Letter

Robustness assessment of complex networks using the idle network

Marcus Engsig [*]
Department of Science and Engineering, Sorbonne University Abu Dhabi, Abu Dhabi, United Arab Emirates and Directed Energy Research Centre, Technology Innovation Institute, Abu Dhabi, United Arab Emirates

Alejandro Tejedor [†]
Department of Science and Engineering, Sorbonne University Abu Dhabi, Abu Dhabi, United Arab Emirates and Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, California 92697, USA

Yamir Moreno [‡]
Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, 50018 Zaragoza, Spain; Department of Theoretical Physics, University of Zaragoza, Zaragoza 50009, Spain; and CENTAI Institute, Turin 10138, Italy

(Received 13 July 2022; accepted 6 December 2022; published 26 December 2022)

Network robustness is an essential system property to sustain functionality in the face of failures or targeted attacks. Currently, only the connectivity of the nodes resilient to an attack is used to assess robustness. We propose to incorporate the properties of the emerging connectivity of the nodes that are affected by the attack (idle network), which is demonstrated to contain relevant information about network robustness, improving the accuracy of its assessment. Our work shows that the information contained in the idle network offers a potential to generalize models, enabling them to estimate robustness for unseen attacks.

DOI: 10.1103/PhysRevResearch.4.L042050

The representation of complex systems as networks, where system components are abstracted as nodes and their interactions as links, has allowed us to advance our understanding of the structure and dynamics of such systems in fields as diverse as biology, engineering, economics, and geosciences [1–9]. Particularly, network theory has been instrumental in developing methodologies to assess the robustness of interconnected systems such as power grids, the internet, and airports, in the face of random failures or targeted attacks [10–16].

The robustness of a network can be defined as its ability to maintain functionality whilst undergoing an attack. In a world where critical infrastructures and their connectivities are potential targets of malicious attacks, it is paramount to identify what are the key network properties that determine network resilience to a given attack. Since the pioneering work by *Albert et al.* [17], a vast literature has presented methodologies and metrics to quantify network robustness [11,12,17–24]. However, current methodologies focus mainly on the connectivity of the nodes that remain unaffected (*active network*) by the attack, while the connectivity of the affected nodes (*idle network*) has received minimal attention [25]. Here, we demonstrate the benefit of including information about the idle network in assessing network robustness.

Let us formally define the active and idle networks, which naturally emerge from an attack process acting on a network [25]. Attacking a network is synonymous to a process of sequential node removal. Consider an initial network $\mathcal{N}$ that consists of N nodes, denoted $\{n_i\} : i = 1, ...N$, connected by a set of links $\{(n_i, n_j)\}$. The sequential node removal process starts at $t = 0$ with the original network $\mathcal{N}$, and an attack strategy $D(\mathcal{N})$. For every discrete time step $t > 0$, the attack eliminates a chosen node n_i and all of its corresponding links $(n_i, \cdot)$, resulting in a new network, formed by the sets of nodes and links that are unaffected by the attack; we denote this the active network $\mathcal{N}_A(t)$. The attack process also gives rise to the idle network $\mathcal{N}_I(t)$ (the projection of the attack on the network $\mathcal{N}$), which consists of the entire set of nodes removed from the network $\mathcal{N}$ up to time t, and the links originally existing among them (see Fig. 1). We can mathematically express a given attack strategy D acting on a network $\mathcal{N}$, as the decomposition of $\mathcal{N}$ into the active $\mathcal{N}_A(t)$ and idle $\mathcal{N}_I(t)$ networks:

$$D : \mathcal{N} \rightarrow \{\mathcal{N}_A(t), \mathcal{N}_I(t)\}, \quad t = 1, ...N. \tag{1}$$

It is clear that with respect to the nodes, the active and idle networks are complementary, implying that the union of the nodes in $\mathcal{N}_A(t)$ and $\mathcal{N}_I(t)$ is the set of nodes in $\mathcal{N}$. However, this is not the case for the connectivity of the nodes, as it is neither complementary nor symmetric. When a node is removed, all of its links are removed from the active network. Yet,

[*]marcus.w.engsig@gmail.com
[†]alej.tejedor@gmail.com
[‡]yamir.moreno@gmail.com

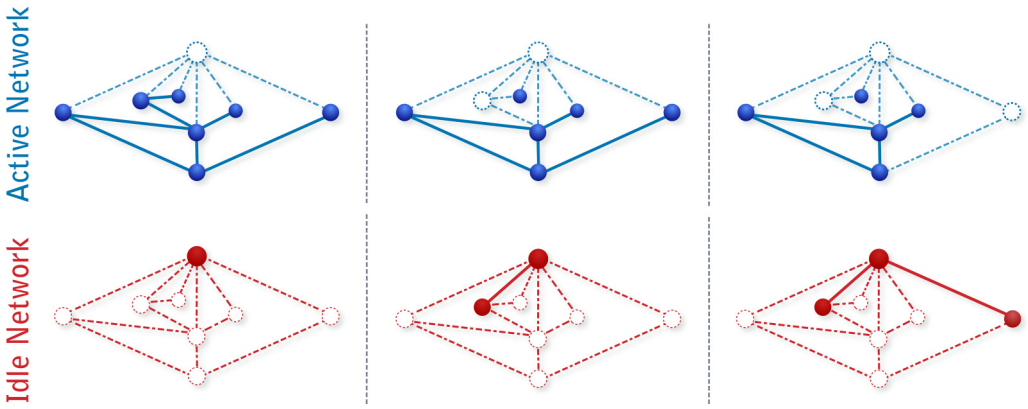

FIG. 1. Schematic illustration of the active and idle networks at different stages of an attack.

from the set of links removed by the attack, only the subset that connects affected (removed) nodes is included in the idle network. We argue that the information about the connectivity of the affected nodes by an attack, which is just available in the idle network, provides important information on the effectiveness of the attack, and therefore on the robustness of the attacked network. Thus, our research hypothesis can be summarized as follows: *There exists pertinent and readily available information on the robustness of a network undergoing an attack in the corresponding idle network structure.*

To test this hypothesis, we extract indicators from the active and idle networks to benchmark our capacity to assess network robustness using only active indicators (traditional approach) versus incorporating idle indicators as well. Particularly, we choose two simple indicators to model robustness: (i) The largest cluster size C, defined as the ratio of the number of nodes in the largest cluster (set of connected nodes) over the number of nodes N in the initial network, quantifies the effect of the attack in breaking down (building up) the active (idle) network in terms of its size. Note that this metric does not encompass the effectiveness of the connectivity of these networks. (ii) The link fraction L is the number of links in the active (idle) network, normalized by the number of links in the initial network $\mathcal{N}$. This indicator describes how the attack removes (adds) links and thus provides information about how well-connected the nodes are in the active (idle) network. These indicators were chosen such that, in complement, they have information on the overall functionality of the network, and therefore on its robustness.

Network robustness, as the network's capacity to maintain functionality whilst undergoing an attack, is a complex, *a priori* unknown function of both the specific attack strategy and network properties. In this study, we use the efficiency E as a proxy for robustness. Recall that E of a network $\mathcal{N}$ with N nodes is defined as the standardized sum of the reciprocal of the shortest paths $d_{i,j}$ between all pair of nodes i and j:

$$E = \frac{1}{N(N-1)} \sum_{i,j \in \mathcal{N}, i \neq j} \frac{1}{d_{i,j}}. \tag{2}$$

In our study, E is normalized to always start at 1, by dividing the efficiencies computed at the different stages of an attack by the value of the efficiency for the intact network. The choice of E as the proxy for robustness is threefold: (i) it is a well-established proxy of robustness [26–29], (ii) it avoids circular reasoning in testing our hypothesis, as it is a function of only the active network, and (iii) its high complexity offers us a playground to emulate realistic network robustness assessments, where low complexity indicators are necessary. Importantly, the framework and results presented in this work are generalizable for different proxies of robustness as they are not particular to the choice of efficiency.

Given the two indicators and the proxy for robustness, we transform our hypothesis into a regression problem. Thus, we evaluate the difference in estimation accuracy achieved via a neural network when only active indicators are included in the training set, and when idle indicators are also included. We use a forward-feeding and back-propagating artificial neural network with 3 hidden layers of 10 neurons per layer, each with ReLu activation functions; set to optimize validation squared residual loss. Each neural network was implemented with a dataset of 200 attack sequences, with a $6:1:1$ train, test, validation split. The output of the neural network is the estimation of the efficiency as the proxy for robustness. In order to verify our hypothesis, the estimation accuracy must increase when the neural network is granted the active and idle indicators, compared to the estimation produced using the active indicators alone.

Our study investigates different stochastically generated synthetic network topologies and attack strategies to test our hypothesis systematically. Namely, we test the robustness estimation for random (Erdős-Rényi [30]), scale-free (using a configuration model [31]), and small world (Strogatz-Watts [1]) topologies, undergoing three different attack strategies: targeted (degree), random failure, and random spreading [25]. Furthermore, the different topologies were explored for varying initial link densities, as characterized by $\bar{k}$ (average degree of the initial network $\mathcal{N}$). The tested link densities for all of the synthetic topologies correspond to $\bar{k} \in \{3, 6, 12, 24\}$. Thus, we have explored 36 combinations of topologies, attacks, and link densities. For each of these combinations, 200 different stochastic topologies were generated and exposed to a full attack evolution, where the indicators and efficiency were calculated at the different stages of the attack (see Supplemental Material (SM) [32]).

Figure 2 displays a representative case to illustrate our results. Note that the sum of the square residuals (*SSR*) has been used to directly compare the performance of the assessments

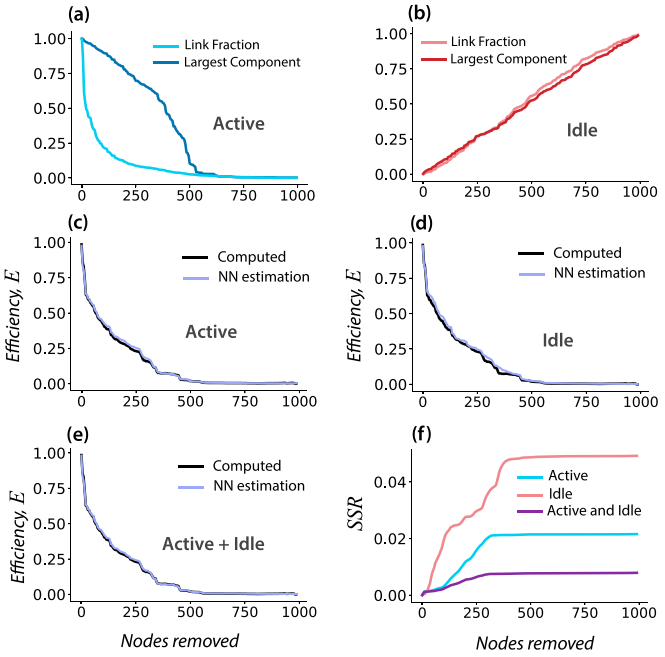

FIG. 2. Evolution of the (a) active and (b) idle indicators (largest component and link fraction) for a scale-free network with $N = 1000$ nodes ($\bar{k} = 12$), undergoing a degree attack. Estimation of the network efficiency as a function of the attack stage via a neural network using indicators from (c) only active, (d) only idle, or (e) both active and the idle networks. The computed values of network efficiency using Eq. (2) are displayed for comparison in panels (c)–(e). (f) Sum of the squared residuals (*SSR*) as a function of the attack stage computed for a testing set.

by the artificial neural network [33–35] depending on the set of indicators used for the regression (all the evaluations of model performance are done over full attack sequences, and thus comparing *SSR* is equivalent to comparing mean square error or root mean square error). As expected, the estimation of network robustness using active indicators [Fig. 2(a)] is quite accurate [Fig. 2(c): $SSR = 0.02 \pm 0.01$]. Noticeably, the idle indicators alone [Fig. 2(b)] allowed us to estimate fairly well ($SSR = 0.06 \pm 0.05$) the trend of the evolution of the efficiency as shown in Fig. 2(d). Most importantly, as hypothesized, by combining the indicators of the active and idle networks [Fig. 2(e)], we obtained a more accurate estimation of network robustness ($SSR = 0.01 \pm 0.006$). Note that the difference between the *SSR* obtained from the model trained with only active indicators and the model trained with both active and idle indicators can be interpreted as the marginal reduction in the *SSR* obtained by the addition of idle indicators. Additionally, the increased accuracy in the estimation is consistent throughout all stages of the attack [Fig. 2(f)]. Our results for the whole data set of network topologies and attacks demonstrate systematically that active indicators, when combined with idle indicators, increase the accuracy in the estimation of robustness from 20% to 900% depending on the topology and attack, verifying our hypothesis (see the SM).

As shown in detail in the SM, we have also observed from the analysis of the different topologies and attacks that the more complex (i.e., more variability at different scales) the efficiency curves are, the greater the improvement in the accuracy of robustness assessment by acknowledging idle indicators. Guided by this finding, we investigate the potential role of idle information in distilling variability in the data set to improve network robustness estimation. To this end, we systematically explore the effect of variability in the training set in estimating robustness. More specifically, we trained neural networks with training sets of increasing variability by combining different topologies, attacks, and link densities (including a data set consisting of all combinations), and then compared the estimation accuracy when only active indicators are considered, and when active and idle indicators are both included.

Figure 3 shows the model outputs for the most generalized case: data for all three topologies, four densities, and three attacks are included in the training set. The results are apparent: the inclusion of idle indicators [see Figs. 3(c) and 3(g)] produce exceedingly good predictions when compared with those achieved via only active indicators [see Figs. 3(a) and 3(b)]. When the difference between the model output and the true value of robustness (*SSR*) is computed as a function of the attack stage [see Figs. 3(d) and 3(h)], a consistent pattern is observed: active and idle indicators combined outperformed the active indicators alone during the most significant part of the attack sequence.

As expected, a general trend is also observed (see the SM): the more heterogeneous the training set is, the less accurate is the estimation of network robustness done by all three neural networks. However, the rate of performance deterioration is not comparable. As soon as variability is introduced in the training set, the neural network using the active indicators exclusively is not able to estimate even the general trend. Whereas the neural network trained using both the active and idle indicators is able to estimate the general trend very well and a majority of the variability. These results are consistent for all of the topologies tested (see Fig. 3 and the SM).

Two further remarks are noteworthy: (i) In several instances, the neural networks trained exclusively with idle indicators outperform their active counterparts in assessing network robustness, highlighting the relevant information content in the idle network. (ii) In select cases, the neural network trained with all topologies, densities, and attacks provides a more accurate estimation of robustness, than the neural network trained for a specific topology, attack, and link density, highlighting the value of idle indicators in interpreting the overall variability in the data set to improve estimations.

Acknowledging that the used synthetic networks lack some properties often exhibited by real-world networks (e.g., modularity), we further test the relevance of idle network information in assessing robustness of real networks. We simulate stochastic degree attacks on real-world topologies, where the probability of removing a given node is proportional to its original degree. We also evaluate the role of idle information in generalizing the estimation robustness for an unseen attack (e.g., based on betweenness centrality). Particularly, we first train a neural network using only active indicators resulting from 200 node removal sequences obtained by following a stochastic degree attack strategy. Our results show a fairly good estimation of our proxy of robustness [see Fig. 4(a): Little Rock Lake Food Web [36]]. That same trained neural

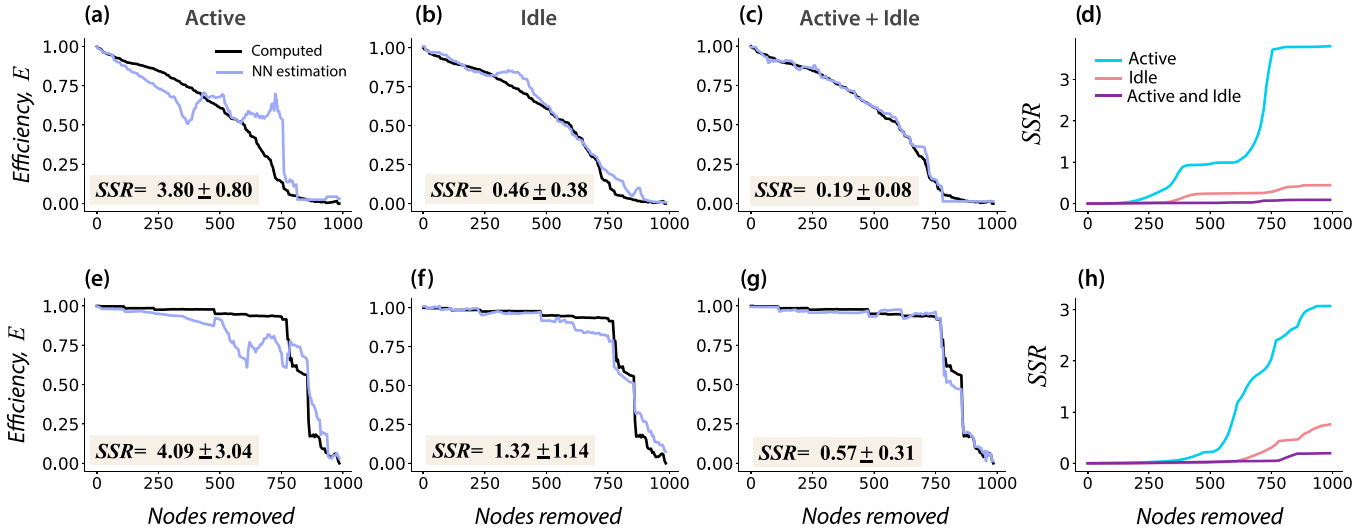

FIG. 3. Performance of a neural network in estimating network robustness when trained with the entire data set of topologies, attacks, and link densities. Results for three different models are presented: neural network trained with (a),(e) active indicators, (b),(f) idle indicators, and (c),(g) both active and idle indicators. The estimations are displayed for an Erdős-Rényi topology with $\bar{k} = 6$ undergoing degree attacks (top panels), and for scale-free networks of $\bar{k} = 12$ undergoing random attacks (bottom panels). The mean and standard deviation of the sum of the squared residuals (*SSR*) computed for a testing set, consisting of 25 synthetic topologies undergoing their respective attacks, is also reported as an inset in the respective panels. Cumulative values of *SSR* as a function of the attack stage for the (d) Erdős-Rényi and (h) scale-free topologies are displayed as well.

network fails to estimate the robustness of the same network under a stochastic betweenness attack (unseen attack) during the vast majority of the attack sequence [see Fig. 4(c)]. Notably, if a neural network is trained with the active and idle indicators of the same 200 node removal sequences (stochastic degree attacks), not only do we obtain better accuracy in estimating network robustness under stochastic degree attacks [see Fig. 4(b)], but also that neural network provides an exceptionally well-maintained accuracy in the estimation of network robustness for a previously unseen attack (stochastic betweenness attack) for the vast majority (and relevant) part of the attack sequence [see Fig. 4(d)]. These results have been tested for several real-world networks (Little Rock Lake Food Web [36], Budapest Connectome [37], and US airports [38], see the SM), corroborating our two previous findings, namely, idle network information (i) systematically improves our capacity to estimate network robustness, and (ii) allows us to retain accuracy in network robustness estimation under scenarios of enhanced variability, both in the training set and out-of-sample (e.g., altered attack strategies).

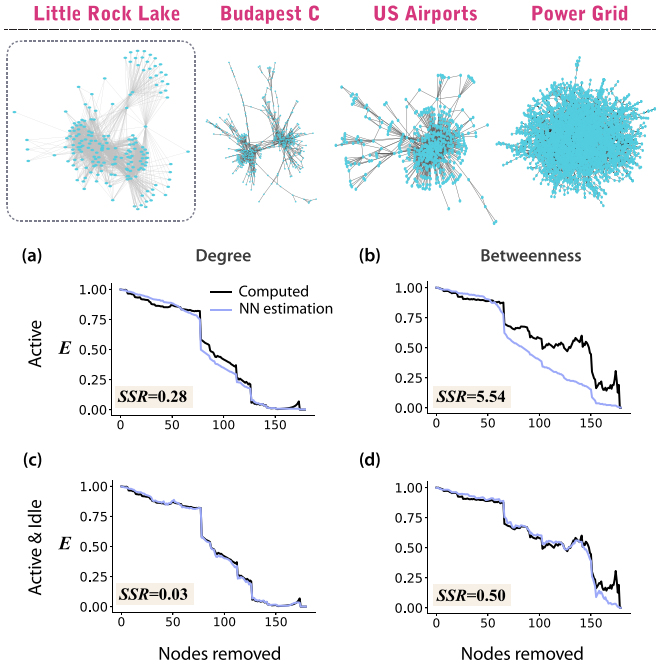

FIG. 4. Illustration of the four real network topologies tested (top panels). An artificial neural network was trained by attacking the Little Rock Lake's topology with 200 stochastic degree attack sequences. The estimation of network robustness are shown for (a),(c) attacks of the same type as those in the training, and (b),(d) previously unseen attack schemes (stochastic betweenness attack).

Our results indicate that the key role of idle indicators is to partially harness the existing information in the internal variability of the training set to gain estimation power (i) in the face of variability in the training set (either from its intrinsic stochastic variability or due to the inclusion of different topologies and attacks in the training set), and (ii) for unforeseen attacks and topological features that generate variability compatible with that observed in the training set. Thus, the idle network information is instrumental for our model (neural network) to interpret variability and improve the robustness assessment. It is worth noting that the indicators chosen in this study (size of the largest cluster and link fraction) could be particularly clumsy in encoding complementary information on network robustness to that encoded by the active indicators for certain network topologies (e.g., spatial networks such as the power grid [1]), and therefore, alternative idle indicators could be proposed to more effectively mine the information available in the idle network in those cases.

We would like to remark that this study uses a neural network as a tool to turn our hypothesis into a regression problem. The chosen neural network architecture and typology to estimate our proxy of robustness is not intended to be optimal, but to demonstrate the information content and role of the idle network in the assessment of network robustness. Thus, we anticipate that using convolutional neural networks may improve the accuracy of robustness estimation. Such further improvements in the accuracy of estimating efficiency can lead to important implications, since neural networks trained for generalized data sets would offer a light way to estimate network efficiency, which otherwise is a computationally very demanding quantity to be calculated.

Finally, we want to emphasize that we have presented this new general framework for network robustness assessment and tested our hypothesis using undirected and unweighted monoplex for clarity and better interpretability of our results. Nevertheless, our framework relies on no assumptions that prevent it from being used in more general setups. In fact, it would be interesting to explore the role of the idle network in assessing network robustness for weighted monoplex and multilayer networks.

Assessing network robustness accurately is essential to ensure the correct and sustained functionality of many natural and engineered systems. Our study shows that there is pertinent and readily available information on the robustness of a network in the so-called idle network. The inclusion of idle network information in models to assess network robustness allows us to improve the accuracy of our estimations for a specific network topology and attack and equips models with the capability to interpret in-sample and out-sample variability to preserve estimation power amid noise and unseen variability. Thus, evaluating network robustness in the light of the idle network constitutes a conceptual paradigm shift that could improve the quality and accuracy of its assessment and might lead to new strategies to guide enhanced network resilience. Along these lines, our work also opens the path to developing new ways to design and reconfigure existing networks in order to maximize or optimize their structural robustness.

A.T. acknowledges financial support from the NSF Earth Sciences Directorate Grant No. EAR-1811909. Y.M. acknowledges support by the Government of Aragón and ERDF A way of making Europe funds through Grant No. E36-20R, by Ministerio de Ciencia e Innovación, Agencia Española de Investigación (MCIN/AEI/10.13039/501100011033) through Grant No. PID2020-115800GB-I00, and by Soremartec S.A. and Soremartec Italia, Ferrero Group. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

[1] D. J. Watts and S. H. Strogatz, Collective dynamics of small-world networks, Nature (London) **393**, 440 (1998).

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

[3] M. Newman, *Networks: An Introduction* (Oxford University Press, Inc., New York, NY, USA, 2010).

[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Complex networks: Structure and dynamics, Phys. Rep. **424**, 175 (2006).

[5] A. Barrat, M. Barthlemy, and A. Vespignani, *Dynamical Processes on Complex Networks*, 1st ed. (Cambridge University Press, New York, NY, USA, 2008).

[6] I. Rodriguez-Iturbe and A. Rinaldo, *Fractal River Basins: Chance and Self-Organization, 2nd ed* (Cambridge Univ Press, New York, 2001), p. 547.

[7] E. Bullmore and O. Sporns, Complex brain networks: Graph theoretical analysis of structural and functional systems, Nat. Rev. Neurosci. **10**, 186 (2009).

[8] M. Kivela, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, Multilayer networks, Journal of Complex Networks **2**, 203 (2014).

[9] A. Tejedor, A. Longjas, D. Edmonds, T. Georgiou, I. Zaliapin, A. Rinaldo, and E. Foufoula-Georgiou, Entropy and optimality in river deltas, Proc. Natl. Acad. Sci. USA **114**, 11651 (2017).

[10] R. V. Solé, M. Rosas-Casals, B. Corominas-Murtra, and S. Valverde, Robustness of the european power grids under intentional attack, Phys. Rev. E **77**, 026102 (2008).

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

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

[13] D. R. Wuellner, S. Roy, and R. M. D'Souza, Resilience and rewiring of the passenger airline networks in the united states, Phys. Rev. E **82**, 056101 (2010).

[14] C. M. Schneider, A. A. Moreira, J. S. Andrade, S. Havlin, and H. J. Herrmann, Mitigation of malicious attacks on networks, Proc. Natl. Acad. Sci. **108**, 3838 (2011).

[15] F. Radicchi and G. Bianconi, Redundant Interdependencies Boost the Robustness of Multiplex Networks, Phys. Rev. X **7**, 011013 (2017).

[16] G. Bertagnolli, R. Gallotti, and M. De Domenico, Quantifying efficient information exchange in real network flows, Commun. Phys. **4**, 125 (2021).

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

[18] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Network Robustness and Fragility: Percolation on Random Graphs, Phys. Rev. Lett. **85**, 5468 (2000).

[19] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, Attack vulnerability of complex networks, Phys. Rev. E **65**, 056109 (2002).

[20] A. E. Motter and Y.-C. Lai, Cascade-based attacks on complex networks, Phys. Rev. E **66**, 065102 (2002).

[21] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, Catastrophic cascade of failures in interdependent networks, Nature (London) **464**, 1025 (2010).

[22] B. Min, S. D. Yi, K.-M. Lee, and K.-I. Goh, Network robustness of multiplex networks with interlayer degree correlations, Phys. Rev. E **89**, 042811 (2014).

[23] M. Bellingeri and D. Cassi, Robustness of weighted networks, Physica A **489**, 47 (2018).

[24] A. Ghavasieh, G. Bertagnolli, and M. D. Domenico, Dismantling the information flow in complex interconnected systems (2022), arXiv:2202.09692.

[25] A. Tejedor, A. Longjas, I. Zaliapin, S. Ambroj, and E. Foufoula-Georgiou, Network robustness assessed within a dual connectivity framework: Joint dynamics of the active and idle networks, Sci. Rep. **7**, 8567 (2017).

[26] S. Trajanovski, J. Martín-Hernández, W. Winterbach, and P. Van Mieghem, Robustness envelopes of networks, Journal of Complex Networks **1**, 44 (2013).

[27] M. Ventresca and D. Aleman, Network robustness versus multi-strategy sequential attack, Journal of Complex Networks **3**, 126 (2015).

[28] M. J. Williams and M. Musolesi, Spatio-temporal networks: Reachability, centrality and robustness, R. Soc. Open Sci. **3**, 160196 (2016).

[29] O. Cats and P. Krishnakumari, Metropolitan rail network robustness, Physica A **549**, 124317 (2020).

[30] P. Erdős and A. Rényi, On random graphs I, Publicationes Mathematicae Debrecen **6**, 290 (1959).

[31] M. Catanzaro, M. Boguñá, and R. Pastor-Satorras, Generation of uncorrelated random scale-free networks, Phys. Rev. E **71**, 027103 (2005).

[32] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.4.L042050 for extended results corresponding to the complete set of experiments performed in this work.

[33] I. J. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, USA, 2016).

[34] T. Chakraborty, A. K. Chakraborty, and S. Chattopadhyay, A novel distribution-free hybrid regression model for manufacturing process efficiency improvement, J. Comput. Appl. Math. **362**, 130 (2019).

[35] R. Iten, T. Metger, H. Wilming, L. del Rio, and R. Renner, Discovering Physical Concepts with Neural Networks, Phys. Rev. Lett. **124**, 010508 (2020).

[36] N. D. Martinez, Artifacts or attributes? effects of resolution on the little rock lake food web, Ecological Monographs **61**, 367 (1991).

[37] B. Szalkai, C. Kerepesi, B. Varga, and V. Grolmusz, The budapest reference connectome server v2.0, Neurosci. Lett. **595**, 60 (2015).

[38] V. Colizza, R. Pastor-Satorras, and A. Vespignani, Reaction–diffusion processes and metapopulation models in heterogeneous networks, Nat. Phys. **3**, 276 (2007).