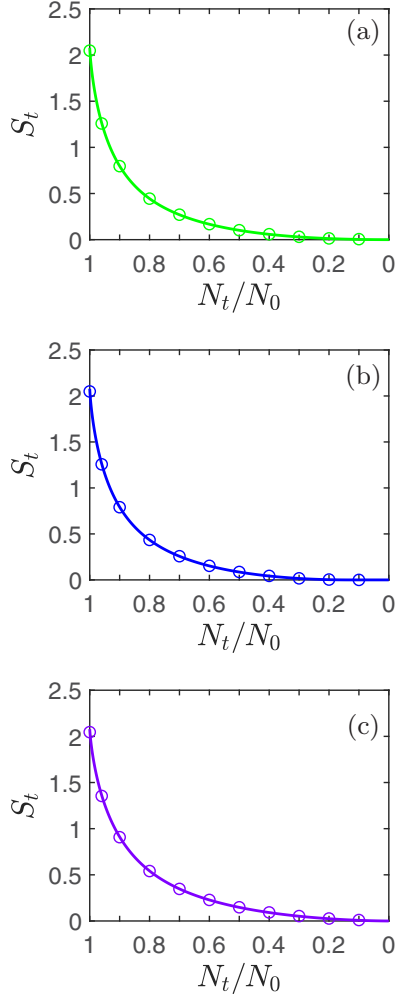
FIG. 2. The relative entropy S_t as a function of time for a random regular graph of initial size $N_0 = 10^4$ and initial degree $c_0 = 10$ that contracts via random deletion (a), preferential deletion (b), and propagating deletion (c), obtained from numerical integration of the master equation (solid lines). In all three cases the relative entropy quickly decays, which implies that the degree distribution of the contracting network converges towards a Poisson distribution. The master equation results are in excellent agreement with the results obtained from computer simulations (circles). Also, the initial value $S_0 \simeq 2.08$ is in perfect agreement with the result obtained from Eq. (72).

B. Configuration model networks with exponential degree distributions

Consider a configuration model network with an exponential degree distribution of the form $P_0(k) \sim e^{-\alpha k}$, where $k \geqslant k_{\min}$ and $k_{\min}$ is the lower cutoff of the initial degree distribution. It is convenient to parametrize the degree distribution using the mean degree $\langle K \rangle_0$, in the form

$$P_0(k) = \begin{cases} 0 & k < k_{\min} \\ D\left(\dfrac{\langle K \rangle_0 - k_{\min}}{\langle K \rangle_0 - k_{\min} + 1}\right)^k & k \geqslant k_{\min} \end{cases}, \quad (73)$$

where D is the normalization constant, given by

$$D = \frac{1}{(\langle K \rangle_0 - k_{\min}) + 1}\left(\frac{\langle K \rangle_0 - k_{\min}}{\langle K \rangle_0 - k_{\min} + 1}\right)^{-k_{\min}}. \quad (74)$$

Below we evaluate the relative entropy of an initial network with an exponential degree distribution with respect to the corresponding Poisson distribution. Inserting the exponential degree distribution of Eq. (73) into Eq. (23) and carrying out the summation, we obtain the Shannon entropy

$$S[P_0(k)] = -\sum_{k=k_{\min}}^{\infty} P_0(k) \ln[P_0(k)]$$
$$= -(\langle K \rangle_0 - k_{\min}) \ln(\langle K \rangle_0 - k_{\min})$$
$$+ (\langle K \rangle_0 - k_{\min} + 1) \ln(\langle K \rangle_0 - k_{\min} + 1). \quad (75)$$

In order to calculate the cross-entropy $C[P_0(k)||\pi(k|\langle K \rangle_0)]$, we insert the exponential distribution $P_0(k)$ of Eq. (73) into Eq. (33). We obtain

$$C[P_0(k)||\pi(k|\langle K \rangle_0)]$$
$$= -\langle K \rangle_0 \ln(\langle K \rangle_0) + \sum_{k=k_{\min}}^{\infty} \left(k + \frac{1}{2}\right)$$
$$\times \ln(k)\left[D\left(\frac{\langle K \rangle_0 - k_{\min}}{\langle K \rangle_0 - k_{\min} + 1}\right)^k\right]$$
$$+ \frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(2\pi)P_0(0) + \left[1 - \frac{1}{2}\ln(2\pi)\right]P_0(1)$$
$$+ \left[2 - \frac{3}{2}\ln(2) - \frac{1}{2}\ln(2\pi)\right]P_0(2). \quad (76)$$

Carrying out the summation, we obtain

$$C[P_0(k)||\pi(k|\langle K \rangle_0)]$$
$$= -\langle K \rangle_0 \ln(\langle K \rangle_0) - \frac{1}{2(\langle K \rangle_0 - k_{\min} + 1)}$$
$$\times \left[2\frac{\partial}{\partial \gamma}\Phi\left(\frac{\langle K \rangle_0 - k_{\min}}{\langle K \rangle_0 - k_{\min} + 1}, \gamma, k_{\min}\right)\bigg|_{\gamma=-1}\right.$$
$$\left. + \frac{\partial}{\partial \gamma}\Phi\left(\frac{\langle K \rangle_0 - k_{\min}}{\langle K \rangle_0 - k_{\min} + 1}, \gamma, k_{\min}\right)\bigg|_{\gamma=0}\right]$$
$$+ \frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(2\pi)P_0(0) + \left[1 - \frac{1}{2}\ln(2\pi)\right]P_0(1)$$
$$+ \left[2 - \frac{3}{2}\ln(2) - \frac{1}{2}\ln(2\pi)\right]P_0(2), \quad (77)$$

where $\Phi(x, \gamma, k)$ is the Lerch transcendent [48]. The relative entropy takes the form $S_0 = -S[P_0(k)] + C[P_0(k)||\pi(k|\langle K \rangle_0)]$, where $S[P_0(k)]$ is given by Eq. (75) and $C[P_0(k)||\pi(k|\langle K \rangle_0)]$ is given by Eq. (77).

Below we analyze the convergence of a configuration model network with an exponential degree distribution towards an ER graph structure upon contraction. In particular, we calculate the time-dependent degree distribution $P_t(k)$ during contraction and examine its convergence towards $\pi(k|\langle K \rangle_t)$. To this end we perform direct numerical integration of the master equation (10) and computer simulations,
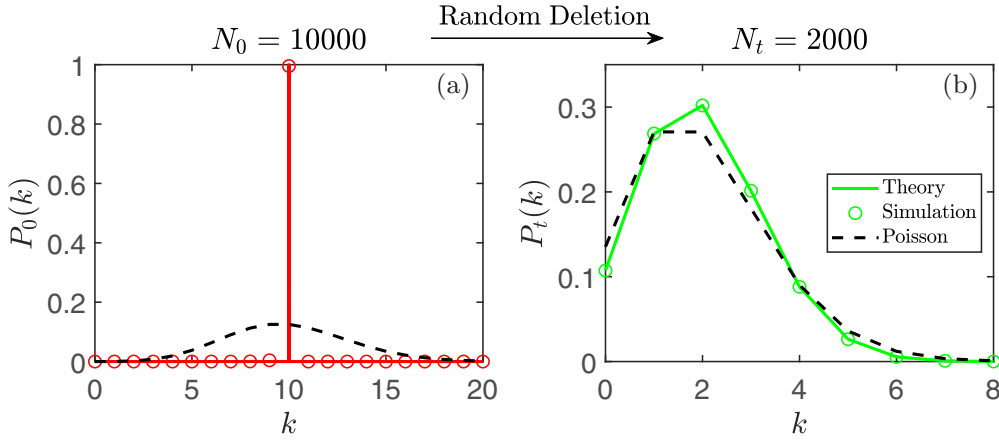
FIG. 3. (a) The degree distribution $P_0(k)$ of a random regular graph (solid line) in which all the nodes are of degree $c_0 = 10$. The circles represent the degree sequence of a single network instance of $N_0 = 10^4$ nodes, which was used in the computer simulations. The corresponding Poisson distribution with the same mean degree is also shown (dashed line). The network contracts via random node deletion. (b) The degree distribution $P_t(k)$ of the contracted network at time $t = 8000$, when the network size is reduced to $N_t = 2000$. The results obtained from numerical integration of the master equation (solid line) are in excellent agreement with the results obtained from computer simulations (circles). They are both very well converged towards the corresponding Poisson distribution $\pi(k|\langle K \rangle_t)$ with the same mean degree $\langle K \rangle_t$ (dashed line).

starting from a configuration model network with a degree distribution given by Eq. (73) and evaluate the time-dependent relative entropy S_t.

In Fig. 4 we present the relative entropy S_t as a function of time for a configuration model network of initial size $N_0 = 10^4$ and initial mean degree $\langle K \rangle_0 = 20$ with an exponential degree distribution that contracts via random deletion (a), preferential deletion (b), and propagating deletion (c), obtained from numerical integration of the master equation (solid lines). In all three cases the relative entropy quickly decays, which implies that the degree distribution of the contracting network converges towards a Poisson distribution. The convergence is dramatically faster in the preferential and the propagating deletion scenarios compared to random deletion scenario. The master equation results are in very good agreement with the results obtained from computer simulations (circles).

In Fig. 5(a) we present the degree distribution $P_0(k)$ of a configuration model network of size $N_0 = 10^4$ and an exponential degree distribution with mean degree $\langle K \rangle_0 = 20$ (solid line). The corresponding Poisson distribution with the same mean degree is also shown (dashed line). The network contracts via preferential node deletion. In Fig. 5(b) we present the degree distribution $P_t(k)$ of the contracted network at time $t = 7000$, when the network size is reduced to $N_t = 3000$, obtained from numerical integration of the master equation (solid line) and from computer simulations (circles). The corresponding Poisson distribution $\pi(k|\langle K \rangle_t)$ with the same mean degree is also shown (dashed line). The master equation results, the computer simulation results and the corresponding Poisson distribution are found to be in very good agreement with each other.

In Fig. 6 we present the time derivative of the relative entropy, $dS_t/dt = \Delta_A(t) + \Delta_B(t)$, as a function of time, for a configuration model network of initial size $N_0 = 10^4$ and exponential degree distribution with mean degree $\langle K \rangle_0 = 20$ that contracts via preferential node deletion, obtained from numerical integration of the master equation (solid lines). The

terms $\Delta_A(t)$ (dashed line) and $\Delta_B(t)$ (dotted line), which sum up to the derivative dS_t/dt, are also shown. As expected, both $\Delta_A(t)$ and $\Delta_B(t)$ are negative at all times during the contraction process.

C. Configuration model networks with power-law degree distributions

Consider a configuration model network with a power-law degree distribution of the form $P_0(k) \sim k^{-\gamma}$, where $1 \leqslant k_{\min} \leqslant k \leqslant k_{\max}$. Here we focus on the case of $\gamma > 2$, in which the mean degree, $\langle K \rangle_0$, is bounded even for $k_{\max} \to \infty$. Power-law distributions do not exhibit a typical scale and are therefore referred to as scale-free networks. The normalized degree distribution is given by

$$P_0(k) = \begin{cases} 0 & k < k_{\min} \\ D\,k^{-\gamma} & k_{\min} \leqslant k \leqslant k_{\max}, \\ 0 & k > k_{\max} \end{cases} \quad (78)$$

where D is the normalization constant, given by

$$D = D(\gamma, k_{\min}, k_{\max}) = \frac{1}{\zeta(\gamma, k_{\min}) - \zeta(\gamma, k_{\max} + 1)}, \quad (79)$$

and $\zeta(\gamma, k)$ is the Hurwitz zeta function [48]. For $2 < \gamma \leqslant 3$ the mean degree is bounded while the second moment, $\langle K^2 \rangle$, diverges in the limit of $k_{\max} \to \infty$. For $\gamma > 3$ both moments are bounded. The mean degree is given by

$$\langle K \rangle_0 = \frac{\zeta(\gamma - 1, k_{\min}) - \zeta(\gamma - 1, k_{\max} + 1)}{\zeta(\gamma, k_{\min}) - \zeta(\gamma, k_{\max} + 1)}. \quad (80)$$

The second moment of the degree distribution, when finite, is

$$\langle K^2 \rangle_0 = \frac{\zeta(\gamma - 2, k_{\min}) - \zeta(\gamma - 2, k_{\max} + 1)}{\zeta(\gamma, k_{\min}) - \zeta(\gamma, k_{\max} + 1)}. \quad (81)$$

Below we evaluate the relative entropy of an initial network with a power-law degree distribution with respect to the corresponding Poisson distribution. In order to calculate the
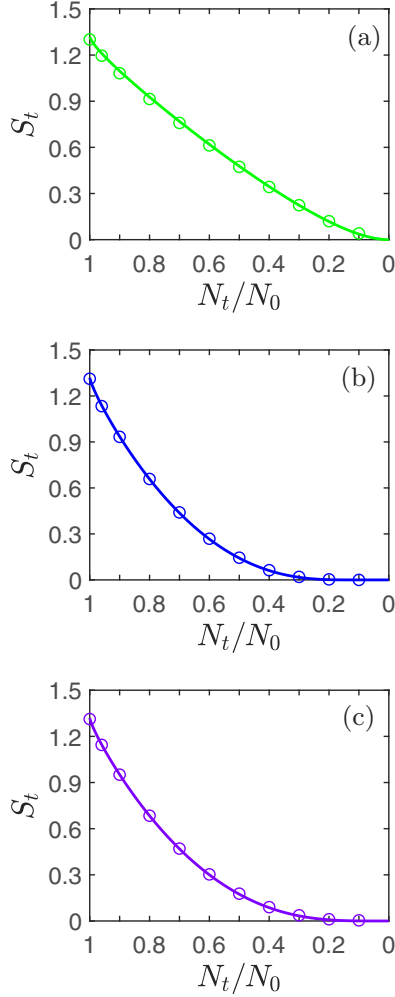
FIG. 4. The relative entropy S_t as a function of time for a configuration model network of initial size $N_0 = 10^4$ and mean degree $\langle K \rangle_0 = 20$ with an exponential degree distribution in which $k_{min} = 10$, that contracts via random deletion (a), preferential deletion (b), and propagating deletion (c), obtained from numerical integration of the master equation (solid lines). In all three cases the relative entropy quickly decays, which implies that the degree distribution of the contracting network converges towards a Poisson distribution. The convergence is dramatically faster in the preferential and the propagating deletion scenarios compared to random deletion scenario. The master equation results are in very good agreement with the results obtained from computer simulations (circles). Also, the initial value $S_0 \simeq 1.32$ is in perfect agreement with the result obtained from Eqs. (75) and (77).

Shannon entropy $S[P_0(k)]$ we insert the power-law distribution of Eq. (78) into Eq. (23). We obtain

$$
\begin{aligned}
S[P_0(k)] &= -\sum_{k=k_{min}}^{\infty} P_0(k) \ln[P_0(k)] \\
&= -\ln(D) + \gamma \sum_{k=k_{min}}^{\infty} D k^{-\gamma} \ln(k).
\end{aligned}
\tag{82}
$$

Since $\ln(1) = 0$ the summation in Eq. (82) actually starts from the larger value between $k = 2$ and k_{min}, denoted by

$\bar{k}_{min} = \max\{2, k_{min}\}$. We thus obtain

$$
S[P_0(k)] = -\ln(D) + \gamma \sum_{k=\bar{k}_{min}}^{\infty} D k^{-\gamma} \ln(k).
\tag{83}
$$

Carrying out the summation, we obtain

$$
S[P_0(k)] = -\ln(D) + \gamma D[\zeta'(\gamma, k_{max} + 1) - \zeta'(\gamma, \bar{k}_{min})],
\tag{84}
$$

where $\zeta'(\gamma, k) = \partial \zeta(\gamma, k)/\partial \gamma$.

In order to calculate the cross-entropy $C[P_0(k)||\pi(k|\langle K \rangle_0)]$, we insert the power-law distribution $P_0(k)$ into Eq. (33). We obtain

$$
\begin{aligned}
&C[P_0(k)||\pi(k|\langle K \rangle_0)] \\
&= -\langle K \rangle_0 \ln(\langle K \rangle_0) + \sum_{k=\bar{k}_{min}}^{\infty} \left(k + \frac{1}{2}\right) \ln(k) D k^{-\gamma} \\
&\quad + \frac{1}{2}\ln(2\pi) + \left[1 - \frac{1}{2}\ln(2\pi)\right] P_0(1) \\
&\quad + \left[2 - \frac{3}{2}\ln(2) - \frac{1}{2}\ln(2\pi)\right] P_0(2).
\end{aligned}
\tag{85}
$$

Carrying out the summation, we obtain

$$
\begin{aligned}
C[P_0(k)||\pi(k|\langle K \rangle_0)] &= -\langle K \rangle_0 \ln(\langle K \rangle_0) + D[\zeta'(\gamma - 1, \bar{k}_{min}) \\
&\quad - \zeta'(\gamma - 1, k_{max} + 1)] \\
&\quad + \frac{D}{2}[\zeta'(\gamma, k_{min}) - \zeta'(\gamma, k_{max} + 1)] \\
&\quad + \frac{1}{2}\ln(2\pi) + \left[1 - \frac{1}{2}\ln(2\pi)\right] P_0(1) \\
&\quad + \left[2 - \frac{3}{2}\ln(2) - \frac{1}{2}\ln(2\pi)\right] P_0(2).
\end{aligned}
\tag{86}
$$

The relative entropy of the initial network with a power-law degree distribution given by Eq. (78) takes the form $S_0 = -S[P_0(k)] + C[P_0(k)||\pi(k|\langle K \rangle_0)]$, where $S[P_0(k)]$ is given by Eq. (84) and $C[P_0(k)||\pi(k|\langle K \rangle_0)]$ is given by Eq. (86).

Below we analyze the convergence of a configuration model network with a power-law degree distribution towards an ER graph structure upon contraction. In particular, we calculate the time-dependent degree distribution $P_t(k)$ during contraction and examine its convergence towards $\pi(k|\langle K \rangle_t)$. To this end we perform direct numerical integration of the master equation (10) and computer simulations, starting from a configuration model network with a degree distribution given by Eq. (78) and evaluate the time-dependent relative entropy S_t.

In Fig. 7 we present the relative entropy S_t as a function of time for a configuration model network with a power-law degree distribution, of initial size $N_0 = 10^4$ and initial mean degree $\langle K \rangle_0 = 20$, where $k_{min} = 10$, $k_{max} = 100$ and $\gamma = 2.65$, that contracts via random deletion (a), preferential deletion (b), and propagating deletion (c), obtained from numerical integration of the master equation (solid lines). In all three cases the relative entropy quickly decays, which implies that the degree distribution of the contracting network
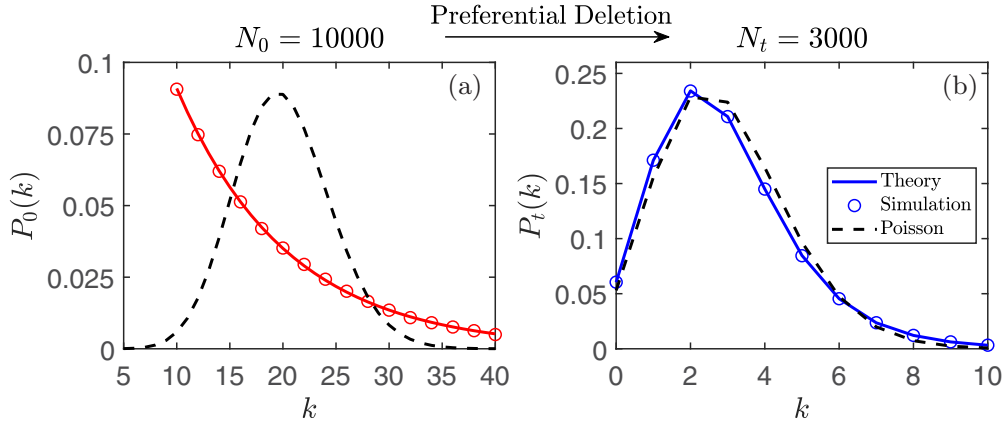
FIG. 5. (a) The degree distribution $P_0(k)$ of a configuration model network with mean degree $\langle K \rangle_0 = 20$ and an exponential degree distribution, given by Eq. (73) with $k_{\min} = 10$ (solid line). The circles represent the degree sequence of the $N_0 = 10^4$ nodes in a single realization of the initial network, which was used in the computer simulation. The corresponding Poisson distribution with the same mean degree is also shown (dashed line). The network contracts via the preferential node deletion scenario. (b) The degree distribution $P_t(k)$ of the contracted network at time $t = 7000$, when the network size is reduced to $N_t = 3000$, obtained from numerical integration of the master equation (solid line). The master equation results are in excellent agreement with the results obtained from computer simulations (circles). The corresponding Poisson distribution $\pi(k|\langle K \rangle_t)$ with the same mean degree is also shown (dashed line). The master equation results and the computer simulation results are in very good agreement with the corresponding Poisson distribution with the same mean degree.

converges towards a Poisson distribution. The convergence is dramatically faster in the preferential and the propagating deletion scenarios compared to random deletion scenario. The master equation results are in very good agreement with the results obtained from computer simulations (circles).

In Fig. 8(a) we present the degree distribution $P_0(k)$ of a configuration model network of size $N_0 = 10^4$ and a power-law degree distribution with mean degree $\langle K \rangle_0 = 20$ (solid line). The corresponding Poisson distribution with the same mean degree is also shown (dashed line). The network
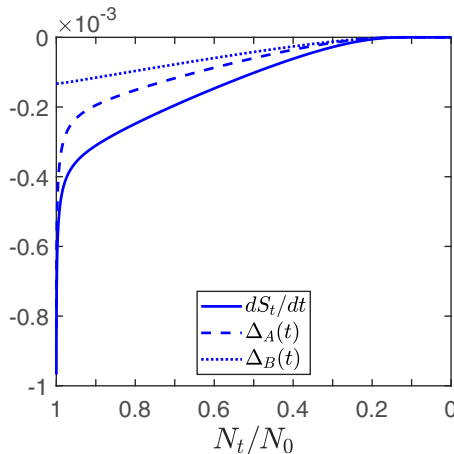

FIG. 6. The time derivative of the relative entropy, $dS_t/dt = \Delta_{\mathrm{A}}(t) + \Delta_{\mathrm{B}}(t)$, as a function of time, for a configuration model network of initial size $N_0 = 10^4$ and exponential degree distribution with mean degree $\langle K \rangle_0 = 20$ and $k_{\min} = 10$, that contracts via preferential node deletion, obtained from numerical integration of the master equation (solid lines). The terms $\Delta_{\mathrm{A}}(t)$ (dashed line) and $\Delta_{\mathrm{B}}(t)$ (dotted line), which sum up to the derivative dS_t/dt are also shown. Note that both $\Delta_{\mathrm{A}}(t)$ and $\Delta_{\mathrm{B}}(t)$ are negative at all times during the contraction process.

contracts via propagating node deletion. In Fig. 8(b) we present the degree distribution $P_t(k)$ of the contracted network at $t = 7000$, when the network size is reduced to $N_t = 3000$, obtained from numerical integration of the master equation (solid line) and from computer simulations (circles). The corresponding Poisson distribution $\pi(k|\langle K \rangle_t)$ with the same mean degree is also shown (dashed line). The master equation results, the computer simulation results, and the corresponding Poisson distribution are found to be in very good agreement with each other.

In Fig. 9 we present the time derivative of the relative entropy, dS_t/dt as a function of time, for a configuration model network of initial size $N = 10^4$ and a power-law degree distribution with mean degree $\langle K \rangle_0 = 20$ that contracts via propagating node deletion, obtained from numerical integration of the master equation (solid lines). The terms $\Delta_{\mathrm{A}}(t)$ (dashed line) and $\Delta_{\mathrm{B}}(t)$ (dotted line), which sum up to the derivative dS_t/dt are also shown. As expected, both $\Delta_{\mathrm{A}}(t)$ and $\Delta_{\mathrm{B}}(t)$ are negative at all times during the contraction process.

VII. DISCUSSION

In Ref. [31] we used direct numerical integration of the master equation and computer simulations to show that the degree distributions of contracting networks converge towards the Poisson distribution. To this end, we used the relative entropy as a distance measure between the degree distribution $P_t(k)$ of the contracting network and the corresponding Poisson distribution $\pi(k|\langle K \rangle_t)$, and showed that this distance decreases as the network contracts.

A computer simulation of network contraction provides results for a single instance of the initial network and a single stochastic path of the contraction process. In order to obtain statistically significant results for a given ensemble of initial networks and given network contraction scenario one needs
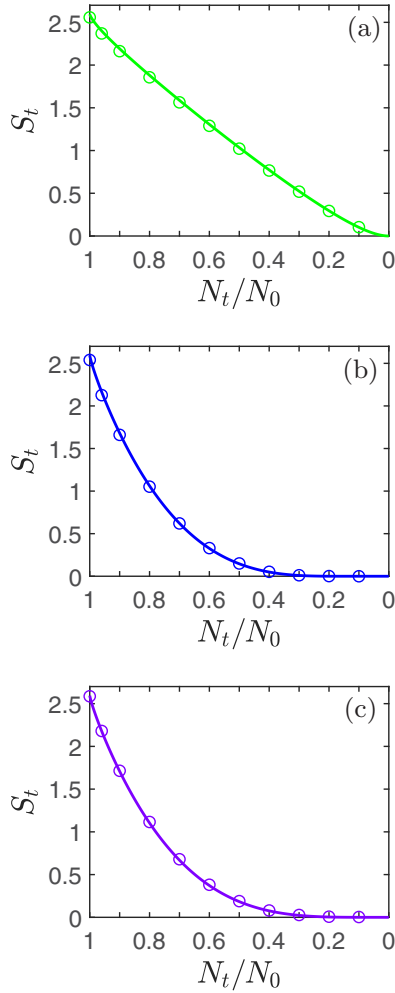
FIG. 7. The relative entropy S_t as a function of time for a configuration model network with a power-law degree distribution of initial size $N_0 = 10^4$ and mean degree $\langle K \rangle_0 = 20$, where $k_{\min} = 10$, $k_{\max} = 100$ and $\gamma = 2.65$, that contracts via random deletion (a), preferential deletion (b), and propagating deletion (c), obtained from numerical integration of the master equation (solid lines). In all three cases the relative entropy quickly decays, which implies that the degree distribution of the contracting network converges towards a Poisson distribution. The convergence is dramatically faster in the preferential and the propagating deletion scenarios compared to random deletion scenario. The master equation results are in very good agreement with the results obtained from computer simulations (circles). Also, the initial value $S_0 \simeq 2.59$ is in perfect agreement with the result obtained from Eqs. (84) and (86).

to combine the results of a large number of independent runs. The direct numerical integration of the master equation is advantageous in the sense that a single run of the numerical integration process provides results for a whole ensemble of initial networks. However, a given network ensemble represents a single point in the high-dimensional parameter space of possible network ensembles. Therefore, in order to explore the general properties of network contraction processes one needs to repeatedly apply the direct integration of the master equation to a large sample of distinct network ensembles.

Our aim in this paper was to obtain rigorous analytical results for the convergence of contracting networks towards the ER network ensemble. To this end we devised a rigorous argument, which is based on the master equation that describes the temporal evolution of the degree distribution $P_t(k)$ and the relative entropy S_t. Such an argument is advantageous over the direct numerical integration of the master equation or computer simulations in the sense that it is universally applicable to all possible degree distributions.

The relative entropy $S[P(k)||Q(k)]$ of a distribution $P(k)$ with respect to a distribution $Q(k)$ is a special case of the Rényi divergence $S_\alpha[P(k)||Q(k)]$, with $\alpha = 1$ [49]. The choice of $\alpha = 1$ is advantageous in the sense that it has a natural information theoretic interpretation [35,36]. The relative entropy is an asymmetric distance measure or quasidistance [50]. Interestingly, the relative entropy is related to other distance measures between discrete probability distributions. For example, the total variation distance between probability distributions $P(k)$ and $Q(k)$ is given by $T[P(k), Q(k)] = \sum_k |P(k) - Q(k)|$, namely, the sum of the differences (in absolute value) between the probabilities assigned to all values of k by the two distributions. Clearly, for any two distributions $P(k)$ and $Q(k)$, the total variation distance satisfies $0 \leqslant T[P(k), Q(k)] \leqslant 2$. The relative entropy provides an additional upper bound on the total variation distance via the Pinsker inequality, which takes the form [51–54]

$$T[P(k), Q(k)] \leqslant \sqrt{\frac{1}{2} S[P(k)||Q(k)]}. \quad (87)$$

This relation implies that whenever the relative entropy between $P(k)$ and $Q(k)$ vanishes, so does the total variation distance between them, meaning that the two distributions become identical in the L_1 norm. This shows that when the relative entropy vanishes the distributions become identical.

In this paper we focused on the case of configuration model networks, which exhibit a given degree distribution and no degree-degree correlations. The theoretical framework presented here may provide the foundations for the study of network contraction processes in a much broader class of complex networks, which exhibit degree-degree correlations as well as other structural correlations. This will require a more general formulation of the relative entropy, expressed in terms of the joint degree distributions of pairs or adjacent nodes, which take into account the correlations between their degrees.

The theoretical framework presented here may be relevant in the broad context of neurodegeneration, which is the progressive loss of structure and function of neurons in the brain. Such processes occur in normal aging [55] as well as in a large number of incurable neurodegenerative diseases such as Alzheimer's, Parkinson's, Huntington's, and amylotrophic lateral sclerosis, which result in a gradual loss of cognitive and motoric functions [56]. These diseases differ in the specific brain regions or circuits in which the degeneration occurs. The characterization of the evolving structure using the relative entropy may provide useful insight into the structural aspects of the loss of neurons and synapses in neurodegenerative processes [57].
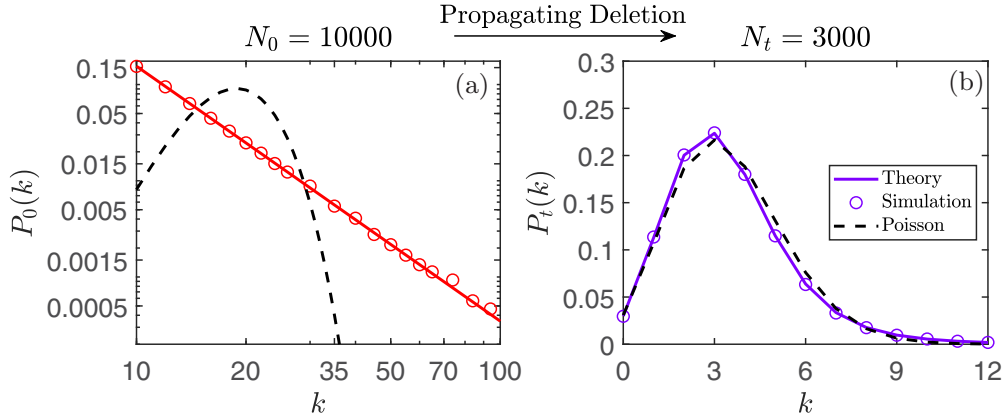
FIG. 8. (a) The degree distribution $P_0(k)$ of a configuration model network with a power-law degree distribution, given by Eq. (78), and mean degree $\langle K \rangle_0 = 20$ (solid line), where $k_{\min} = 10$, $k_{\max} = 100$ and $\gamma = 2.65$, is shown on a log-log scale. The circles represent the degree sequence of the $N_0 = 10^4$ nodes in a single realization of the initial network, which was used in the computer simulation. The corresponding Poisson distribution with the same mean degree is also shown (dashed line). The network contracts via the propagating node deletion scenario. (b) The degree distribution $P_t(k)$ of the contracted network at time $t = 7000$, when the network size is reduced to $N_t = 3000$, obtained from numerical integration of the master equation is shown on a linear scale (solid line). The master equation results are in excellent agreement with the results obtained from computer simulations (circles). The corresponding Poisson distribution $\pi(k|\langle K \rangle_t)$ with the same mean degree is also shown (dashed line). The master equation results and the computer simulation results are in very good agreement with the corresponding Poisson distribution with the same mean degree.

It is worth mentioning that there is another class of network dismantling processes that involve optimized attacks, which maximize the damage to the network for a minimal set of deleted nodes [29,30]. Such optimization is achieved by first decycling the network, namely, by selectively deleting nodes that reside on cycles, thus driving the giant component into a tree structure. The branches of the tree are then trimmed such that the giant component is quickly disintegrates. Clearly,

these optimized dismantling processes do not converge towards an ER structure.

VIII. SUMMARY

In summary, we have analyzed the structural evolution of complex networks undergoing contraction processes via generic node deletion scenarios, namely, random deletion, preferential deletion, and propagating deletion. Focusing on configuration model networks we have shown using a rigorous argument that upon contraction the degree distributions of these networks converge towards a Poisson distribution. In this analysis we used the relative entropy $S_t = S[P_t(k)||\pi(k|\langle K \rangle_t)]$ of the degree distribution $P_t(k)$ of the contracting network at time t with respect to the corresponding Poisson distribution $\pi(k|\langle K \rangle_t)$ with the same mean degree $\langle K \rangle_t$ as a distance measure between $P_t(k)$ and Poisson. The relative entropy is suitable as a distance measure since it satisfies $S_t \geqslant 0$ for any degree distribution $P_t(k)$, while equality is obtained only for $P_t(k) = \pi(k|\langle K \rangle_t)$. We derived an equation for the time evolution of the relative entropy S_t during network contraction and expressed its time derivative dS_t/dt as a sum of two terms, $\Delta_A(t)$ and $\Delta_B(t)$. We have shown that the first term satisfies $\Delta_A(t) < 0$ for any degree distribution $P_t(k)$. This means that the $\Delta_A(t)$ term always pushes the relative entropy down towards zero, driving the convergence of $P_t(k)$ towards Poisson. For the $\Delta_B(t)$ term we provide a condition that can be used for any given degree distribution $P_t(k)$ to determine whether this term would accelerate the convergence to Poisson or slow it down. The condition implies that for degree distributions $P_t(k)$ whose tail falls more slowly than the tail of the corresponding Poisson distribution, the $\Delta_B(t)$ term would accelerate the convergence to Poisson, while in the case that the tail falls more quickly than Poisson the $\Delta_B(t)$ term would slow down the convergence. We analyzed the
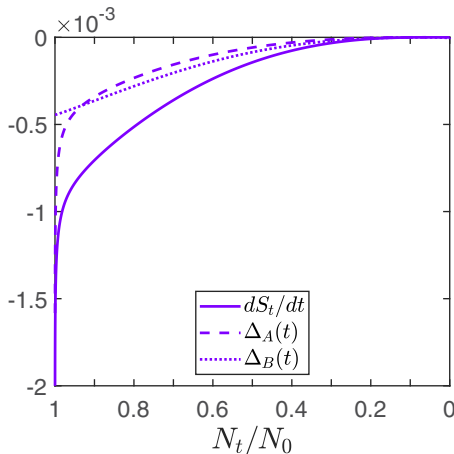

FIG. 9. The time derivative of the relative entropy dS_t/dt as a function of time, for a configuration model network of initial size $N_0 = 10^4$ and a power-law degree distribution with mean degree $\langle K \rangle_0 = 20$, where $k_{\min} = 10$, $k_{\max} = 100$ and $\gamma = 2.65$, that contracts via propagating node deletion, obtained from numerical integration of the master equation (solid lines). The terms $\Delta_A(t)$ (dashed line) and $\Delta_B(t)$ (dotted line), which sum up to the derivative dS_t/dt, are also shown. Note that both $\Delta_A(t)$ and $\Delta_B(t)$ are negative at all times during the contraction process.

convergence for configuration model networks with degenerate degree distributions (random regular graphs), exponential degree distributions, and power-law degree distributions (scale-free networks) and showed that the relative entropy decreases monotonically to zero during the contraction process, reflecting the convergence of the degree distribution towards a Poisson distribution. Since the contracting networks remain uncorrelated, this means that their structures converge towards an Erdős-Rényi graph structure, substantiating earlier results obtained using direct integration of the master equation and computer simulations [31].

[1] S. Havlin and R. Cohen, *Complex Networks: Structure, Robustness and Function* (Cambridge University Press, New York, 2010).

[2] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).

[3] E. Estrada, *The Structure of Complex Networks: Theory and Applications* (Oxford University Press, Oxford, 2011).

[4] P. Erdős and A. Rényi, On random graphs I, Publ. Math. Debrecen **6**, 290 (1959).

[5] P. Erdős and A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).

[6] P. Erdős and A. Rényi, On the evolution of random graphs II, Bull. Int. Stat. Inst. **38**, 343 (1961).

[7] B. Bollobás, *Random Graphs*, 2nd ed. (Academic Press, London, 2001).

[8] M. Bauer and D. Bernard, Maximal entropy random networks with given degree distribution, arXiv:cond-mat/0206150 (2002).

[9] L. Bogacz, Z. Burda, and B. Wacław, Homogeneous complex networks, Physica A **366**, 587 (2006).

[10] G. Bianconi, The entropy of randomized network ensembles, Europhys. Lett. **81**, 28005 (2008).

[11] G. Bianconi, Entropy of network ensembles, Phys. Rev. E **79**, 036114 (2009).

[12] R. Albert, H. Jeong, and A.-L. Barabási, Diameter of the World-Wide Web, Nature (London) **401**, 130 (1999).

[13] S. Redner, How popular is your paper? An empirical study of the citation distribution, Eur. Phys. J. B **4**, 131 (1998).

[14] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys. **74**, 47 (2002).

[15] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, Science **286**, 509 (1999).

[16] P. L. Krapivsky, S. Redner, and F. Leyvraz, Connectivity of Growing Random Networks, Phys. Rev. Lett. **85**, 4629 (2000).

[17] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Structure of Growing Networks with Preferential Linking, Phys. Rev. Lett. **85**, 4633 (2000).

[18] A.-L. Barabási, Scale-free networks: A decade and beyond, Science **325**, 412 (2009).

[19] J. Török and J. Kertész, Cascading collapse of online social networks, Sci. Rep. **7**, 16743 (2017).

[20] L. Lőrincz, J. Koltai, A. F. Győr, and K. Takács, Collapse of an online social network: Burning social capital to create it? Social Netw. **57**, 43 (2019).

[21] L. Daqing, J. Yinan, K. Rui, and S. Havlin, Spatial correlation analysis of cascading failures: Congestions and blackouts, Sci. Rep. **4**, 5381 (2014).

[22] B. Schäfer, D. Witthaut, M. Timme, and V. Latora, Dynamically induced cascading failures in power grids, Nat. Commun. **9**, 1975 (2018).

[23] R. Pastor-Satorras and A. Vespignani, Epidemic Spreading in Scale-Free Networks, Phys. Rev. Lett. **86**, 3200 (2001).

[24] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, Rev. Mod. Phys. **87**, 925 (2015).

[25] R. Pastor-Satorras and A. Vespignani, Immunization of complex networks, Phys. Rev. E **65**, 036104 (2002).

[26] R. Albert, H. Jeong, and A.-L. Barabási, Error and attack tolerance of complex networks, Nature (London) **406**, 378 (2000).

[27] R. Cohen, K. Erez, D. ben-Avraham, and D. Havlin, Resilience of the Internet to Random Breakdowns, Phys. Rev. Lett. **85**, 4626 (2000).

[28] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Breakdown of the Internet under Intentional Attack, Phys. Rev. Lett. **86**, 3682 (2001).

[29] A. Braunstein, L. Dall'Asta, G. Semerjian, and L. Zdeborová, Network dismantling, Proc. Natl. Acad. Sci. U. S. A. **113**, 12368 (2016).

[30] L. Zdeborová, P. Zhang, and H.-J. Zhou, Fast and simple decycling and dismantling of networks, Sci. Rep. **6**, 37954 (2016).

[31] I. Tishby, O. Biham, and E. Katzav, Convergence towards an Erdős-Rényi graph structure in network contraction processes, Phys. Rev. E **100**, 032314 (2019).

[32] M. Molloy and B. Reed, A critical point for random graphs with a given degree sequence, Random Struct. Alg. **6**, 161 (1995).

[33] B. Molloy and A. Reed, The size of the giant component of a random graph with a given degree sequence, Combin. Prob. Comput. **7**, 295 (1998).

[34] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Random graphs with arbitrary degree distributions and their applications, Phys. Rev. E **64**, 026118 (2001).

[35] A. Annibale, A. C. C. Coolen, L. P. Fernandes, F. Fraternali, and J. Kleinjung, Tailored graph ensembles as proxies or null models for real networks I: Tools for quantifying structure, J. Phys. A **42**, 485001 (2009).

[36] E. S. Roberts, T. Schlitt, and A. C. C Coolen, Tailored graph ensembles as proxies or null models for real networks II: Results on directed graphs, J. Phys. A **44**, 275002 (2011).

[37] A. C. C. Coolen, A. Annibale, and E. S. Roberts, *Generating Random Networks and Graphs* (Oxford University Press, Oxford, 2017).

[38] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (North-Holland, Amsterdam, 2007).

[39] C. Gardiner, *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences*, 3rd ed. (Springer-Verlag, Berlin, 2004).

[40] J. C. Butcher, *Numerical Methods for Ordinary Differential Equations* (John Wiley & Sons, Chichester, UK, 2003).

[41] Trickle down: Diffuse downward through some hierarchical structure, in *The American Heritage Dictionary of Phrasal Verbs* (Houghton Mifflin Harcourt, New York, 2005).

[42] S. Kullback and R. A. Leibler, On information and sufficiency, Ann. Math. Stat. **22**, 79 (1951).

[43] I. Csiszár and P. C. Shields, Information theory and statistics: A tutorial, in *Foundations and Trends in Communications and Information Theory*, Vol. 1, issue 4, edited by S. Verdú (Now Publishers, Hanover, MA, 2004).

[44] S. Kullback, *Information Theory and Statistics* (Dover Publications, Mineola, NY, 1969).

[45] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley, Hoboken, NJ, 2006).

[46] C. E. Shannon, A Mathematical theory of communication, Bell Syst. Tech. J. **27**, 379 (1948).

[47] J. E. Shore and R. W. Johnson, Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy, IEEE Trans. Inf. Theory **26**, 26 (1980).

[48] F. W. J. Olver, D. M. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions* (Cambridge University Press, New York, 2010).

[49] A. Rényi, On measures of information and entropy, in *Proceedings of the Fourth Berkeley Symposium on Mathematics, Statistics and Probability, June 20–July 30, 1960*, edited by J. Neyman (University of California Press, Berkeley, 1961), Vol. I, pp. 547–561.

[50] M. M. Deza and E. Deza, *Encyclopedia of Distances*, 4th ed. (Springer-Verlag, Berlin, 2016).

[51] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes* (Holden-Day, San Francisco, 1964).

[52] I. Csiszár, Information-type measures of difference of probability distributions and indirect observations, Studia Sci. Math. Hungar. **2**, 299 (1967).

[53] S. Kullback, A lower bound for discrimination information in terms of variation, IEEE Trans. Inf. Theory **13**, 126 (1967).

[54] I. Vajda, Note on discrimination and variation, IEEE Trans. Inf. Theory **16**, 771 (1970).

[55] J. H. Morrison and P. R. Hof, Life and death of neurons in the aging brain, Science **278**, 412 (1997).

[56] M.-T. Heemels, Neurodegenerative diseases, Nature (London) **539**, 179 (2016).

[57] T. Arendt, M. K. Brückner, M. Morawski, C. Jäger, and H.-J. Gertz, Early neurone loss in Alzheimer's disease: Cortical or subcortical? Acta Neuropathol. Commun. **3**, 10 (2015).