Degree-Preserving Distance Compression and Topological Compressibility of Complex Networks

Jian-Hui Li(李建辉)[1,2,3], Zu-Guo Yu(喻祖国)[1,2*], and Yu-Chu Tian(田玉楚)[3*]

[1]*National Center for Applied Mathematics in Hunan, Xiangtan University, Hunan 411105, China*
[2]*Key Laboratory of Intelligent Computing and Information Processing of the Ministry of Education of China, Xiangtan University, Hunan 411105, China*
[3]*School of Computer Sciences, Queensland University of Technology, GPO Box 2434, Brisbane, QLD 4001, Australia*

(Received 10 August 2025; accepted manuscript online 10 November 2025)

Accurately modeling real network dynamics is a grand challenge in network science. The network dynamics arise from node interactions, which are shaped by network topology. Real networks tend to exhibit compact or highly optimized topologies. But the key problems arise: how to compress a network to best enhance its compactness, and what the compression limit of the network reflects? We abstract the topological compression of complex networks as a dynamic process of making them more compact and propose the local compression modulus that plays a key role in effective compression evolution of networks. Subsequently, we identify topological compressibility—a general property of complex networks that characterizes the extent to which a network can be compressed—and provide its approximate quantification. We anticipate that our findings and established theory will provide valuable insights into both dynamics and various applications of complex networks.

DOI: 10.1088/0256-307X/42/12/120003 CSTR: 32039.14.0256-307X.42.12.120003

An open challenge in network science is to model and predict network dynamics.[1] The rich dynamics of complex networks arise not only from the behavior of individual nodes[2,3] but also from their interactions. Network topology plays a crucial role in shaping these intricate interactions.[4] A compact topology of complex networks is one important factor for diverse dynamical processes of complex systems across various domains such as physics, biochemical sciences, and engineering.[5] It can be characterized by a small average distance of the network, often implying high network efficiency or performance. A significant advantage of the compact network topology in complex systems is their efficient transport of information.[6–12]

The significance of compact topology has shown in other complex systems like microorganisms in the body,[13–15] cellular networks,[16] and sensor networks.[17,18] More specifically, networks exhibiting small-world[19] characteristics demonstrate high efficiency of information transport.[20]

Hence, reducing the average path length of a network is a significant driver for improving functional properties such as communication efficiency, motivating the search for effective compression approaches. For many topological optimization problems in real-world networks, however, preserving the degree distribution is a crucial constraint, as these networks are typically engineered to meet underlying design specifications, such as the topology optimization of communication networks for smart grids in the engineering field.[10] Furthermore, for certain network topology analyses, it is essential to keep node degrees unchanged. This practice isolates the influence of other topological features on network function.[21–23] Hence, the central objective of this study is to achieve degree-preserving network compaction.

Naturally, insightful questions emanate from the above understanding: how to make a complex network more compact and what the limit of network compression can reflect? Intuitively, how much a network can be compressed depends on how compact the network currently is. If a network is already compact, further compression would be a challenge. Ultra-small-world[5] networks are an example of such networks with a compact topology.

In this Letter, we address the aforementioned significant issues widespread in complex networks. Overall, we develop an effective compression approach that can be used to analyze the topological compressibility of complex networks. Serving as a critical foundation, we establish a theory of increase-upper and decrease-lower bounds in network average distance variations resulting from edge removal and addition. Built upon the theory, we present the local compression modulus (LCM) to guide efficient topological compression via edge rewiring steps. Hence, the theoretical results shed light on a generalized evolutionary principle to compress the entire network the most. Subsequently, we identify topological compressibility as an inherent and general property of complex networks. It quantifies the extent to which a network can be compressed.

*Corresponding authors. Email: yuzg@xtu.edu.cn; y.tian@qut.edu.au
© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

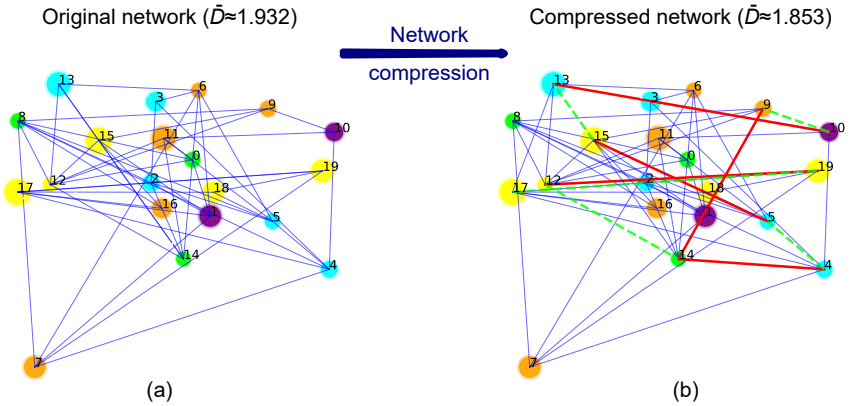

Fig. 1. A diagram of network topological compression. (a) Original network. (b) Compressed network. In (b), the dashed lines and the red solid lines represent the edges removed from the original network and the edges added in the compressed network, respectively.

Network Topological Compressibility. For complex networks, their topological compression is abstracted as a dynamic process to tighten them with invariance in the involved nodes, the number of edges, network connectivity, and degree distribution. It reshapes the topology of the network, i.e., transforms the topology for a reduced network average distance, such as an example in Fig. 1. The compression does not segment the network or alter the ensemble to which the network belongs. It corresponds to a series of edge rewiring[21,24,25] to make the network average distance shorter while keeping the invariance in the network properties mentioned above.

For a network, once it can be compressed to its limit, its topological compressibility can be defined. Hence, we can give the metric that can be used to characterize the topological compressibility of a complex network $G(V,E)$. Let $G_{\rm lim}(V_{\rm lim},E_{\rm lim})$ be the limit compression network of $G(V,E)$. Thus, the metric of network topological compressibility Ψ_G is defined by

$$\Psi_G = \frac{\bar{D}(G)-\bar{D}(G_{\rm lim})}{\bar{D}(G)}.$$

As a general property of complex networks, the topological compressibility is universal to all networks, characterizing the extent to which a network can be compressed. It is the relative reduction rate in average distance from the original network topology to its compression limit. The quantified value of topological compressibility varies with the current compactness of the network. The higher the topological compressibility value is, the more the network can be compressed.

However, obtaining the compression limit is challenging for a network. This makes the theoretical computation of topological compressibility practically infeasible. Here, we design a method to approximate the topological compressibility of complex networks.

Local Compression Modulus of Complex Networks. We consider complex networks represented by a connected simple graph denoted by $G(V,E)$, where V and E are the sets of vertices and edges, respectively. We focus on reducing the distances between node pairs by a rewiring process of network topological reshaping for network compression. The challenge in this process is: which edge is cut and where to add a new edge towards our goal of maximizing network compression?

To tackle this challenge, we present the following observations, which are insightful to our theoretical developments. Firstly, following the removal or addition of an edge in a network, some node pairs will become further away or closer whereas some other node pairs still keep their distances unchanged (Figs. 1a and 1b in Ref. [25]). Regardless, cutting an edge will always lead to an increase in the average distance between all node pairs. Conversely, adding a new non-duplicate edge between two existing nodes will always shorten the average distance of all node pairs.

How much will the average distance of a network increase or decrease when an edge is cut or added? We have established a theory to quantitatively characterize the distance variation. Consider edge $e(a,b)\in E$ between node pair (a,b). Let $B_G(e)$ denote the betweenness of e. Following the removal of edge $e(a,b)$, the increase in the sum of the distances between all pairs of nodes is bounded by $[B_G(e)+2][d_{G-e}(a,b)-1]$, where $d_{G-e}(a,b)$ is the distance between node pair (a,b) after edge $e(a,b)$ is removed from G. This implies that the increase of the network average distance due to the removal of edge e is bounded above by $\frac{1}{|V|(|V|-1)}[B_G(e)+2][d_{G-e}(a,b)-1] \triangleq \underline{\Delta D_e^-}$ (Theorem 1 in Ref. [25]), where $|V|$ is the size of V.

If there is no edge between node pair (a,b), then following the addition of such a new edge, the total reduction in the sum of the distances between all node pairs is bounded below by $-[B_{G+e}(e)+2]\Delta d_{\max,e}$, where e is the new edge that joins node pair (a,b), $B_{G+e}(e)$ is computed in $G+e$ for this new edge e, $\Delta d_{\max,e}=\max\{\Delta d(v_l,v_m)\mid l,m\in\{1,\cdots,|V|\},l\neq m,\Delta d(v_l,v_m)\neq 0\}$, $\Delta d(v_l,v_m)<0$ is the distance variation due to the addition of e for node pair (v_l,v_m). Consequently, the amount of network average distance reduction due to the addition of e is bounded below by $-\frac{1}{|V|(|V|-1)}[B_{G+e}(e)+2]\Delta d_{\max,e} \triangleq \underline{\Delta D_e^+}$ (Theorem 2 in Ref. [25]).

Generally, to reduce the average distance $\bar{D}$ of a network the most through rewiring, the best performance is achieved by removing the edge with the smallest $\overline{\Delta D_e^-}$, and then adding a new non-duplicate edge that gives the largest $\underline{\Delta D_e^+}$ in each iterative rewiring step. A smaller $B_G(e)d_{G-e}(a,b)$ generally means a smaller $\overline{\Delta D_e^-}$ when removing an edge, and a larger $B_{G+e}(e)d_G(a,b)$ corresponds to a larger $\underline{\Delta D_e^+}$ when adding a new edge.

For simplicity of presentation, in the following we use $B(e)$ to denote both $B_G(e)$ and $B_{G+e}(e)$ when there is no confusion. For the edge removal process, $B(e)$ indicates the edge betweenness before the edge removal (i.e., $B_G(e)$); whereas in the edge addition process, $B(e)$ represents the edge betweenness after the edge addition (i.e., $B_{G+e}(e)$). Similarly, we use $d(a,b)$ to denote both $d_{G-e}(a,b)$ for the edge removal process and $d_G(a,b)$ for the edge addition process. In the edge removal process, $d(a,b)$ is the node pair distance after edge removal (i.e., $d_{G-e}(a,b)$); while in the edge addition process, $d(a,b)$ represents the distance of the node pair before the addition (i.e., $d_G(a,b)$). Finally, we call $B(e)d(a,b) \triangleq \psi(e(a,b))$ the LCM respect to edge $e(a,b)$ in a complex network.

Built on these theoretical results, we have developed an approach to effectively compress complex networks and further approximate their topological compressibility. The approach is applicable to general complex networks. It is further extended to more special complex networks with node constraints.

Topological Compressibility of General Networks. Following the definition of topological compressibility of a complex network, for a given complex network $G(V,E)$, we need to obtain the compression limit of G and then calculate its topological compressibility Ψ_G. However, obtaining the compression limit presents a significant challenge. Thus, we employ the evolved network through the effective network compression approach [see the Supplementary Materials (SM)] to approximate the limit network, thereby providing an approximate measure $\hat{\Psi}_G$ of the compressibility.

The well-known models of complex networks include the Barabási–Albert (BA) scale-free model[26] and the Watts–Strogatz (WS) small-world model.[27] They describe topological characteristics of numerous real-world complex networks. We consider these two types of complex networks because they are commonly believed to be very compact with small $\bar{D}$. Hence, to investigate whether or not they can be further compressed, we apply our approach to compress these networks, and then quantify their topological compressibility.

We examine the compression results and topological compressibility of these networks based on two indexes, i.e., $\bar{D}$ and $\hat{\Psi}_G$. The smaller the value of $\bar{D}$ is, the more compact the network topology is. The bigger the value of $\hat{\Psi}_G$ is, the higher the network topological compressibility is. A higher topological compressibility implies that the network is more compressible.

Figure 2 illustrates the computational results of these networks under various settings of model parameters. It is seen from Figs. 2(a) and 2(b) that the compression approach leads to reduced $\bar{D}$ for both scale-free and small-world networks, indicating compressed network topology for these networks. The observed small variations in $\bar{D}$ following compression indicate limited topological compressibility across these networks. Despite their lower topological compressibility, these networks are not incompressible.

We approximate the topological compressibility; the corresponding results are shown in Figs. 2(d) and 2(e). They show that the topological compressibility of these small-world networks is between $[0.05, 0.07]$. In comparison, the value of scale-free networks is about $[0.01, 0.04]$. The topological compressibility of scale-free networks is lower than that of small-world networks. The observed effect stems from the intrinsic tightness of scale-free networks. Therefore, from the perspective of topological compressibility, scale-free networks are more compact and thus less compressible than small-world networks.

Moreover, we discuss the case of the Erdös–Rényi (ER) networks. The experimental results are shown in Figs. 2(c) and 2(f). The results show that the ER network exhibits a topological compressibility intermediate between that of BA and WS networks, reflecting a similar ranking in terms of structural compactness.

Demonstration of Node-Constrained Networks. To demonstrate the compression evolution of node-constrained networks, we constructed multi-population networks, in which each population module is a WS network of equal size (refer to Sec. IV A in Ref. [25] for details). More specifically, we investigate two-, three-, four-, and five-population networks.

Consider a multi-population network $G_{\mathrm{MP}}(V,E)$ consisting of N_S population modules $P_s^{\mathrm{M}}(V_s^{\mathrm{M}}, E_s^{\mathrm{M}})$, each with $|V_1^{\mathrm{M}}|$ nodes, where $s \in \{1,\ldots,N_S\}$. As a toy example, we assign a free value $\delta = s$ to the node $v \in V_s^{\mathrm{M}}$. Then, we define the constrained function $f_{v_i(\delta_i)}(v(\delta))$ for a fixed node v_i as a binary population-similarity constraint function, i.e., $\forall\ i,j \in \{1,\ldots,|V|\}, f_{v_i(\delta_i)}(v_j(\delta_j)) = 1$ if $\delta_j = \delta_i$, or $f_{v_i(\delta_i)}(v_j(\delta_j)) = 0$ otherwise. A node subset $V_i \subset V$, with respect to v_i, is constructed through the node constraint condition, i.e., $V_i = \{v_j | v_j \in V, f_{v_i(\delta_i)}(v_j(\delta_j)) < c_i, j = 1,2,\ldots,|V|\}$, where $c_i < 1$ is a positive number. Simply put, we can set all c_i as the same number.

We apply our approach to compress these node-constrained multi-population networks. The experimental results are shown in Fig. 3. Figure 3(a) shows that these networks are compressed rapidly with the increase of P_{rew} when $P_{\mathrm{rew}} < 0.2$. After P_{rew} reaches 0.2, these networks are approaching their compression limit. The lower the compactness of the network, the higher its topological compressibility $\hat{\Psi}$; see Fig. 3(b). The figure shows that $\hat{\Psi}$ reaches almost 45% for five-population networks. Despite the higher time complexity[28] to calculate the LCM, incorporating node constraints can effectively reduce the algorithm's runtime; see Fig. 3(c). At each step, the two-population network exhibits the smallest search space, resulting in minimal running time.

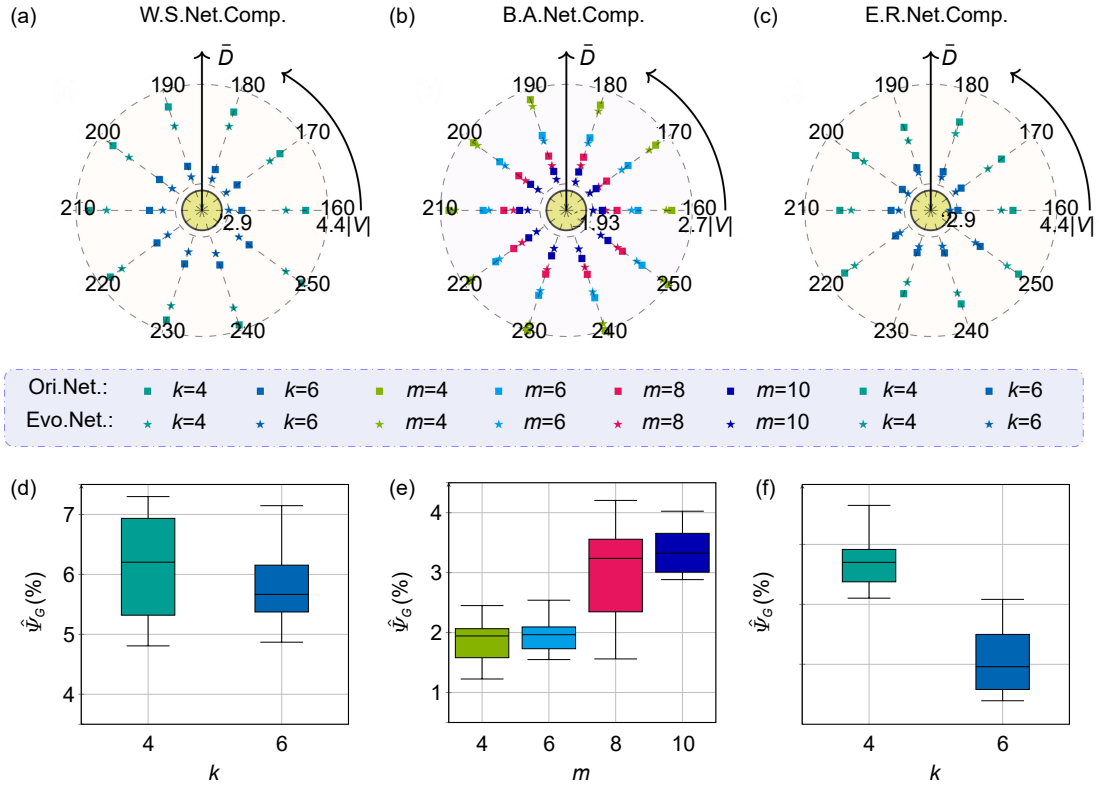

Fig. 2. Validation of topological compressibility on WS, BA and ER networks. For WS networks, we generate 10 networks of different sizes for each $k = 4, 6$. For BA networks, we also generate 10 networks of different sizes for each $m = 4, 6, 8, 10$. For ER networks, we generate ER networks with different sizes. The parameter k follows the relationship established in Ref. [25] (Sec. IV, part D). Then, use our approach to compress theses networks, and quantify their topological compressibility. (a)–(c) Illustrate the gap in $\bar{D}$ before and after compression for all networks. (d)–(f) Display the $\hat{\Psi}_G$ (%) of these synthetic networks, respectively.

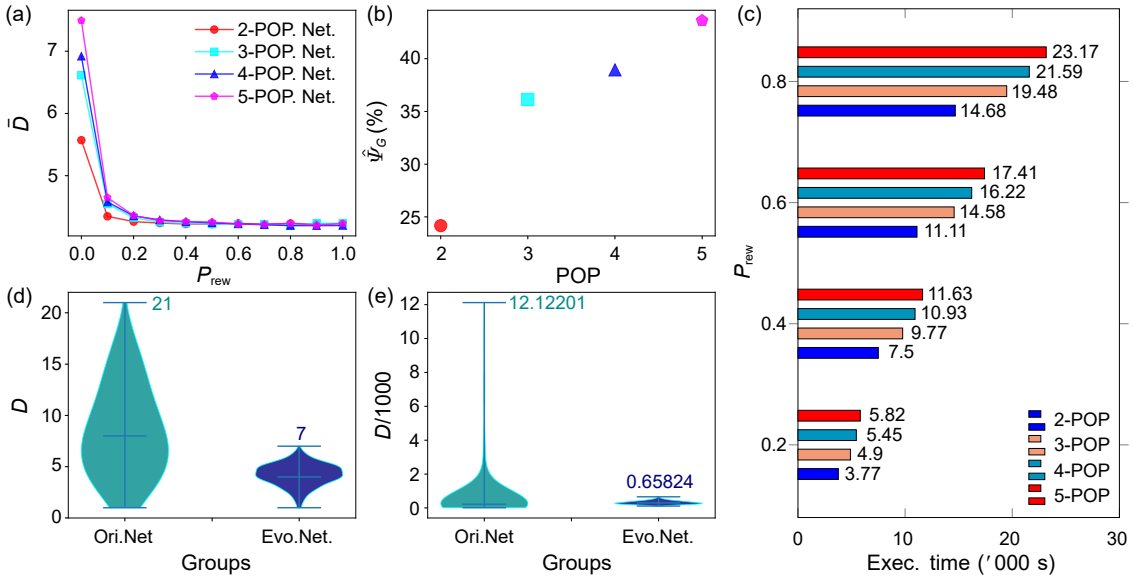

Fig. 3. Demonstration of compression evolution of node-constrained networks. We generate some multi-population networks with 300 nodes. We set the numbers of populations to be $N_S = 2, 3, 4, 5$ and the sizes of the population $|V_1^M|$ such that $N_S \times |V_1^M| = 300$. The connections between different population modules are set to two. (a) The impact of the evolution fraction (P_{rew}) on the network average distance ($\bar{D}$). (b) The approximate topological compressibility of multi-population networks. We set $P_{rew} = 0.2, 0.4, 0.6, 0.8$ to compress these networks. The execution time of each experiment is shown in (c). (d) and (e) The distribution of the distances of all node pairs, and the distribution of the edge betweenness of the two-population network, respectively, where $P_{rew} = 0.6$.

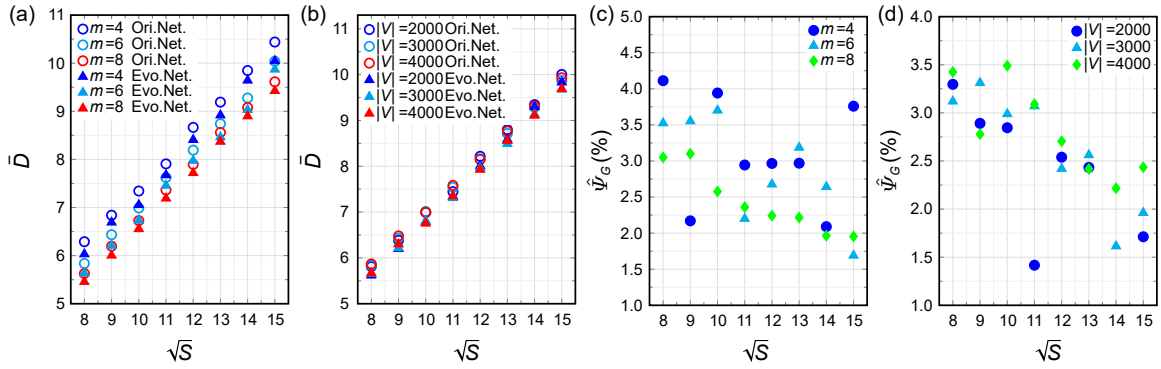

Fig. 4. The $\bar{D}$ and $\hat{\Psi}_G$ versus the side length ($\sqrt{S}$) of the geographic area for SF-NANs. We generate SF-NANs in a geographic area with variable settings of $|V|$, $\sqrt{S}$, and the number of nodes (m) that a node is attached to. Then, using our approach to compress these networks. After that, the $\bar{D}$ and $\hat{\Psi}_G$ of the resulting networks are calculated. (a) and (c) The results for the fixed parameter $|V| = 3000$. (b) and (d) The results for the fixed parameter $m = 6$.

Table 1. Real-world network results summary.

Network	$\|V\|$	$\|E\|$	$\bar{D}$	Network description	$\hat{\Psi}_G$ (%)
Net. 1	175	258	5.2717	The PIN for S. cerevisiae	22.926
Net. 2	495	1 060	4.4076	The PIN for C. elegans	16.279
Net. 3	297	2 148	2.4553	The Caenorhabditis elegans worm's neural network	1.418
Net. 4	95	1170	1.8645	Connection matrices of cat cortex	3.976
Net. 5	71	438	2.2447	Macaque cortical connectivity	9.681

We also analyze the compression from an internal structural perspective. Figures 3(d) and 3(e) present the distance distribution among all node pairs and the betweenness distribution of edges, respectively. These plots clearly show significant compression of the network in node-pair distances. The maximum node-pair distance is reduced from over 21 down to about 7, and the distance distribution has been compressed as shown in Fig. 3(d). With the compression of the network, the maximal edge betweenness is compressed from the order of 10^5 down to 10^3. The resulting small betweenness indicates the nonexistence of many significant edges after network evolution. The network after evolution seems to have multiple paths of similar distances between node pairs. This implies that multiple paths of similar lengths can be used between node pairs for communications. This is particularly beneficial for improving network functional reliability—if one path fails, an alternative path of a similar length can work to support network function.

Topological Compressibility of Smart-Grid Neighborhood Area Networks. Neighborhood area networks (NANs) are communication infrastructures in the smart grid. Previous work has investigated how to establish scale-free NANs (SF-NANs) for a given set of NAN nodes, as well as their connectivity properties and requirements.[29]

We use the SF-NAN model to demonstrate the topological compressibility of smart-grid communication networks. The compression of SF-NANs must be conducted under node-based constraints. The node-free vectors and constraint functions are set as the coordinates of nodes and Euclidian distance, respectively. The results are illustrated in Fig. 4. It can be seen from Fig. 4 that SF-NANs are compressible. Under various settings for the number of NAN nodes ($|V|$) and the size of NAN geographical area (S), the topological compressibility of NANs is close to that of general scale-free networks.

Topological Compressibility of Real-World Networks. We have examined the topological compressibility in several real-world networks from various fields, such as protein interaction networks (PINs) and brain neural networks. The corresponding network information and experimental results are presented in Table 1.

As shown in Table 1, we have renamed these networks. The data source for all networks are as follows:

- Net1, Net2: https://people.duke.edu/~jmoody77/Prot/index.htm
- Net3: https://toreopsahl.com/datasets/
- Net4, Net5: https://sites.google.com/site/bctnet/datasets-and-demos

The results reveal a marked contrast in topological compressibility: it is significantly larger in PINs but smaller in brain networks. The small compressibility observed in brain networks implies a more compact topology that correlates with high efficiency in information transport.

In summary, we explored topological compression—defined as distance reduction under the constraint of degree distribution preservation—and further studied topological compressibility. The LCM plays a key role in guiding effective network compression. Based on this metric, we propose two distinct approaches for network topological compression, each tailored to different operational scenarios. We comprehensively quantified the topological compressibility exhibited by various classes of networks. We found that more compact networks exhibit significantly smaller topological compressibility, rendering them

more resistant to compression. Conversely, looser network structures demonstrate markedly larger compressibility, making compression more readily achievable. Random, small-world, and scale-free networks, which are commonly considered topologically compact, all admit compression, albeit with limited compressibility. Scale-free networks, which possess the least topological compressibility, necessitate more potent strategies for effective compression. The observed ultra-compact topology provides a functional advantage as it is directly linked to enhanced performance in network dynamics such as information propagation. Our findings also serve as a wellspring of inspiration for tailoring node constraints to optimize network structure and performance; offering useful guidance in this realm that will be explored in our future work.

Acknowledgements. The authors are particularly grateful to Professor Marian Boguñá for his insightful comments and suggestions. This work was supported in part by the National Natural Science Foundation of China (Grant No. 12371088), the Innovative Research Group Project of Natural Science Foundation of Hunan Province of China (Grant No. 2024JJ1008), and in part by the Australian Research Council (ARC) through the Discovery Projects scheme (Grant No. DP220100580).

References

[1] Papaefthymiou E S, Iordanou C, and Papadopoulos F 2024 *Phys. Rev. Lett.* **132** 257401
[2] Barzon G, Artime O, Suweis S, and Domenico M D 2024 *Proc. Natl. Acad. Sci. USA* **121** e2317608121
[3] Ghavasieh A and Domenico M D 2024 *Nat. Phys.* **20** 512
[4] Zhu C L, Shen X Y, Zhu G M, and Li B W 2023 *Chin. Phys. Lett.* **40** 124402
[5] Cohen R and Havlin S 2003 *Phys. Rev. Lett.* **90** 058701
[6] Seguin C, Sporns O, and Zalesky A 2023 *Nat. Rev. Neurosci.* **24** 557
[7] Hahn G, Ponce-Alvarez A, Deco G, Aertsen A, and Kumar A 2018 *Nat. Rev. Neurosci.* **20** 117
[8] Doğaner B A, Yan L K Q, and Youk H 2016 *Trends Cell Biol.* **26** 262
[9] Cooper G M and Hausman R E 2010 (*Sunderland, MA: Sinauer Associates*)
[10] Li J H, Ding Y M, Yu Z G, Zhang W Z, Ren H L, and Tian Y C 2025 *IEEE Trans. Consum. Electron.* **70** 2726
[11] Miller M B and Bassler B L 2001 *Annu. Rev. Microbiol.* **55** 165
[12] Saputro N and Akkaya K 2017 *IEEE Internet Things J.* **4** 894
[13] Papenfort K and Bassler B L 2016 *Nat. Rev. Microbiol.* **14** 576
[14] Zarkan A, Liu J Y, Matuszewska M, Gaimster H, and Summers D K 2020 *Trends Microbiol.* **28** 566
[15] Stephens K and Bentley W E 2020 *Trends Microbiol.* **28** 633
[16] Rowland M A, Greenbaum J M, and Deeds E J 2017 *Nat. Commun.* **8** 16009
[17] Kim J, Campbell A S, de Ávila B E F, and Wang J 2019 *Nat. Biotechnol.* **37** 389
[18] Sempionatto J R, Lin M Y, Yin L, De la Paz E, Pei K X, Sonsa-Ard T, de Loyola Silva A N, Khorshed A A, Zhang F Y, Tostado N *et al.* 2021 *Nat. Biomed. Eng.* **5** 737
[19] Zou Z Y, Mao B H, Hao H M, Gao J Z, and Yang J J 2009 *Chin. Phys. Lett.* **26** 110502
[20] Achard S, Salvador R, Whitcher B, Suckling J, and Bullmore E D 2006 *J. Neurosci.* **26** 63
[21] Gallos L K, Makse H A, and Sigman M 2012 *Proc. Natl. Acad. Sci. USA* **109** 2825
[22] Li J H, Yu Z G, Anh V V, Liu J L, and Peng A Q 2023 *Fractals* **31** 2350089
[23] Hou L, Pan X, Liu K C, Yang Z M, Liu J G, and Zhou T 2023 *iScience* **26** 105893
[24] Lynn C W, Holmes C M, and Palmer S E 2024 *Nat. Phys.* **20** 484
[25] Li J H, Yu Z G, and Tian Y C 2025 arXiv:2508.04029 [math.GN]
[26] Barabási A L and Albert R 1999 *Science* **286** 509
[27] Watts D J and Strogatz S H 1998 *Nature* **393** 440
[28] Brandes U 2001 *J. Math. Sociol.* **25** 163
[29] Ding Y M, Li X H, Tian Y C, Ledwich G, Mishra Y, and Zhou C J 2018 *IEEE Trans. Smart Grid* **10** 4245