PNAS
www.pnas.org

Supplementary Information for

What Do Network Motifs Tell Us about Resilience and Reliability of Complex Networks?

Asim K. Dey, Yulia R. Gel and H. Vincent Poor

H. Vincent Poor.
E-mail: poor@princeton.edu

This PDF file includes:

Supplementary text
Figs. S1 to S8
Tables S1 to S3
References for SI reference citations

www.pnas.org/cgi/doi/10.1073/pnas.1819529116

Supporting Information Text

A Case Study: Motif-Based Reliability of European Power Grid Networks

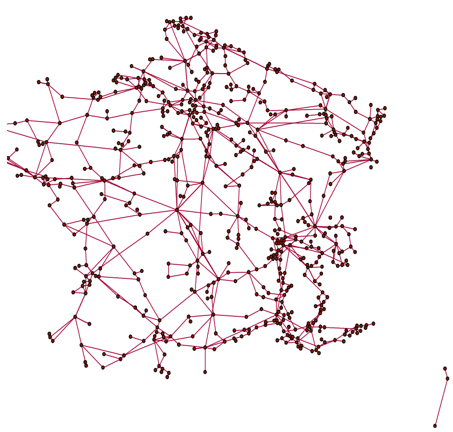
We apply our methods to electricity transmission networks of four European countries, i.e., Germany, Italy, France, and Spain, where nodes corresponds to power stations/sub-stations, and edges correspond to physical connections between nodes. We compare the results with commonly used resilience measures. The data are obtained from the Union for the Coordination of the Transmission of Electricity (UCTE). The numbers of nodes and edges of the four power system networks are listed in Table S1. The topological transmission networks for the four power grids are shown in Fig. S1.

Table S1. Network descriptions.

Power grid	# of nodes	# of edges
Germany	445	567
Italy	273	375
France	677	913
Spain	472	676

(a) German power grid

(b) Italian power grid



(c) French power grid

(d) Spanish power grid

Fig. S1. Maps representing four European country's power grid networks.

Table S2 presents conventional global network-based vulnerability metrics for the four power grids, that is, γ, average path length (APL), diameter (D), clustering coefficient (CC), betweenness centrality (BC), and critical threshold f_c based on the giant component assessment (1–3). Although lower APL and higher CC are typically considered to be associated with small world-ness and higher resilience, there exists no clear understanding whether APL or CC is the primary indicator of the world-ness and to what extent (4–6). In turn, the giant component of a network is a connected component that contains the vast majority of nodes. According to the *Molloy-Reed criterion*, a random network has a giant component if $K = \langle k^2 \rangle / \langle k \rangle > 2$, where, $\langle k \rangle$ and $\langle k^2 \rangle$ are the mean and second moment of the degree distribution $P(k)$, respectively. According to *percolation theory*, when the fraction of removed nodes reaches a critical value f_c, the network fragments into many isolated subgraphs, i.e., the network looses its giant component. When the nodes are removed randomly, the *critical threshold* can be defined as $f_c = 1 - (1/\langle k^2 \rangle / \langle k \rangle - 1)$ (7). In turn, recall that if the characteristic parameter γ of the Poisson degree distribution is less than 1.5 the network is robust, otherwise the network is fragile (8–10).

Table S2 shows that the Spanish power grid has the highest CC and lowest APL, suggesting that this power grid exhibits the small world-ness property and, as such, shall be classified as the most resilient one. In turn, CC of the French and Italian power grid networks are equal, and APL metrics are almost the same – thereby, suggesting their similar levels of resilience in terms of these metrics. The German power grid network has the highest APL and lowest CC, and as such shall be classified as fragile. Rankings of the four power grid networks in terms of critical thresholds f_c and D mirror those based on APL and CC. However, we find no monotonicity in BC – that is, the highest BC is delivered by the French power grid, followed by the German, Spanish, and Italian power grids. Moreover, ranking of resilience levels in terms of the γ parameter of (9, 11) among the four power grid networks is opposite – that is, the German and Italian power grids are classified as robust (i.e., γ of 1.32 and 1.21, respectively) and the French and Spanish grids are classified as fragile (i.e., γ of 2.16 and 2.22, respectively. These contradictory findings echo discussion by (5, 6, 12) that a deeper insight into higher level network topology is needed to better quantify power system sensitivity to various attacks and failures.

Table S2. Vulnerability metrics for the power grid networks

Power System	γ	APL	D	CC	BC	Critical Threshold, f_c
Germany	1.32	11.75	30	0.07	2235.80	0.58
Italy	1.21	9.74	28	0.08	981.85	0.61
France	2.16	9.59	26	0.08	2750.01	0.66
Spain	2.22	8.26	18	0.09	1670.02	0.70

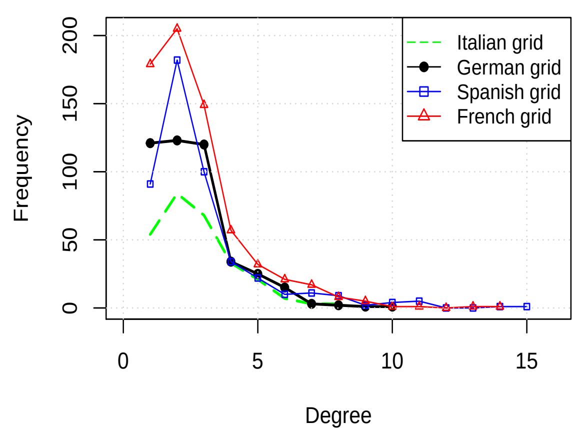
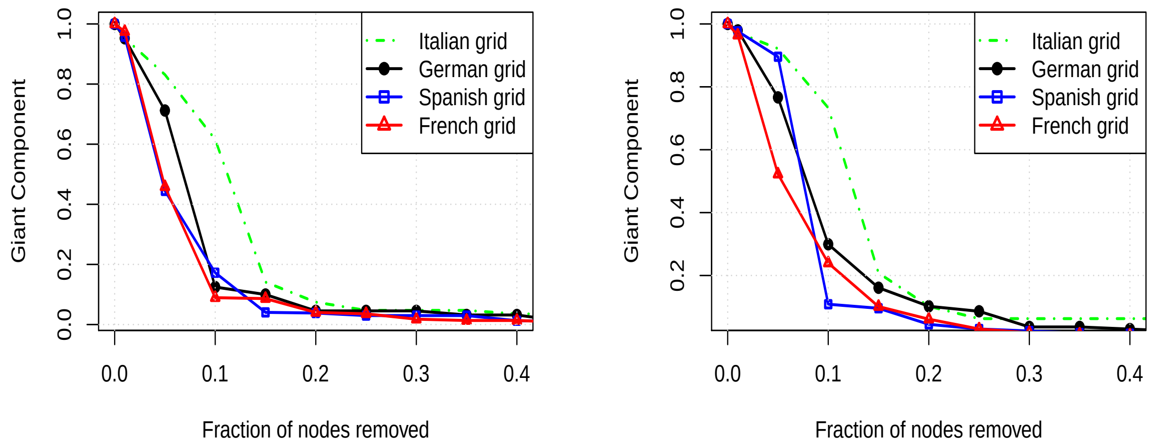


Fig. S2. Degree distributions of the power grid networks.

As hypothesized by an anonymous reviewer, the motif based results might be an expected consequence of the degree distribution. That is, graphs with larger characteristic parameters naturally might be more vulnerable due to higher occurrence of higher-degree nodes which participate in more motifs. Fig. S2 compares the degree distributions of four power grid networks. The degree distributions of the French and Spanish power grids exhibit somewhat longer tails, which implies that these networks have higher degree nodes compared to the Italian and German power grids. That is, the reviewer's hypothesis appears to be indeed plausible in this case. To assess generality of this hypothesis, we studied degree distribution and motif based vulnerability measures of different random networks, e.g., Erdos-Renyi, geometric random graph, etc. However, we could not establish a universal property for all networks. Indeed, notice that motifs quantify joint degree distributions, and the relationship between properties of marginal and joint degree distributions can be very complex. As such, we leave this as a conjecture requiring more detailed theoretical and empirical analysis.



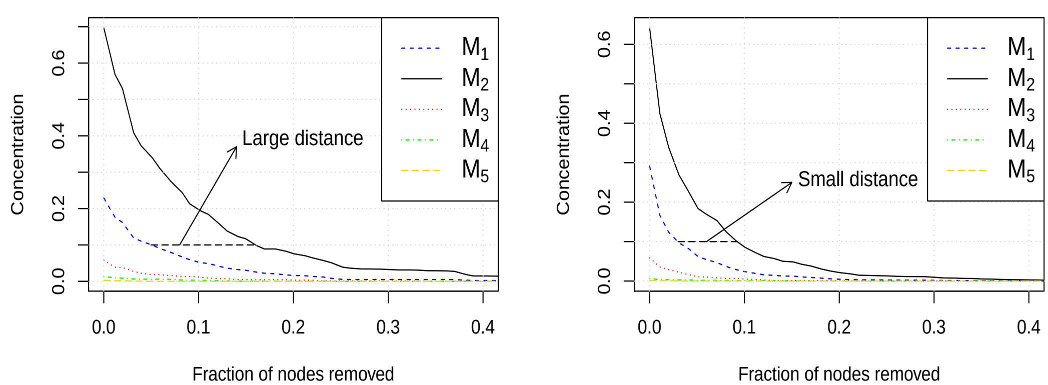
(a) Degree based attack

(b) Betweenness based attack

Fig. S3. Giant components under attacks.

Fig. S3 shows the *normalized* giant component, after a fraction of nodes have failed, where the normalized giant component is calculated as $\Delta S_p = S^p/S^0$, with S^p the giant component S after p percent nodes have failed and S^0 the initial S. Note that while the giant component of the Italian power grid decays most slowly, suggesting the highest degree of resilience, there exist no clear discrimination among resilience levels of the German, French, and Spanish power grids – that is, the giant component overall delivers inconclusive results. Indeed, while it is hard to objectively judge the "true" resilience of the power grid networks in the absence of ground truth data, it is unlikely that rankings of the performance curves exhibits such non-monotonic behavior as depicted in Fig. S3. In turn, the motif-based method (see Fig. 4 in the main body of the manuscript) clearly differentiates among performance under attacks exhibited by these power grid networks.

Fig. S4 and Fig. S5 are the counterpart of Fig. 2 in the main document, that is, it shows the decay of motif concentrations of power grids under degree based and betweenness based attacks. Fig. S4, Fig. S5 and Fig. 2 in the main document suggest that the motif decaying rate for the Italian and German power grid is much slower than for the French and Spanish power grid.



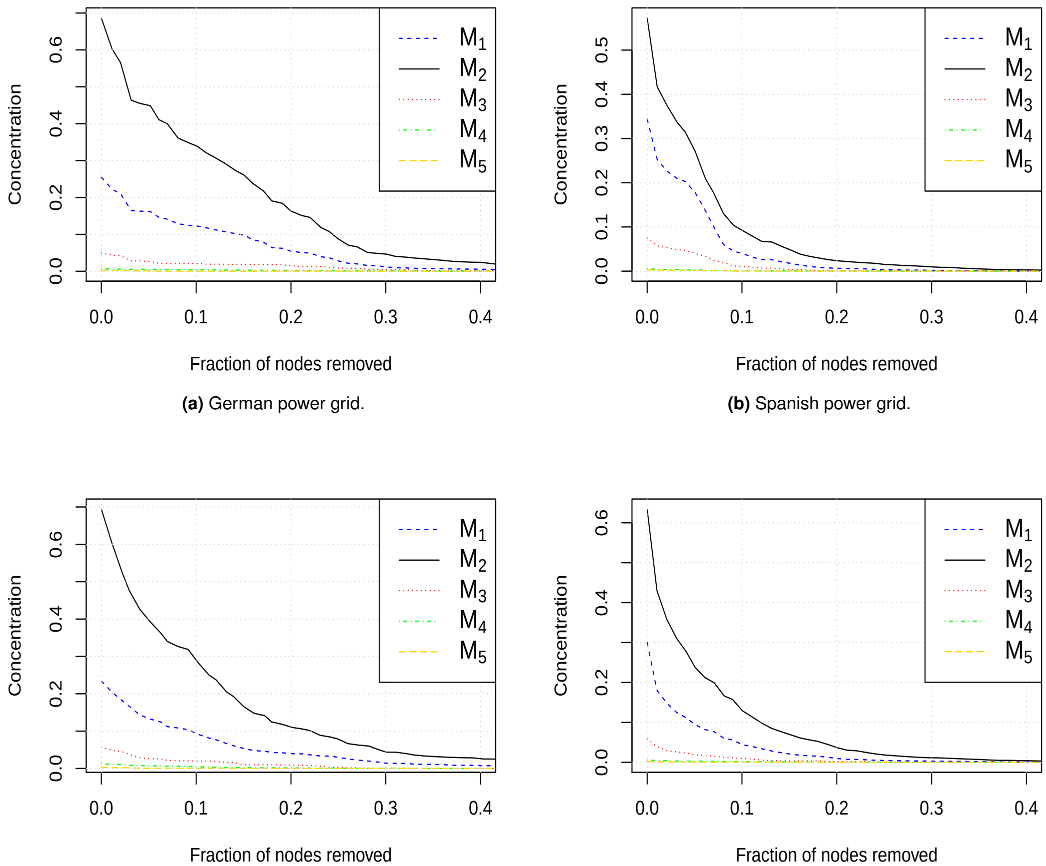
(a) Italian power grid.

(b) French power grid.

Fig. S4. Dynamics of 4-node motif concentrations under degree based attack.

We also estimate the mean lifetimes of motif distributions based on the exponential model. Table S3 provides the estimated mean lifetimes of the five motifs. The motif M_6 does not exist in the four power grid networks. We find that under both degree based and betweenness based attacks the mean motif lifetimes for the German and Italian networks are found to be greater than the mean motif lifetimes for the French and Spanish networks.

The goodness of fit of exponential models for motif lifetimes are assessed graphically. For exponential lifetimes the survival function of motif M_k can be written as $-\log(R_k(t)) = \lambda_k t$, which is an equation of a straight line that passes through the origin. If the plot of $-\log(\hat{R}_k(t))$ versus t is approximately linear and passes through the origin we can say that the exponential model



(a) German power grid.
(b) Spanish power grid.
(c) Italian power grid.
(d) French power grid.

Fig. S5. Dynamics of 4-node motif concentrations under betweenness based attack.

Table S3. Estimated mean lifetimes of motifs for four European power grid networks.

	Degree based attacks				Betweenness based attacks			
Motif	Italian	German	Spanish	French	Italian	German	Spanish	French
M_1	7.27	5.94	3.29	3.98	10.48	12.20	5.84	5.21
M_2	8.72	7.74	4.69	5.22	10.15	12.93	6.42	6.43
M_3	6.02	5.47	3.35	3.93	9.02	12.88	6.27	5.45
M_4	6.14	7.35	3.20	5.73	9.28	16.64	6.22	8.37
M_5	2.83	3.22	3.20	2.86	5.17	6.0	6.87	5.00

is appropriate for the motif lifetimes. Fig. S6 shows the goodness of fit plots for the German network motifs under degree based attacks. All the plots for German network motifs are approximately 45°- lines, which implies the exponential models fit the motif lifetime data well. Goodness-of-fit testing of other power grid networks (under both types of attack strategies) yielded analogous results and are omitted for brevity.

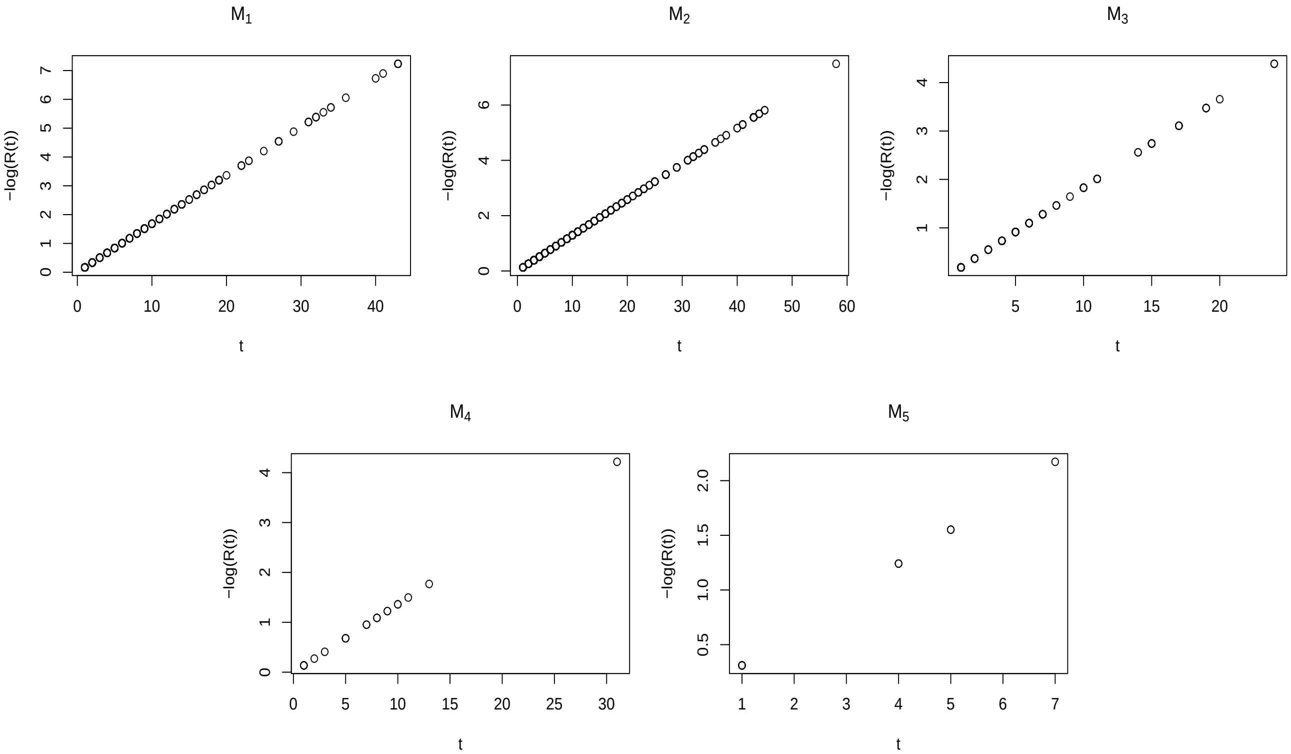
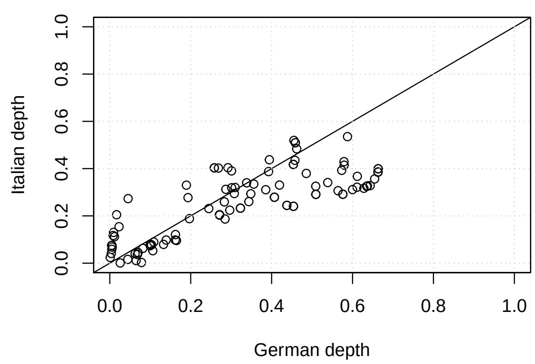
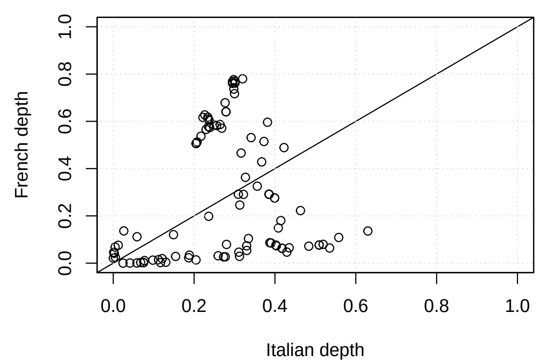


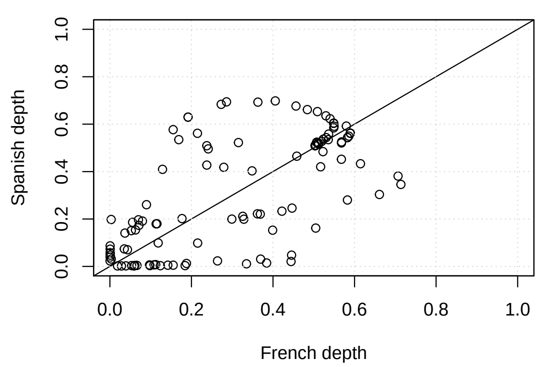
Fig. S6. Goodness of fit plots for the German network motifs, under degree based attacks.



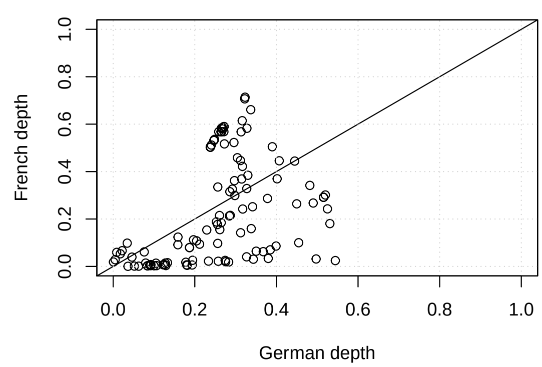
(a) DD plots, degree based attack



(b) DD plots, degree based attack



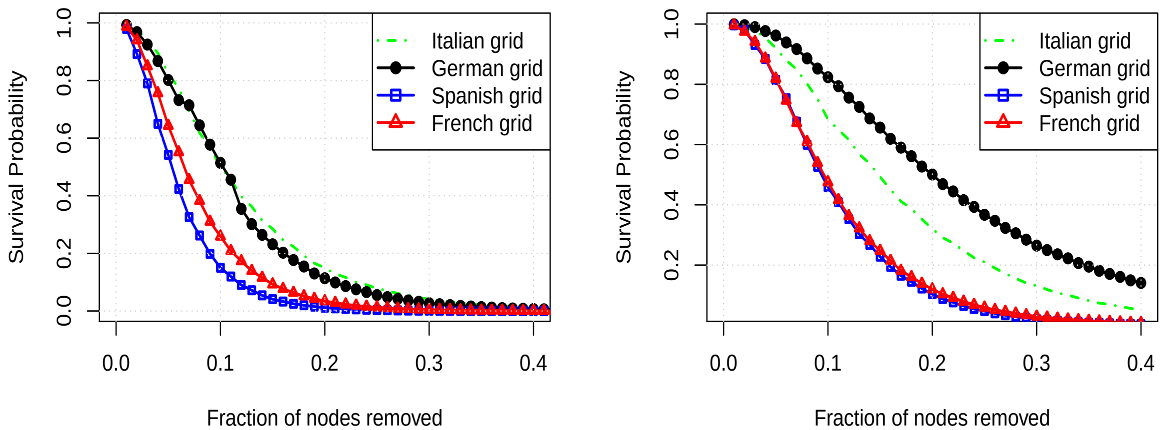
(c) DD plots, betweenness based attack



(d) DD plots, betweenness based attack

Fig. S7. DD plots for four power grid concentration distributions

We now study the concentration of the five motifs with a five-dimensional multivariate distribution. We compute multivariate concentration distributions of different power grid networks based on data depth techniques e.g., the DD plot. Fig. S7 shows the corresponding pairwise DD plots. As Fig. S7 indicates, there exists a noticeable difference between concentration of the German and Spanish power grids, as well as in the French and Italian concentrations. However, the concentrations of the German and Italian power grids look similar, as do the French and Spanish power grids.



(a) Reliability under degree based attack.

(b) Reliability under betweenness based attack

Fig. S8. Reliability curves of four power grid networks.

Finally, we also evaluate network reliability by approximating the network by a parallel system, where each 4-node connected motifs, i.e., M_1, M_2, M_3, M_4, M_5 and M_6, is viewed as a component and at least one motif must succeed for the system to survive. The reliability function of the motif can be written as $R_k(t) = P_r(A_k) = P_r(T_k > t)$, where T_k is the lifetime of the motif M_k, $k = 1, 2, \cdots, 6$ (13). The reliability function of the parallel network can be defined as

$$\begin{aligned} R_s(t) &= P_r\left(T_s > t\right) \\ &= 1 - P_r\left(\text{all the motif fail by time t}\right) \\ &= 1 - P_r\left(\bigcap_{k=1}^{6} A_k^c\right), \end{aligned} \qquad [1]$$

where T_s is the lifetime of the entire network. If the lifetimes of the six types of motifs T_k are mutually independent, $P_r\left(\bigcap_{k=1}^{6} A_k^c\right)$ is the product of individual $P_r(A_k^c)$, $k = 1, 2, \cdots, 6$ (14, 15). However, in practice motif lifetimes are not mutually independent. According to (16) $P_r\left(\bigcap_{k=1}^{6} A_k^c\right)$ can be approximated by the geometric mean of its upper and lower bounds as follows:

$$\begin{aligned} R_s(t) &= 1 - P_r\left(\bigcap_{k=1}^{6} A_k^c\right) \\ &= 1 - \left[\left(\prod_{k=1}^{6} P_r(A_k^c)\right) \min\left\{P_r(A_1^c), \cdots, P_r(A_6^c)\right\}\right]^{\frac{1}{2}}, \end{aligned} \qquad [2]$$

where $P_r(A_k^c) = 1 - R_k(t)$, $k = 1, 2, \cdots, 6$. Fig. S8 compares the four power grid reliabilities under degree based and betweenness based targeted attacks. The Italian and German power grids appear to be more resilient than the French and Spanish power grids, under both degree based and betweenness based targeted attacks – thereby, mirroring the conclusions delivered by the data depth analysis of network performance under hazardous scenarios.

Note that these findings echo the classification results of power grid vulnerability by (11). That is, according to (9, 11), a power grid network is robust if the characteristic parameter γ, estimated from the fit of an exponential model to a power grid cumulative degree distribution, is less than 1.5; and vulnerable, otherwise. Both the German and Italian power grid networks deliver γ of 1.32 and 1.20, respectively; while the Spanish and French power grid networks yield γ of 2.22 and 2.16, respectively. However, in contrast to the classification of (11), our new approach allows us to provide deeper insight into network vulnerability, both at a level of network local topological structure and as a function of the fraction of failed nodes.

References

1. Cuadra L, Salcedo-Sanz S, Del Ser J, Jiménez-Fernández S, Geem ZW (2015) A critical review of robustness in power grids using complex networks concepts. *Energies* 8(9):9211–9265.
2. Cohen R, Erez K, ben Avraham D, Havlin H (2000) Resilience of the internet to random breakdowns. *Physical Review Letters* 85 21:4626–8.
3. Piraveenan M, Uddin S, Chung KSK (2012) Measuring topological robustness of networks under sustained targeted attacks in *Proc. 2012 IEEE/ACM ASONAM Conference*. pp. 38–45.
4. Francisco Gutierrez, E. Barocio FU, Zuniga P (2013) Vulnerability analysis of power grids using modified centrality measures. *Discrete Dynamics in Nature and Society* vol. 2013:1–11.
5. Kim CJ, Obah OB (2007) Vulnerability assessment of power grid using graph topological indices. *International Journal of Emerging Electric Power Systems* 8(6):1–15.
6. Pagani GA, Aiello M (2013) The power grid as a complex network: A survey. *Physica A* 392(11):2688–2700.
7. Albert R, Barabási AL (2002) Statistical mechanics of complex networks. *Rev. Mod. Phys.* 74(1):47–97.
8. Albert R, Jeong H, Barabási AL (2000) Error and attack tolerance of complex networks. *Nature* 406:378–382.
9. Rosas-Casals M, Corominas-Murtra B (2009) Assessing European power grid reliability by means of topological measures. *WIT Transactions on Ecology and the Environment* 121:527–537.
10. Newman M (2010) *Networks: An Introduction.* (Oxford University Press, Inc., New York, NY, USA).
11. Sóle RV, Rosas-Casals M, Corominas-Murtra B, Valverde S (2008) Robustness of the European power grids under intentional attack. *Physical Review E* 77:026102.
12. Abedi A, Gaudard L, Romerio F (2019) Review of major approaches to analyze vulnerability in power system. *Reliability Engineering & System Safety* 183:153 – 172.
13. Lawless JF (2003) *Statistical Models and Methods for Lifetime Data.* (John Wiley & Sons), 2nd edition.
14. Rausand M, Høyland A (2004) *System Reliability Theory: Models, Statistical Methods, and Applications.* (John Wiley & Sons, Ltd), 2nd edition.
15. LaRocca S, Johansson J, Hassel H, Guikema S (2014) Topological performance measures as surrogates for physical flow models for risk and vulnerability analysis for electric power systems. *Risk Analysis* 35(4):608–623.
16. Rasmussen N (1975) Reactor safety study: An assessment of accident risks in U.S. commercial nuclear power plants, wash-1400. *United States Nuclear Regulatory Commission.*