Supplementary Material: Embedding-aided network dismantling

Saeed Osat,[1] Fragkiskos Papadopoulos,[2] Andreia Sofia Teixeira,[3] and Filippo Radicchi[4]

[1] *Max Planck Institute for Dynamics and Self-Organization (MPIDS), 37077 Göttingen, Germany*
[2] *Department of Electrical Engineering, Computer Engineering and Informatics, Cyprus University of Technology, 33 Saripolou Street, 3036 Limassol, Cyprus*
[3] *LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal*
[4] *Center for Complex Networks and Systems Research, Luddy School of Informatics, Computing, and Engineering, Indiana University, Bloomington, Indiana 47408, USA*

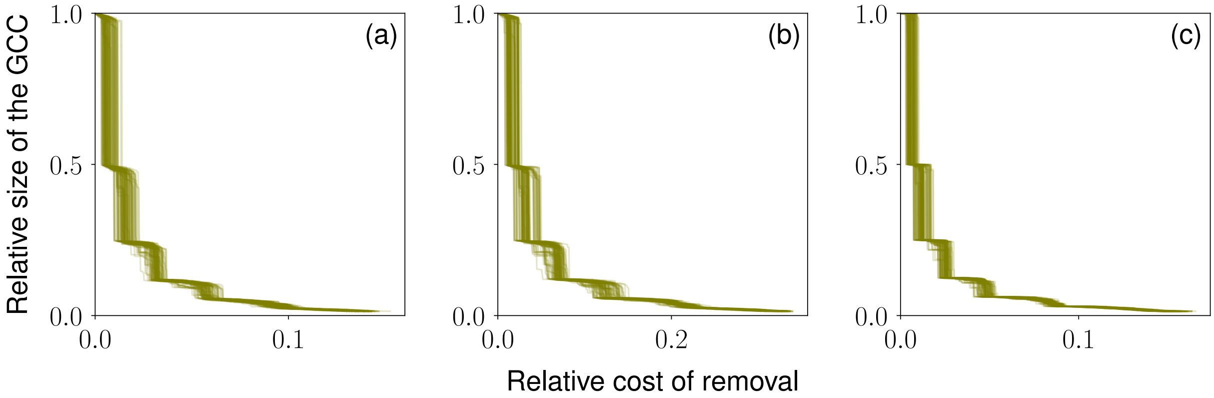

FIG. S1. **Optimal percolation on the US power grid via hyperbolic-embedding-aided dismantling.** (a) Relative size of the giant connected component (GCC) as a function of the fraction of removed nodes. Different curves correspond to solutions obtained by randomizing the initial division of the hyperbolic disk in two slices. (b) Same as in panel (a), but for the optimal site-percolation problem with degree cost. (c) Same as in panel (a), but for optimal bond percolation with unit cost.

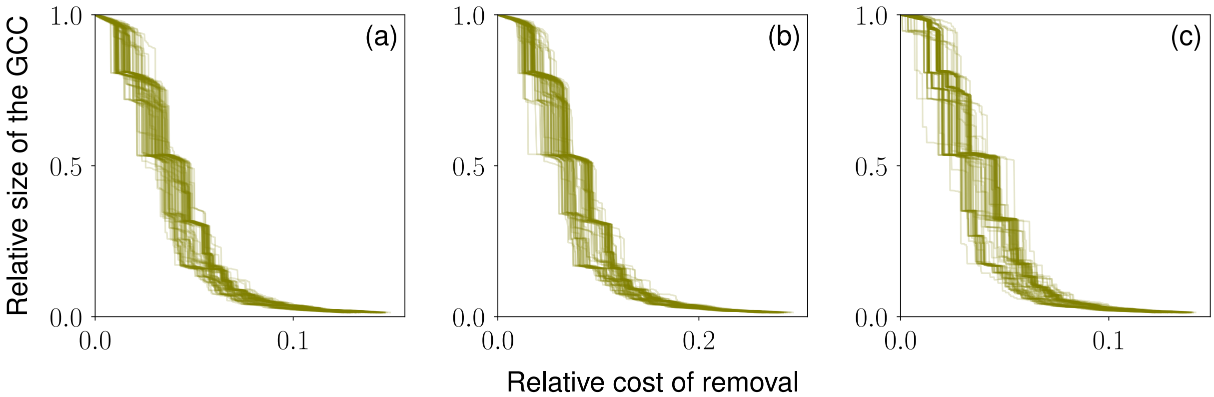

FIG. S2. **Optimal percolation on the US power grid via Node2vec-embedding-aided dismantling.** (a) Relative size of the giant connected component (GCC) as a function of the fraction of removed nodes. Different curves correspond to solutions obtained for different realizations of the dismantling algorithm. Fluctuations are due the intrinsic randomness of both the Node2vec and the k-means algorithm. (b) Same as in panel (a), but for the optimal site-percolation problem with degree cost. (c) Same as in panel (a), but for optimal bond percolation with unit cost.

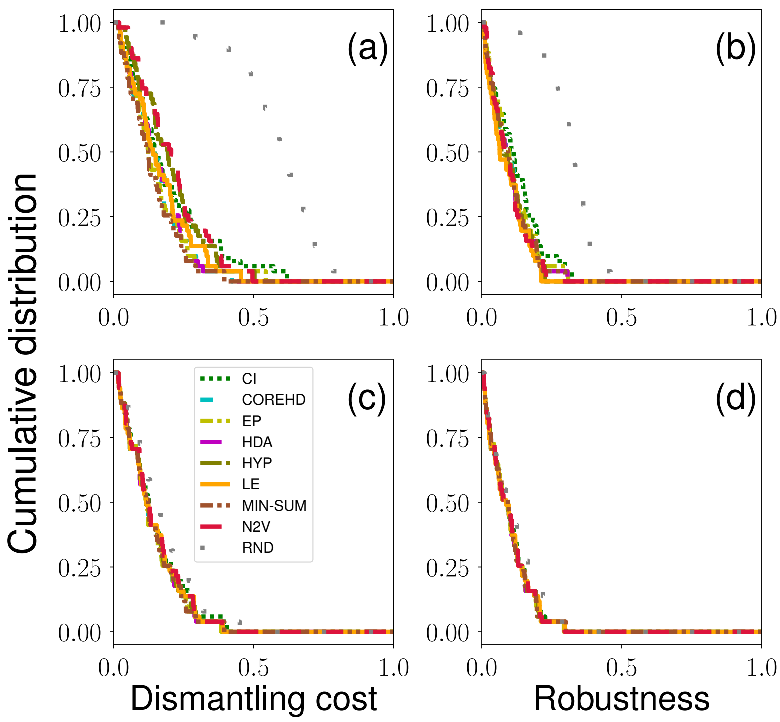

FIG. S3. **Optimal site percolation with unit cost in real-world networks.** (a) Cumulative distribution function of the dismantling cost of the various algorithms in the solution of the optimal site percolation with unit cost. The distribution is evaluated on a corpus of 50 real-world networks. (b) Same as in (a), but for the robustness metric. (c and d) Same as in (a) and (b), respectively, but for solutions obtained after the application of the greedy post-processing technique.

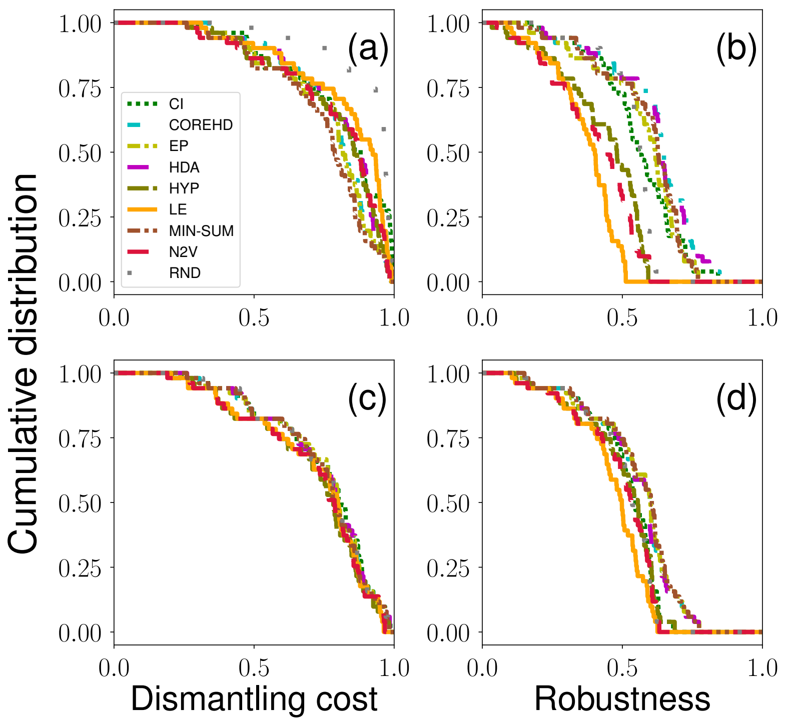

FIG. S4. **Optimal site percolation with degree cost in real-world networks.** (a) Cumulative distribution function of the dismantling cost of the various algorithms in the solution of the optimal site percolation with degree cost. The distribution is evaluated on a corpus of 50 real-world networks. (b) Same as in (a), but for the robustness metric. (c and d) Same as in (a) and (b), respectively, but for solutions obtained after the application of the greedy post-processing technique.

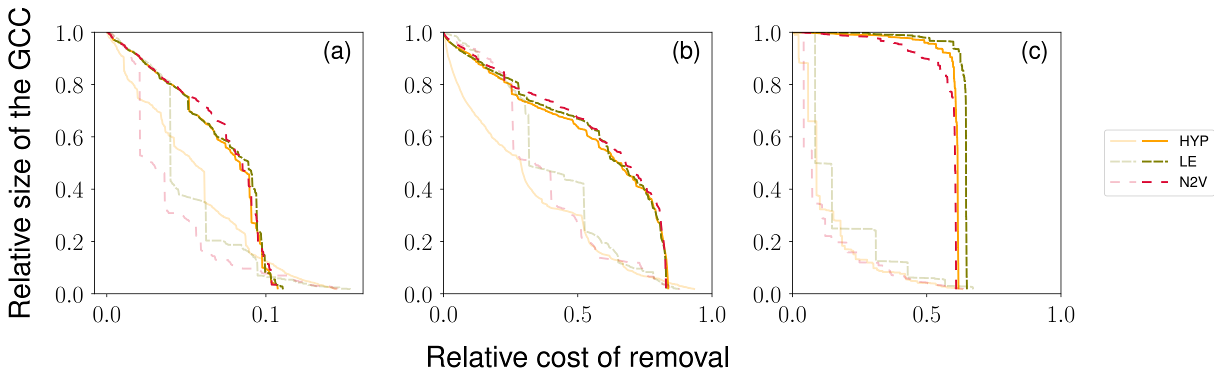

FIG. S5. **Optimal percolation in the air transportation network.** (a) Relative size of the giant connected component (GCC) as a function of dismantling cost for optimal site percolation with unit cost. The network analyzed is the air transportation network [1]. Different curves correspond to solutions obtained for different realizations of the dismantling algorithm. We display as transparent curves the solutions obtained with the original algorithm, and as non-transparent curves those resulting from the application of the post-processing technique aimed at minimizing the dismantling cost of the structural sets identified by the algorithms. (b) Same as in panel (a), but for the optimal site-percolation problem with degree cost. (c) Same as in panel (a), but for optimal bond percolation with unit cost.

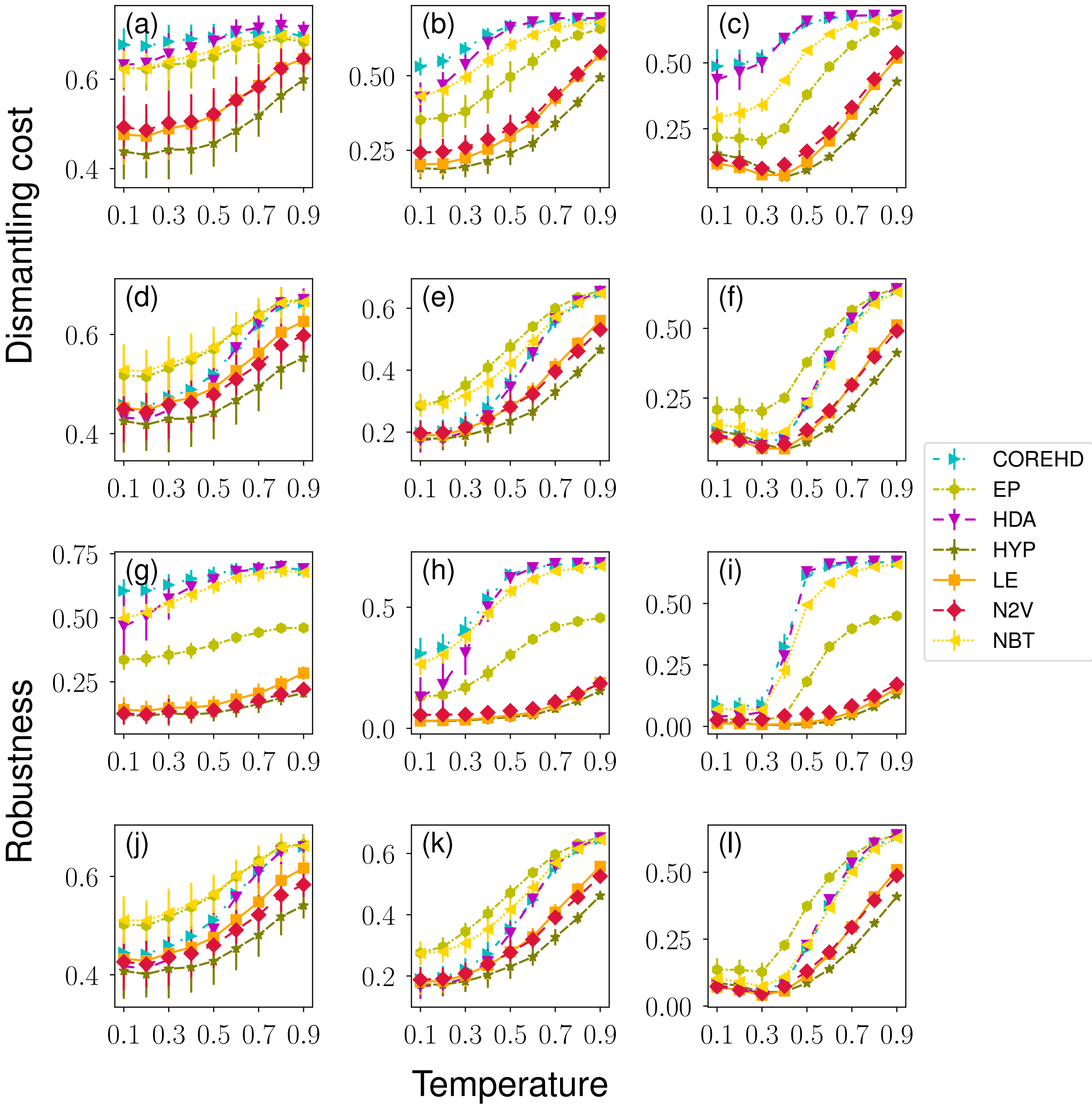

FIG. S6. **Optimal bond percolation in the $\mathbb{H}^2$ model.** (a) We apply the various dismantling methods to instances of the $\mathbb{H}^2$ model. The networks are the same as in Fig. 5 of the main paper; their degree distribution is a power law with exponent $\gamma = 2.2$. We construct networks for different values of the temperature parameter of the model, and measure the dismantling cost of the various algorithms as a function of the temperature parameter of the model. Each point is an average over 100 different instances of the $\mathbb{H}^2$ model. (d) We report the value of the dismantling cost after the application the greedy post-processing technique for the same set of experiments as in panel (a). (g) We report the value of the robustness metric for the same set of experiments as in panel (a). (j) We report the value of the robustness metric for the same set of experiments as in panel (d). (b, e, h and k) Same as in (a, d, g and j), respectively, but for $\gamma = 2.6$. (c, f, i and l) Same as in (a, d, g and j), respectively, but for $\gamma = 3.5$.

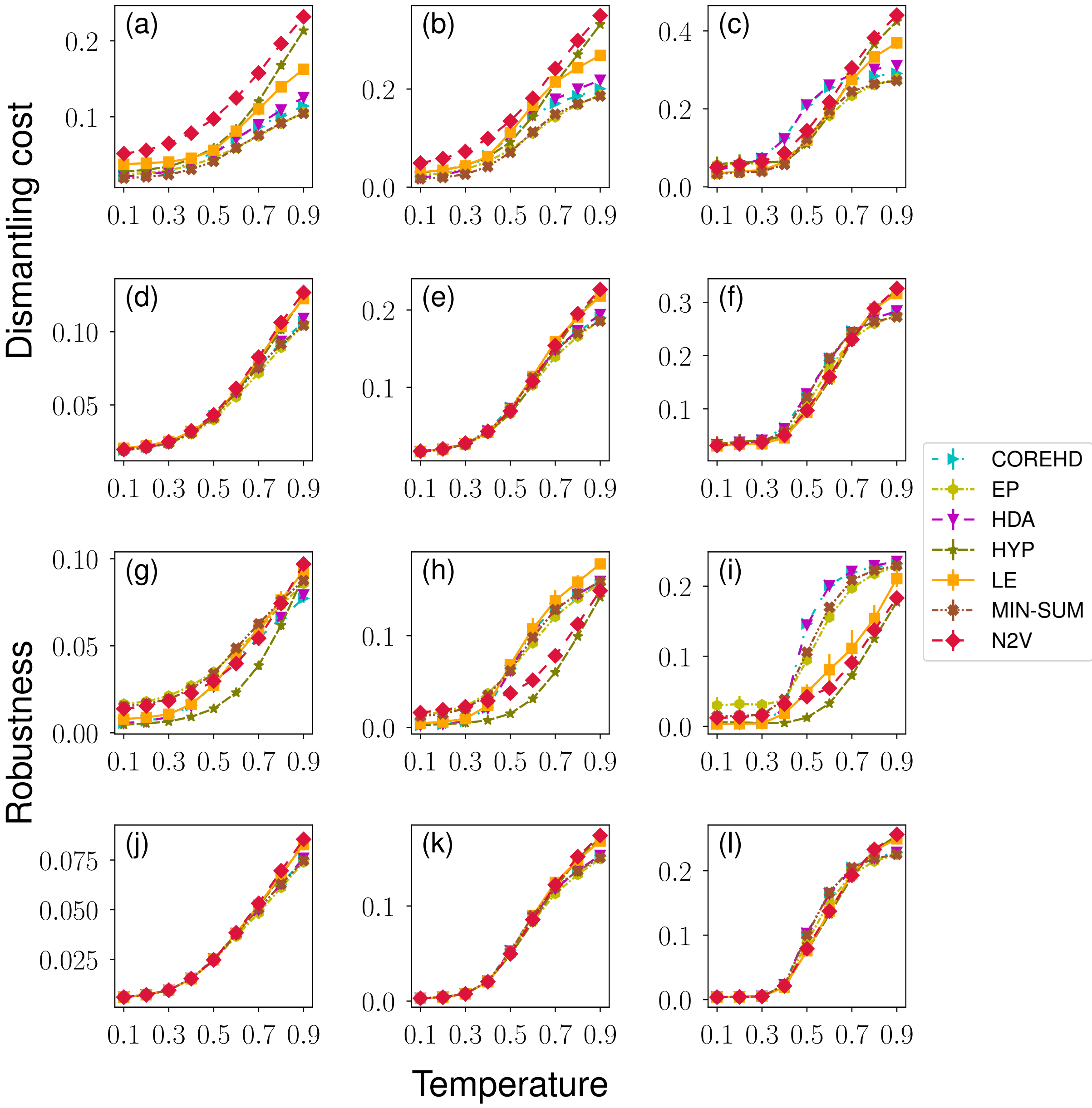

FIG. S7. **Optimal site percolation in the $\mathbb{H}^2$ model.** Same as in Fig. S6, but for the optimal site-percolation problem with unit cost.

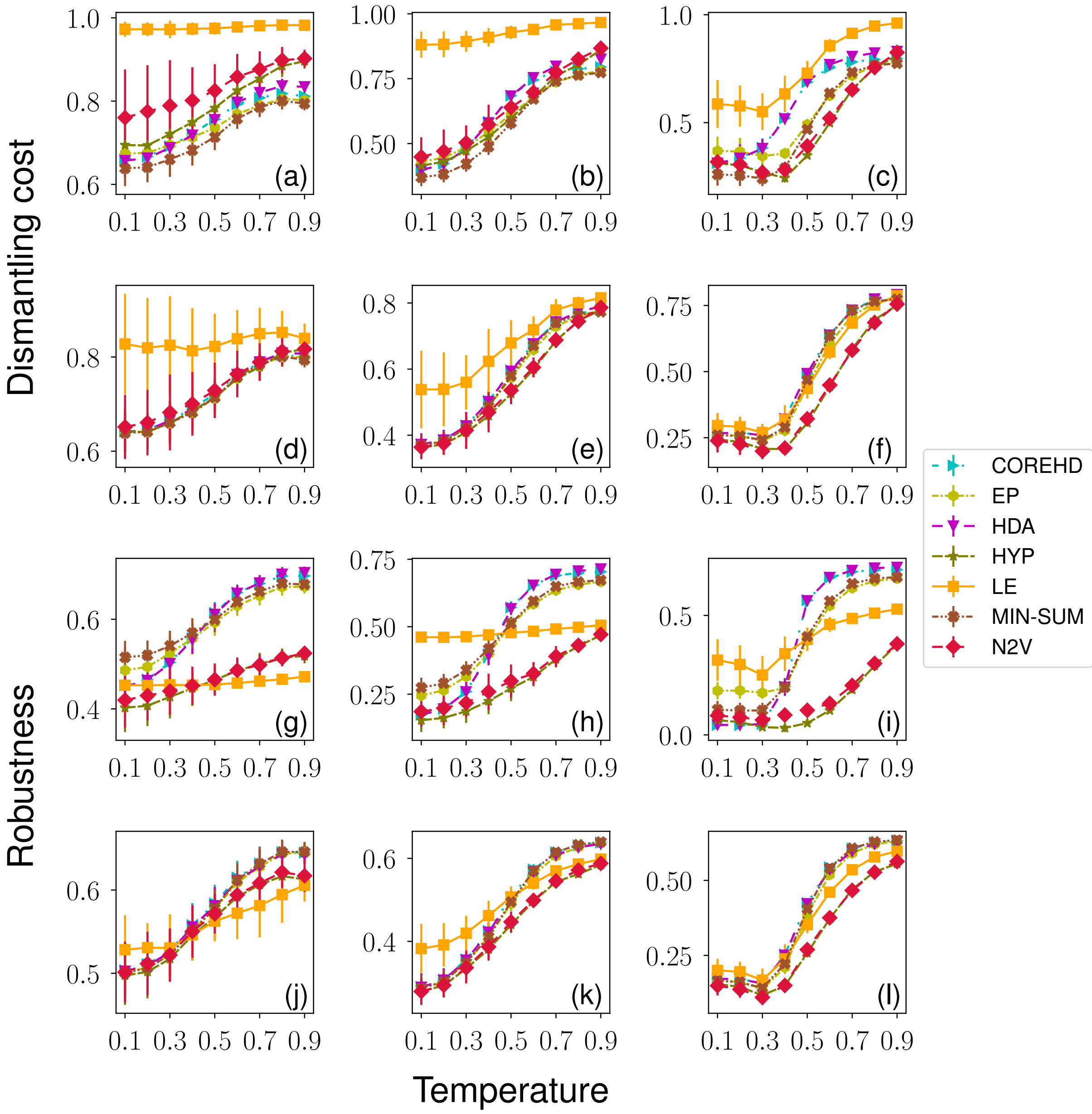

FIG. S8. **Optimal site percolation in the $\mathbb{H}^2$ model with degree cost of removal.** Same as in Fig. S6, but for the optimal site-percolation problem with degree cost.

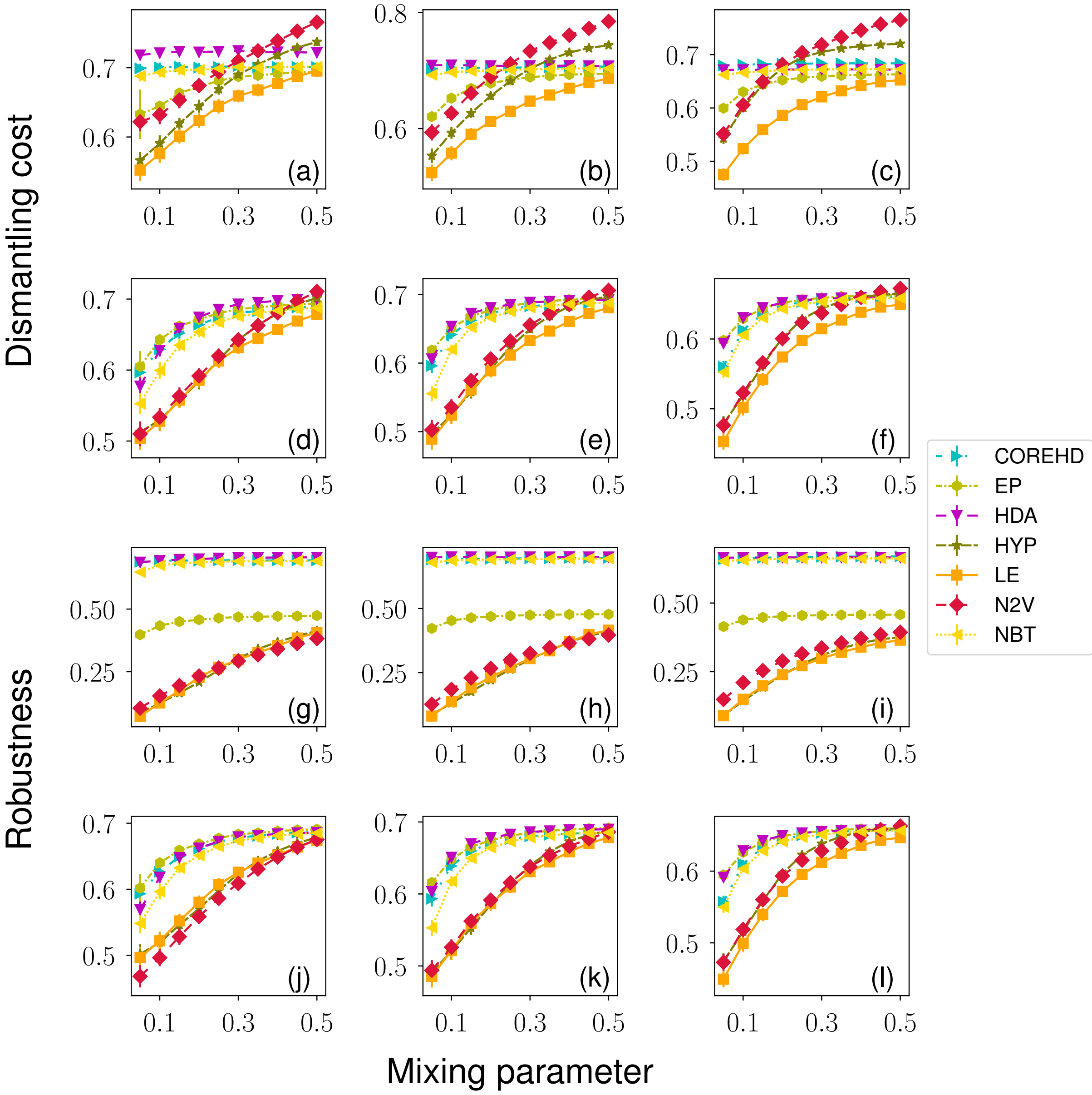

FIG. S9. **Optimal bond percolation in the LFR model.** (a) We apply the various dismantling methods to instances of the Lancichinetti-Fortunato-Radicchi (LFR) model. The degree distribution is a power law with exponent $\gamma = 2.2$. We measure the dismantling cost as a function of the mixing parameter of the LFR model. (d) We report the value of the dismantling cost after the application the greedy post-processing technique for the same set of experiments as in panel (a). (g) We report the value of the robustness metric for the same set of experiments as in panel (a). (j) We report the value of the robustness metric for the same set of experiments as in panel (d). (b, e, h and k) Same as in (a, d, g and j), respectively, but for $\gamma = 2.6$. (c, f, i and l) Same as in (a, d, g and j), respectively, but for $\gamma = 3.5$.

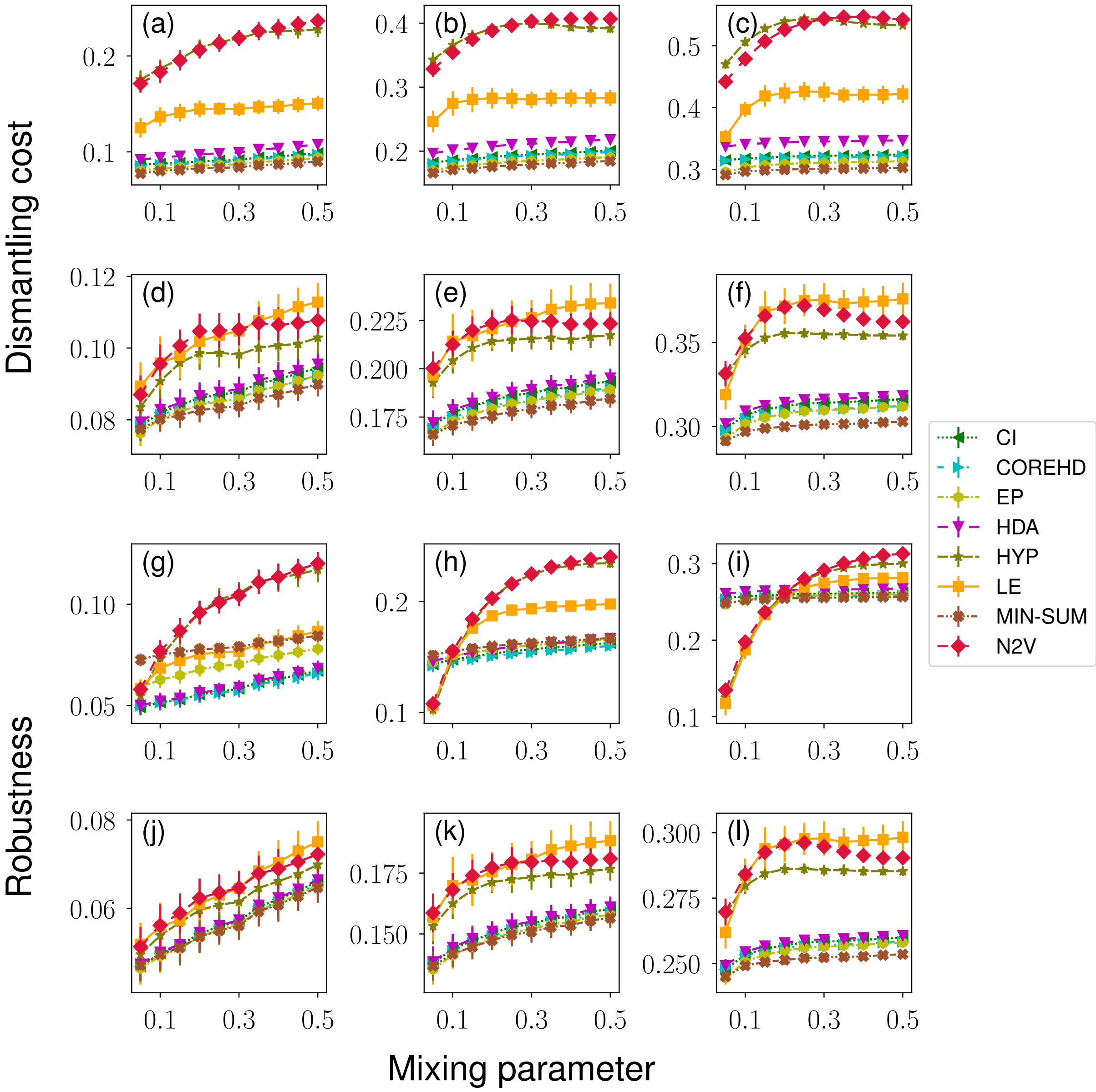

FIG. S10. **Optimal site percolation in the LFR model.** Same is Fig. S9, but for the optimal site-percolation problem with unit cost.

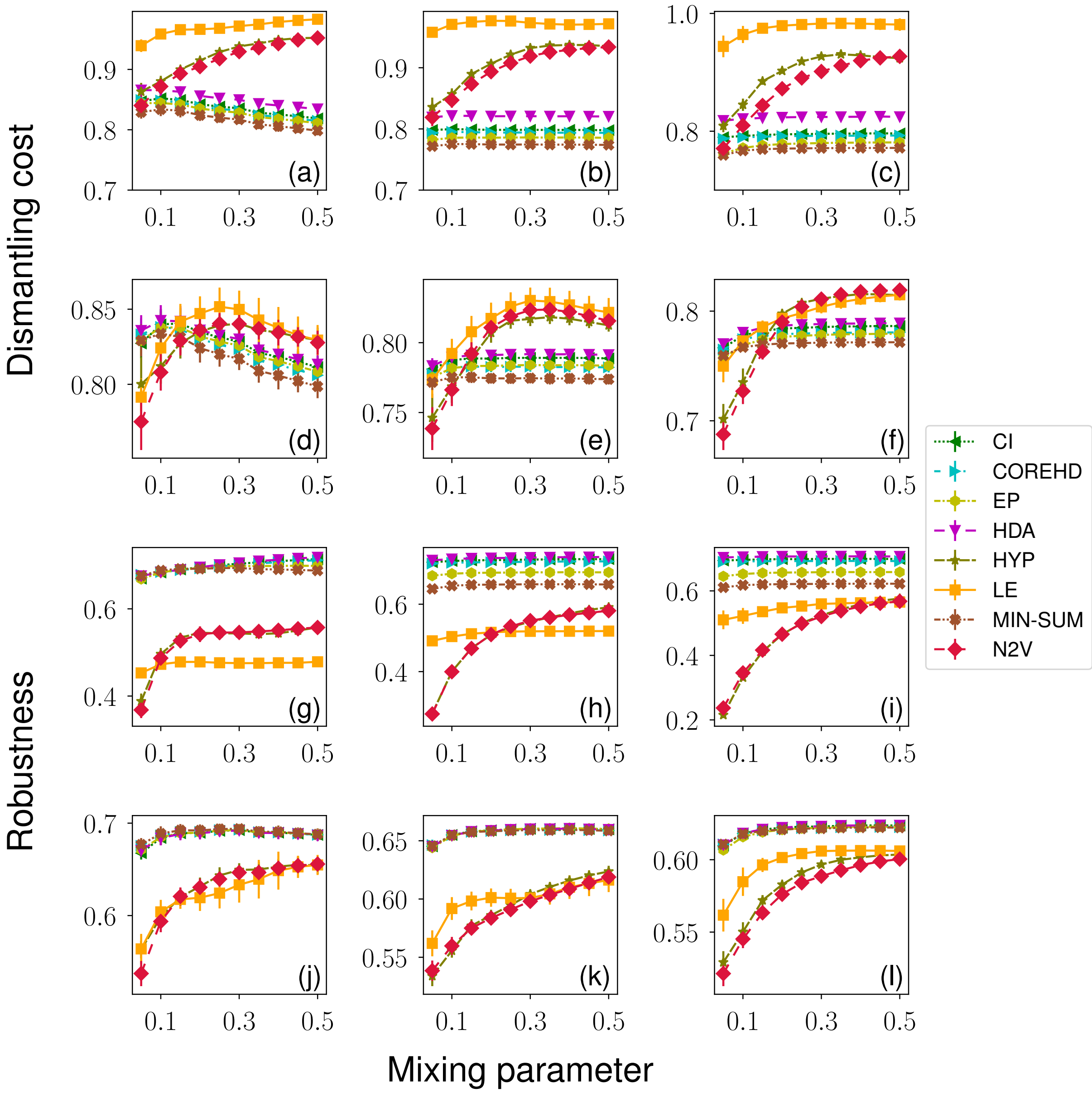

FIG. S11. **Optimal site percolation with degree cost in the LFR model.** Same as in Fig. S9, but for the optimal site-percolation problem with degree cost.

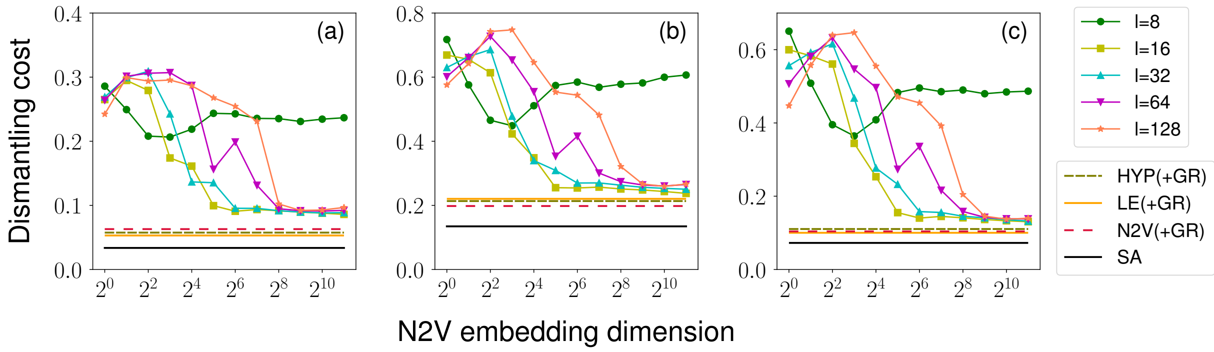

FIG. S12. **Optimal percolation on the US power grid via Node2vec-embedding-aided dismantling.** (a) Dismantling cost for the optimal bond-percolation problem as a function of the embedding dimension. Different curves correspond to different choices of the parameter l corresponding to the length of the random walks used by the Node2vec algorithm. Each point corresponds to the average value of the dismantling cost over 2^6 Node2vec embeddings. As references, we report also solutions obtained by other dismantling methods. Solutions of these algorithms have post-processed with the greedy optimization algorithm of Ref. [2]. (b) Same as in (a), but for site percolation with unit cost. (c) Same as in (a), but for site percolation with degree cost.

Name	N	E	Ref.	URL	Fig.
Corruption	309	3281	[3]	URL	S13
US Air Transportation	500	2980	[4]	-	S14
S 838	512	819	[5]	URL	S15
Yeast, transcription	662	1062	[6]	URL	S16
Crime	754	2127	[1]	URL	S17
Road EU	1039	1305	[7]	URL	S18
URV email	1133	5451	[8]	URL	S19
Political Blogs	1222	16714	[9]	URL	S20
Metabolic	1436	4718	[10]	-	S21
Yeast, protein	1458	1948	[11]	URL	S22
Drosophila	1770	8905	[12]	-	S23
Petster-Hamster	2000	16098	[1]	URL	S24
Yeast, protein	2224	6609	[13]	URL	S25
Music	2476	20624	[14]	URL	S26
Japanese	2698	7995	[5]	URL	S27
Airports	3397	19230	[1]	URL	S28
Proteome	4100	13358	[15]	URL	Fig. 3 in main text
GR-QC, 1993-2003	4158	13422	[16]	URL	S29
Tennis	4338	81865	[17]	URL	S30
Power Grid	4941	6594	[1, 18]	URL	Fig. 2 in main text
HT09	5352	18481	[19]	URL	S31
Hep-Th, 1995-1999	5835	13815	[20]	URL	S32
Gnutella, Aug. 8, 2002	6299	20776	[16, 21]	URL	S33
Autonomous Systems (AS)	6474	12572	[22]	URL	S34
English	7377	44205	[5]	URL	S35
Gnutella, Aug. 9, 2002	8104	26008	[16, 21]	URL	S36
French	8308	23832	[5]	URL	S37
Hep-Th, 1993-2003	8638	24806	[16]	URL	S38
Gnutella, Aug. 6, 2002	8717	31525	[16, 21]	URL	S39
Gnutella, Aug. 5, 2002	8842	31837	[16, 21]	URL	S40
IPv6	10259	27027	[23]	URL	S41
PGP	10680	24316	[24]	URL	S42
Gnutella, August 4 2002	10876	39994	[16, 21]	URL	S43
Hep-Ph, 1993-2003	11204	117619	[16]	URL	S44
Spanish	11558	43050	[5]	URL	S45
Cond-Mat, 1995-1999	13861	44619	[20]	URL	S46
Astrophysics	14845	119652	[20]	URL	S47
Google	15763	148585	[25]	URL	S48
ca-AstroPh	17903	196972	[16]	URL	S49
Authors	21363	91286	[16]	URL	S50
Gnutella, Aug. 25, 2002	22663	54693	[16, 21]	URL	S51
Internet	22963	48436	[1]	URL	S52
Internet	23748	58414	[23]	URL	S53
caida	26475	53381	[1, 16]	URL	S54
Gnutella, Aug. 24, 2002	26498	65359	[16, 21]	URL	S55
Cond-Mat, 1995-2003	27519	116181	[20]	URL	S56
Libimseti	34401	420784	[1, 26, 27]	URL	S57
Cond-Mat, 1995-2005	36458	171735	[20]	URL	S58
Gnutella, Aug. 30, 2002	36646	88303	[16, 21]	URL	S59
Gnutella, Aug. 31, 2002	62561	147878	[16, 21]	URL	S60

TABLE S1. **Real-world networks.** From left to right, we report: the name of the network data set, the number of nodes and edges in the largest connected component of the network, the references where the data set was first considered, the URL of the webpage where the data set can be retrieved (just click on the word URL to be redirected to the web page), and the reference to the figure where the data set is analyzed. Please consider that we analyzed only the largest connected component of each network data set.

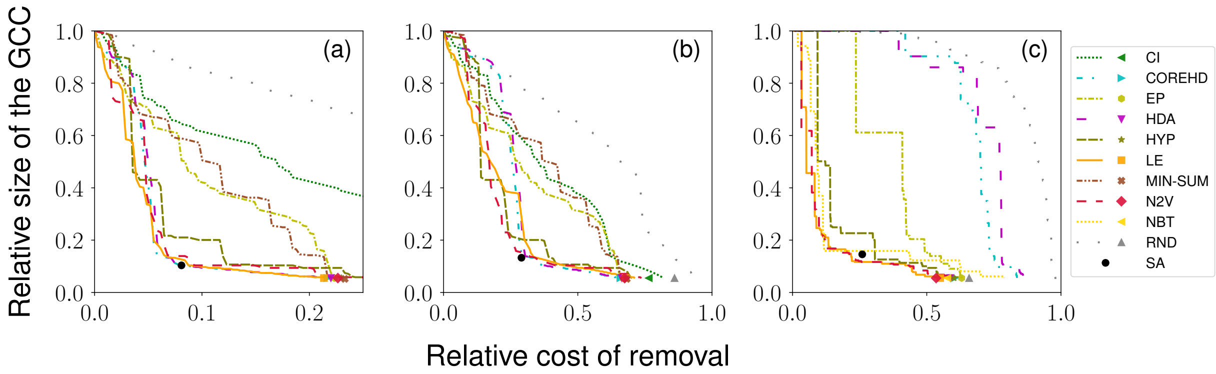

FIG. S13. (a) Relative size of the giant connected component (GCC) as a function of dismantling cost for optimal site percolation with unit cost. Different curves correspond to solutions obtained via the various dismantling methods. The network considered here is obtained from the `Corruption` data set. (b) Same as in panel (a), but for the optimal site-percolation problem with degree cost. (c) Same as in panel (a), but for optimal bond percolation with unit cost.

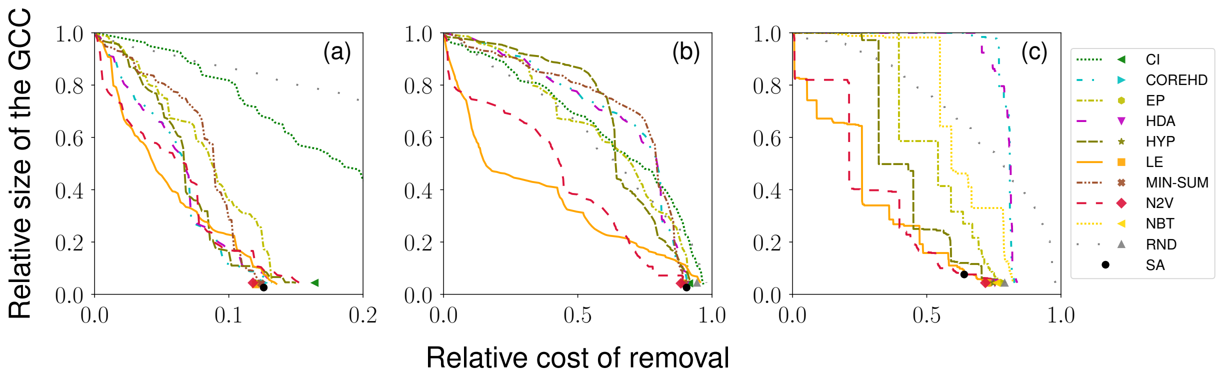

FIG. S14. Same as in Fig. S13, but for the network `US Air Trasportation`.

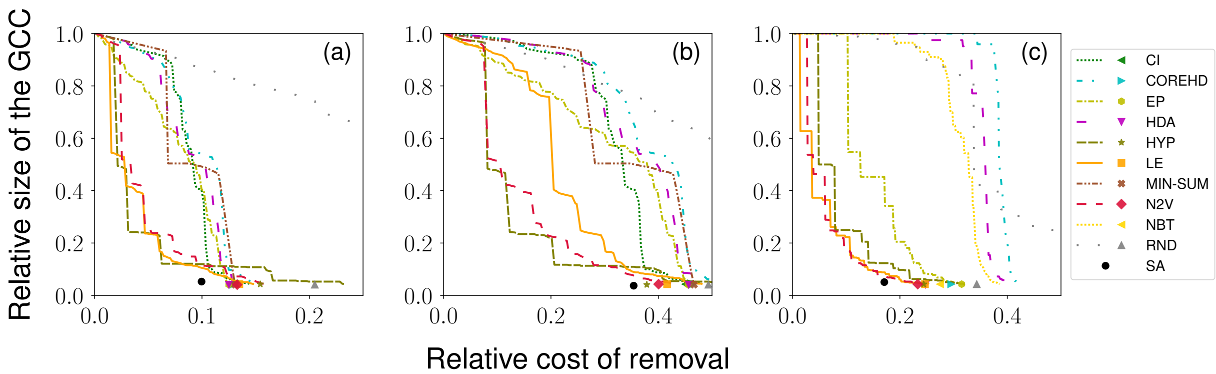

FIG. S15. Same as in Fig. S13, but for the network `S 838`.

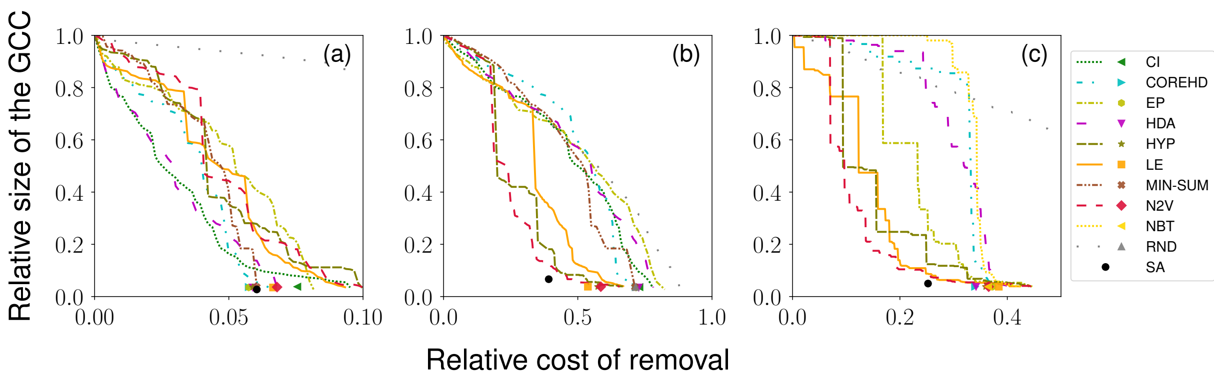

FIG. S16. Same as in Fig. S13, but for the network `Yeast, transcription`.

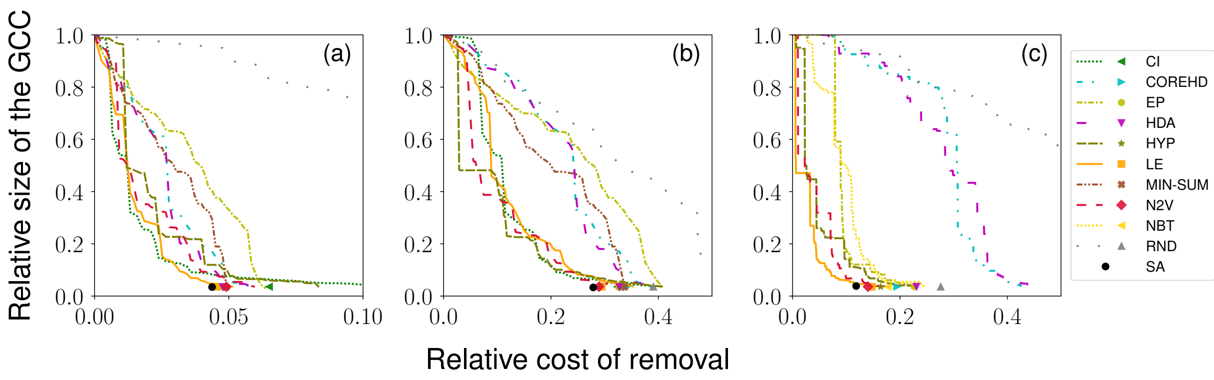

FIG. S17. Same as in Fig. S13, but for the network `Crime`.

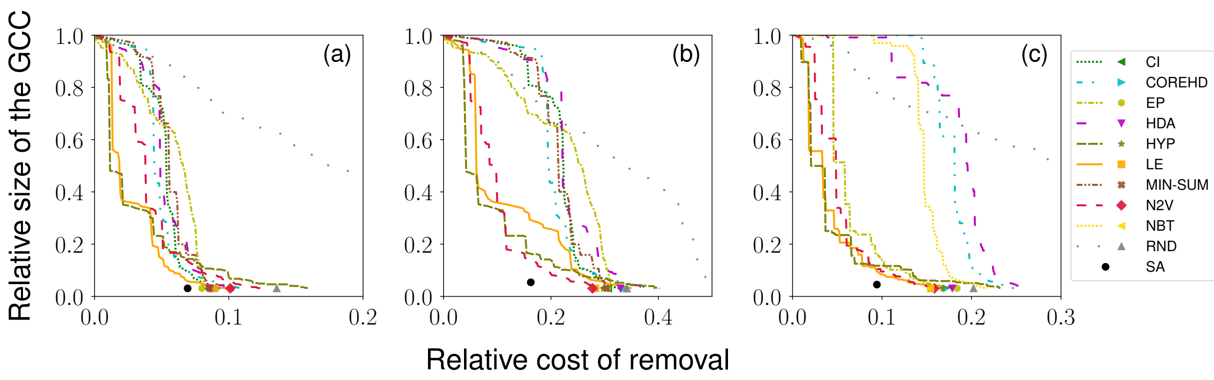

FIG. S18. Same as in Fig. S13, but for the network `Road EU`.

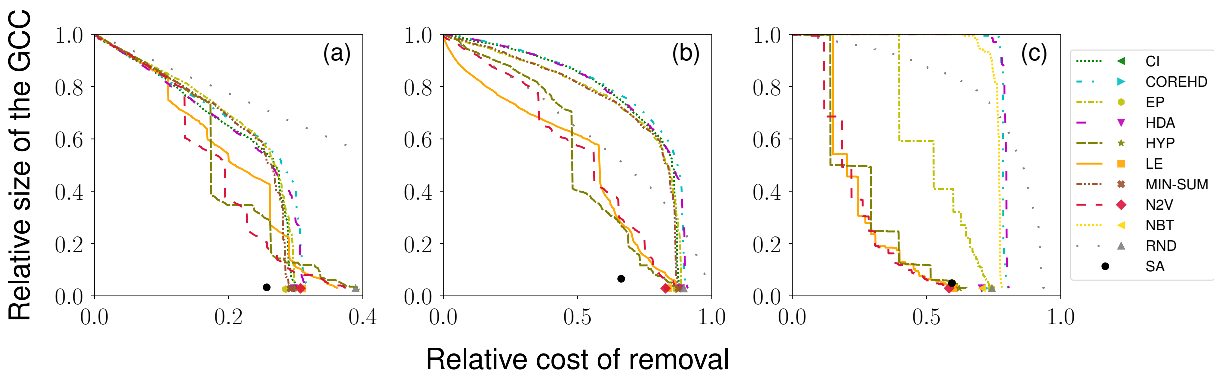

FIG. S19. Same as in Fig. S13, but for the network `URV email`.

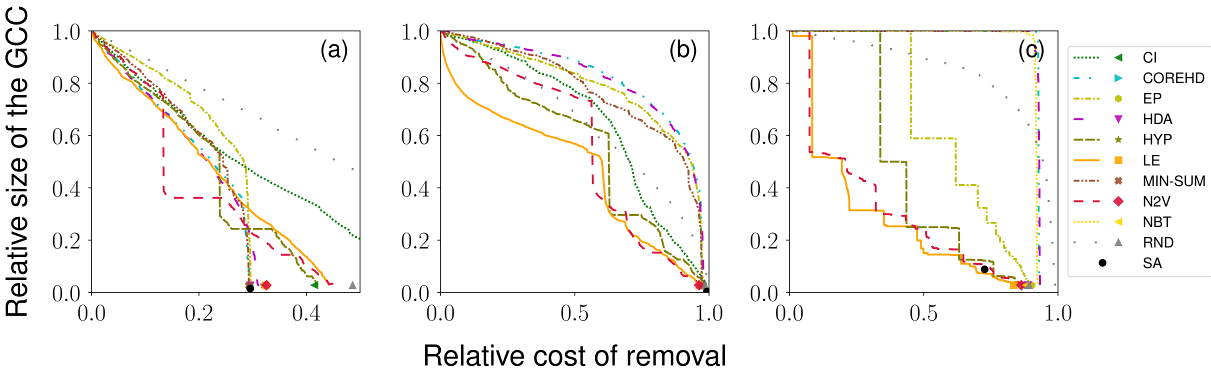

FIG. S20. Same as in Fig. S13, but for the network `Political Blogs`.

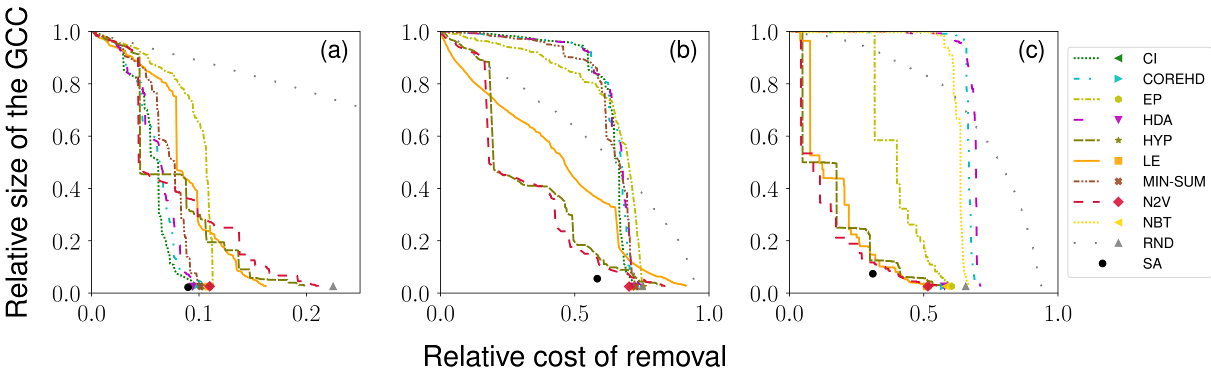

FIG. S21. Same as in Fig. S13, but for the network `Metabolic`.

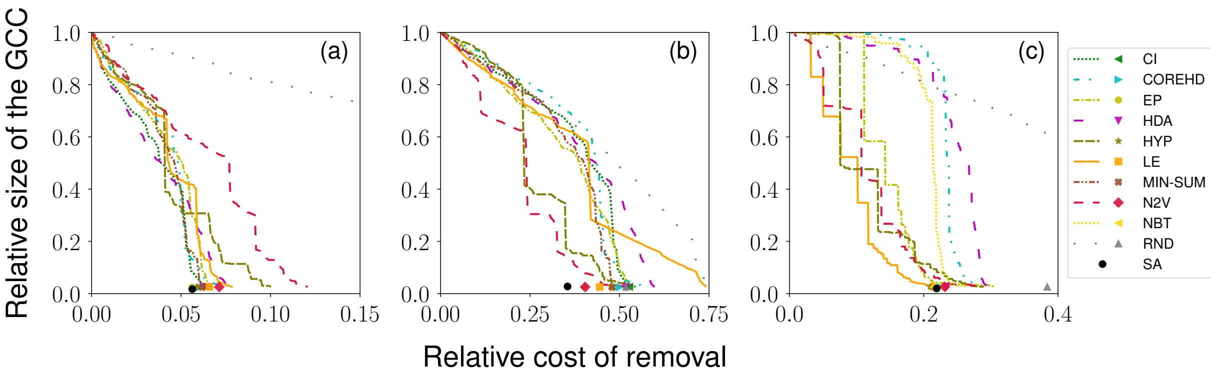

FIG. S22. Same as in Fig. S13, but for the network `Yeast, protein`.

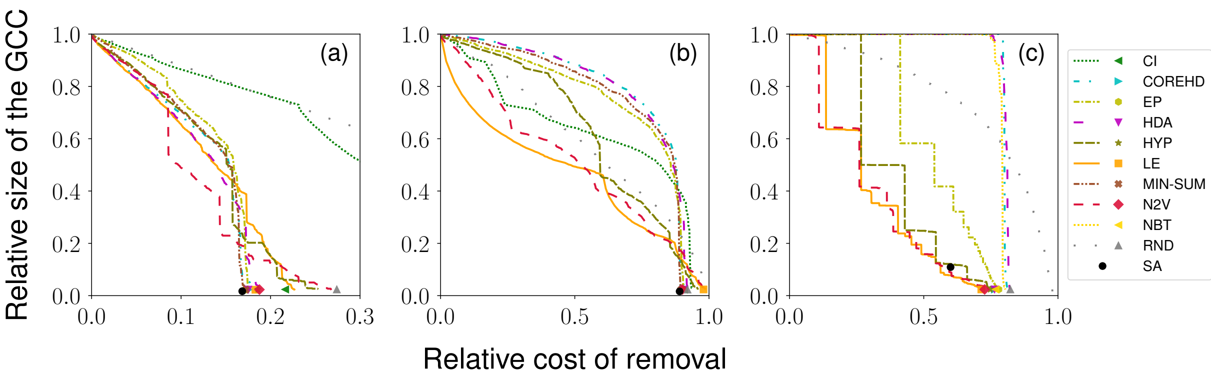

FIG. S23. Same as in Fig. S13, but for the network `Drosophila`.

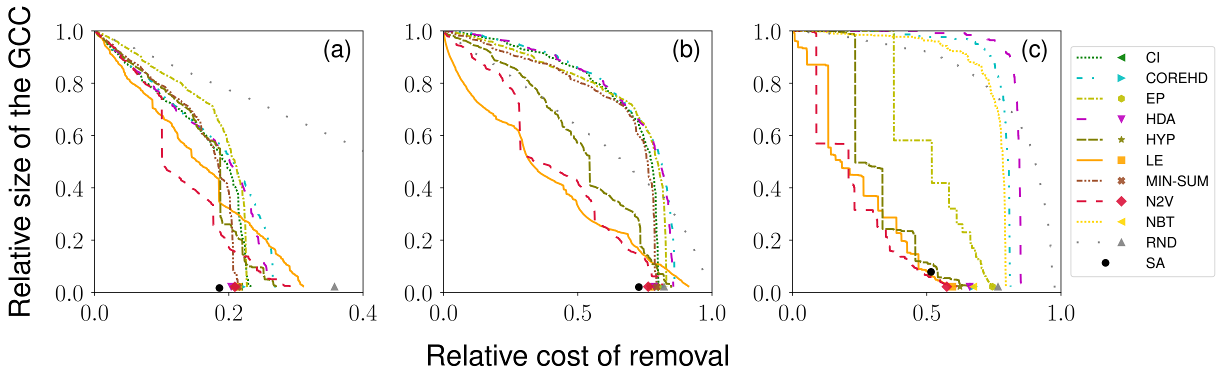

FIG. S24. Same as in Fig. S13, but for the network `Petster-Hamster`.

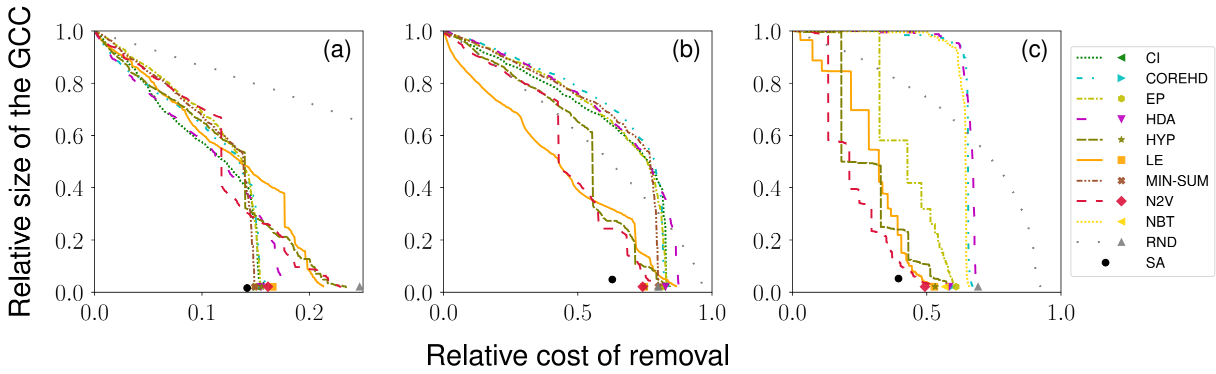

FIG. S25. Same as in Fig. S13, but for the network `Yeast, protein`.

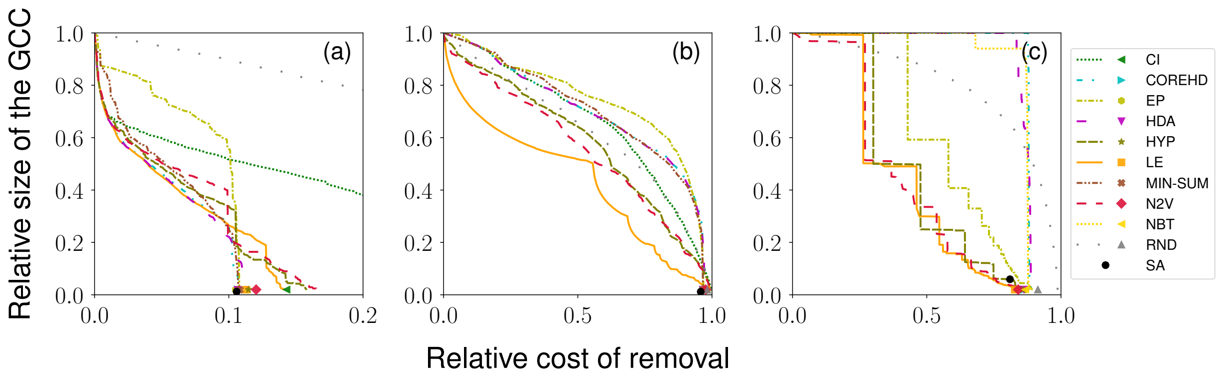

FIG. S26. Same as in Fig. S13, but for the network `Music`.

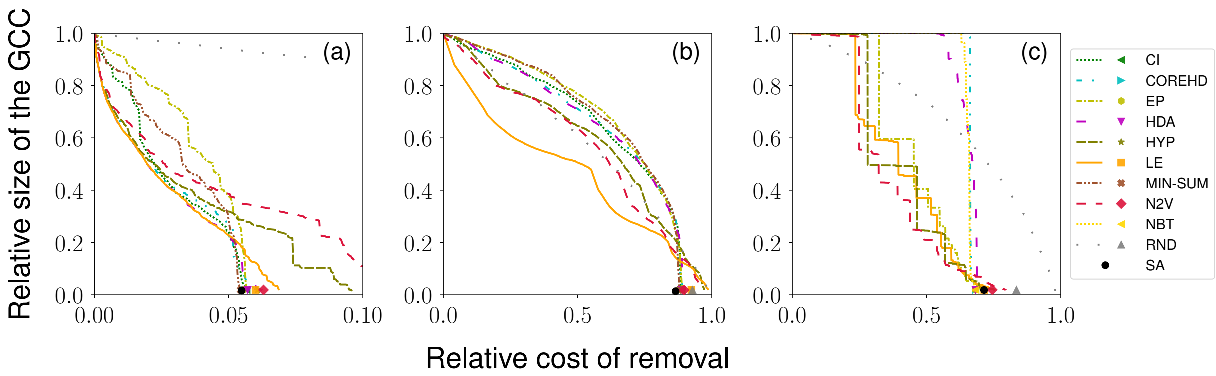

FIG. S27. Same as in Fig. S13, but for the network `Japanese`.

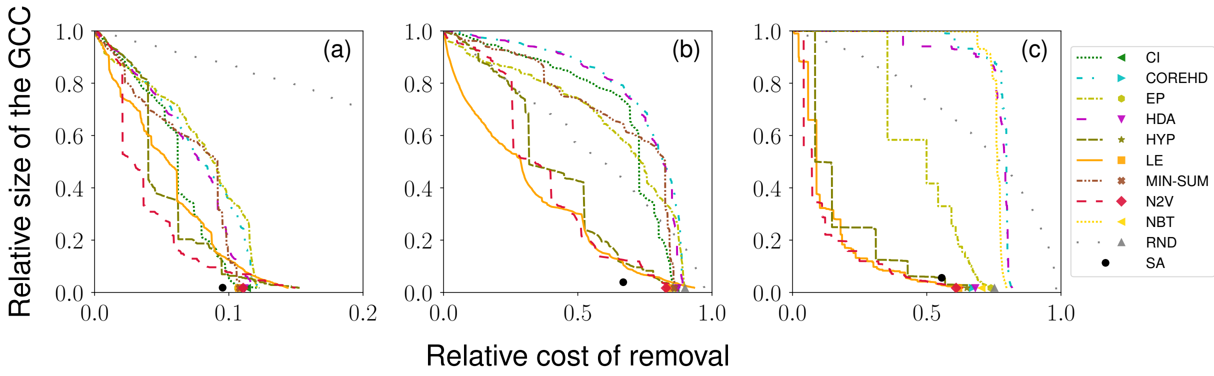

FIG. S28. Same as in Fig. S13, but for the network `Airports`.

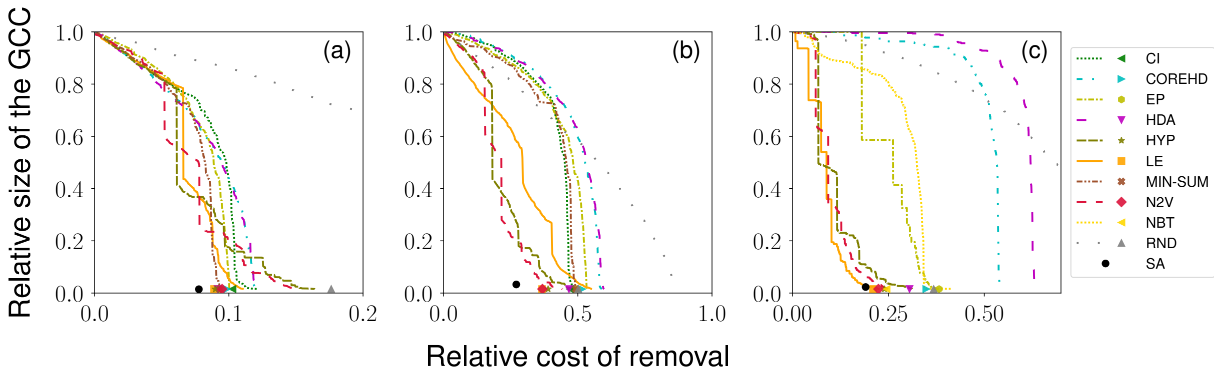

FIG. S29. Same as in Fig. S13, but for the network `GR-QC, 1993-2003`.

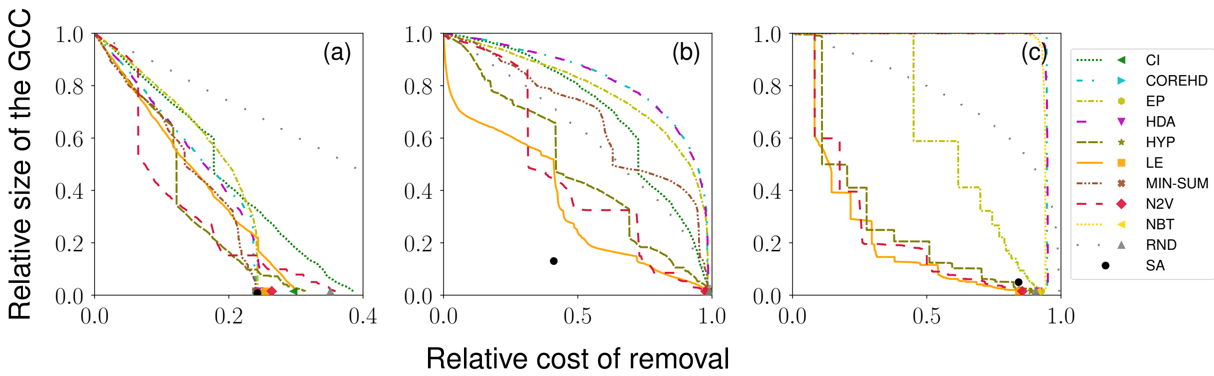

FIG. S30. Same as in Fig. S13, but for the network `Tennis`.

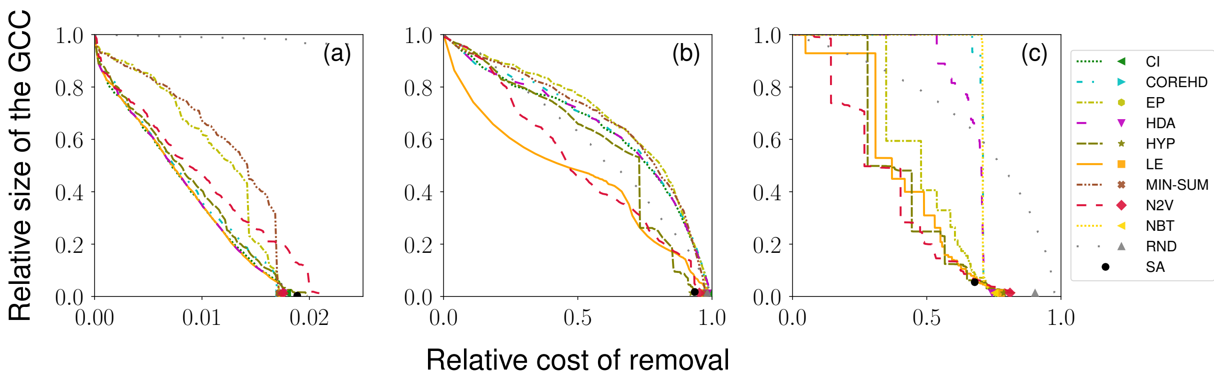

FIG. S31. Same as in Fig. S13, but for the network `HT09`.

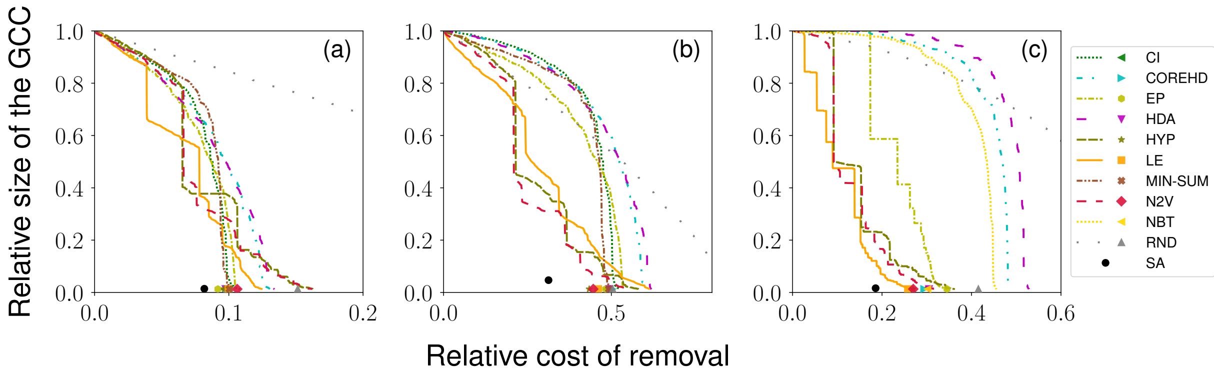

FIG. S32. Same as in Fig. S13, but for the network `Hep-Th, 1995-1999`.

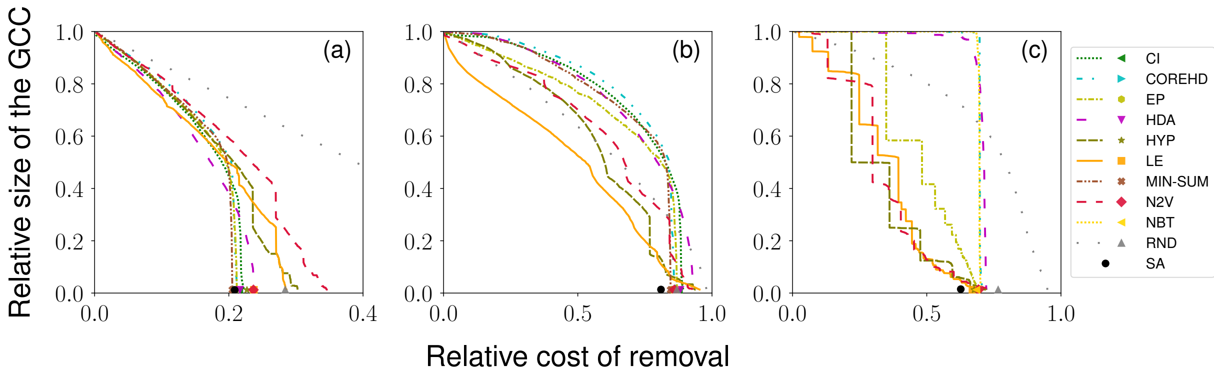

FIG. S33. Same as in Fig. S13, but for the network `Gnutella, Aug. 8, 2002`.

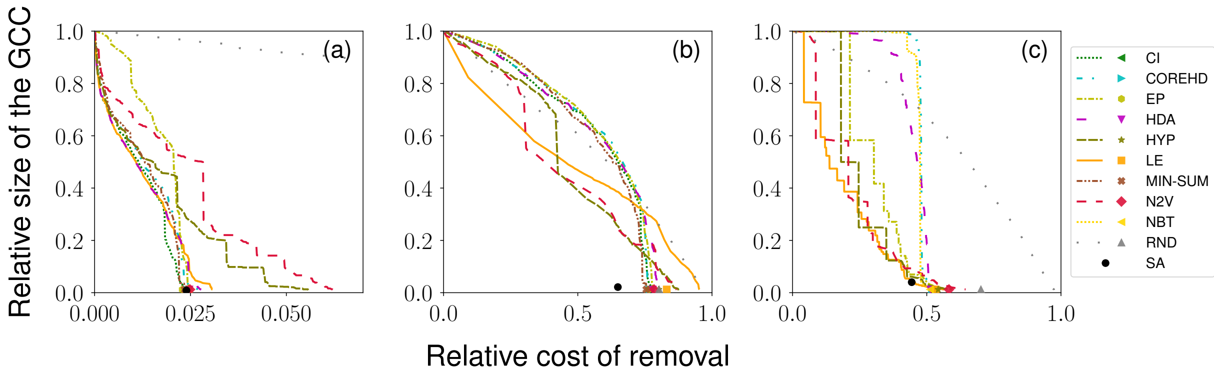

FIG. S34. Same as in Fig. S13, but for the network `AS`.

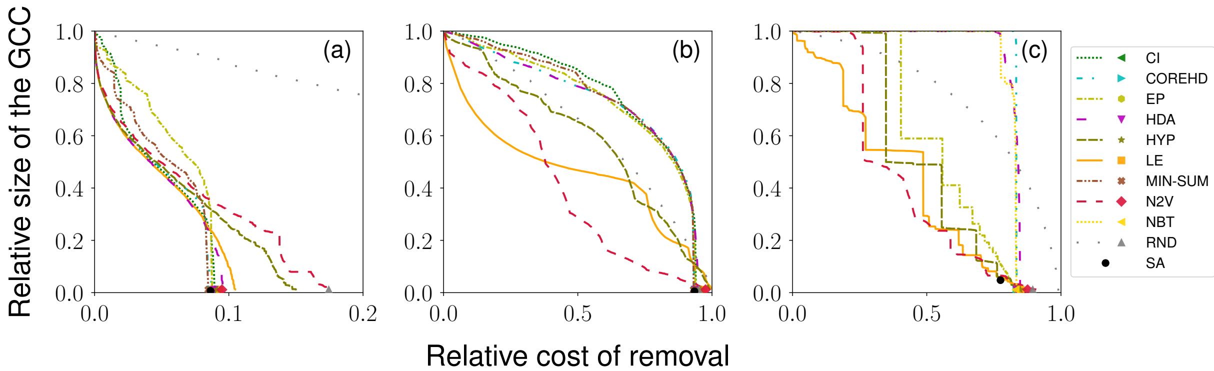

FIG. S35. Same as in Fig. S13, but for the network `English`.

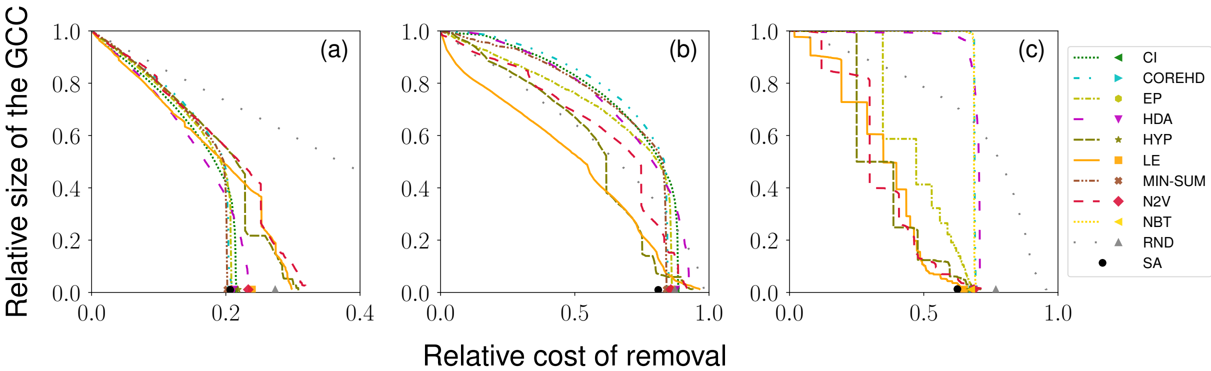

FIG. S36. Same as in Fig. S13, but for the network `Gnutella, Aug. 9, 2002`.

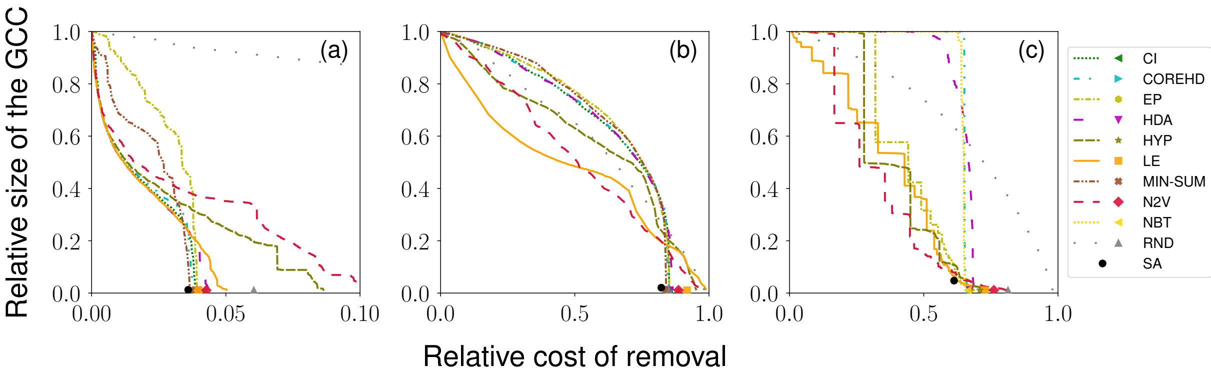

FIG. S37. Same as in Fig. S13, but for the network `French`.

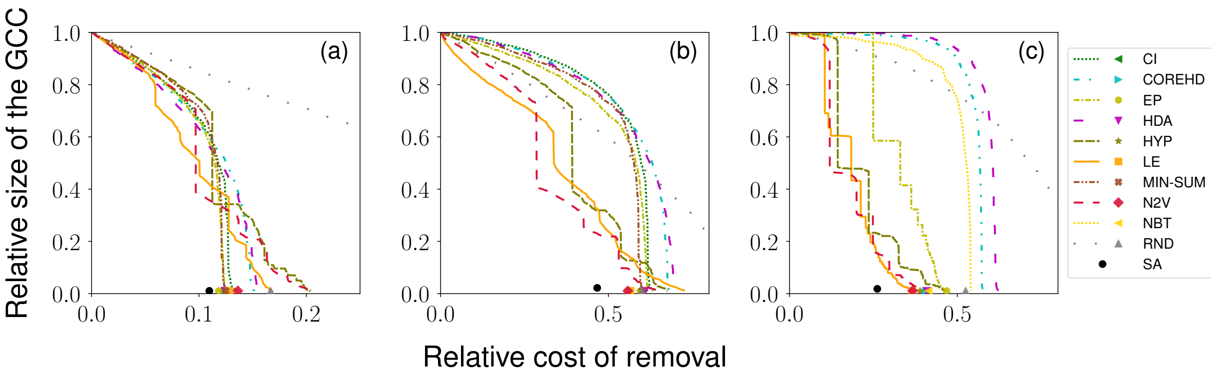

FIG. S38. Same as in Fig. S13, but for the network `Hep-Th, 1993-2003`.

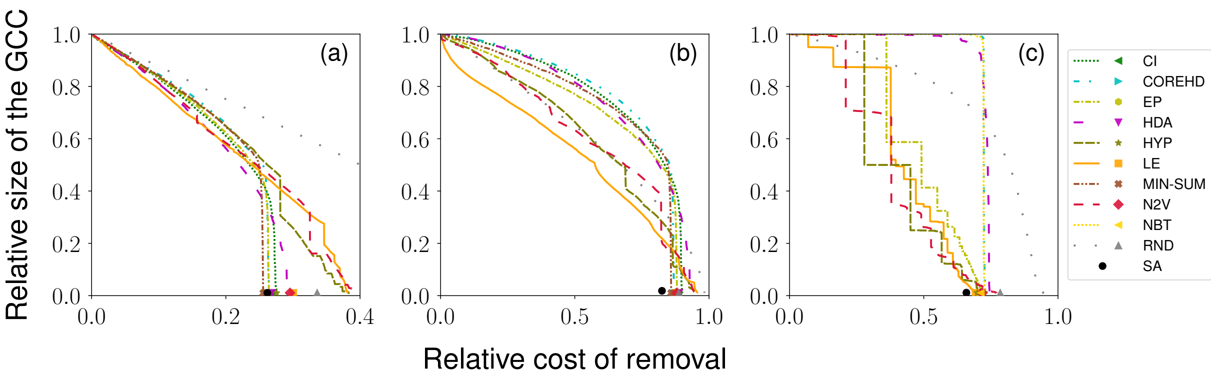

FIG. S39. Same as in Fig. S13, but for the network `Gnutella, Aug. 6, 2002`.

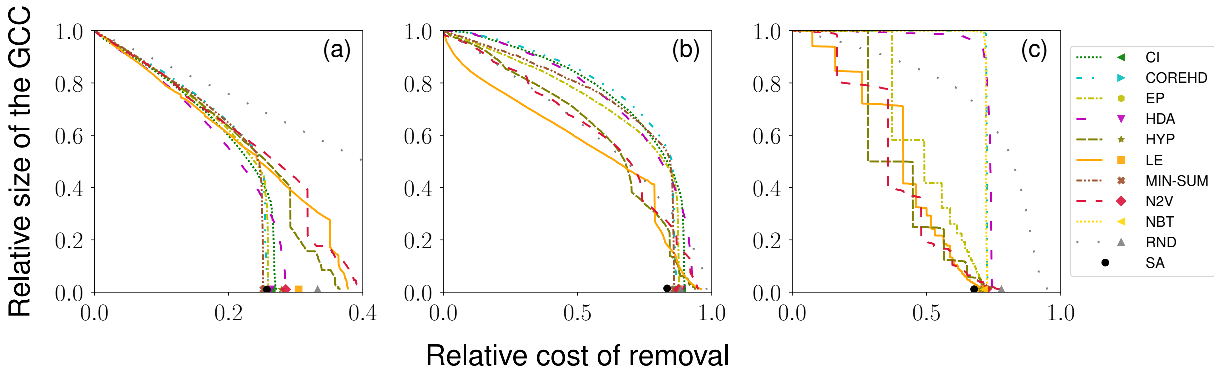

FIG. S40. Same as in Fig. S13, but for the network `Gnutella, Aug. 5, 2002`.

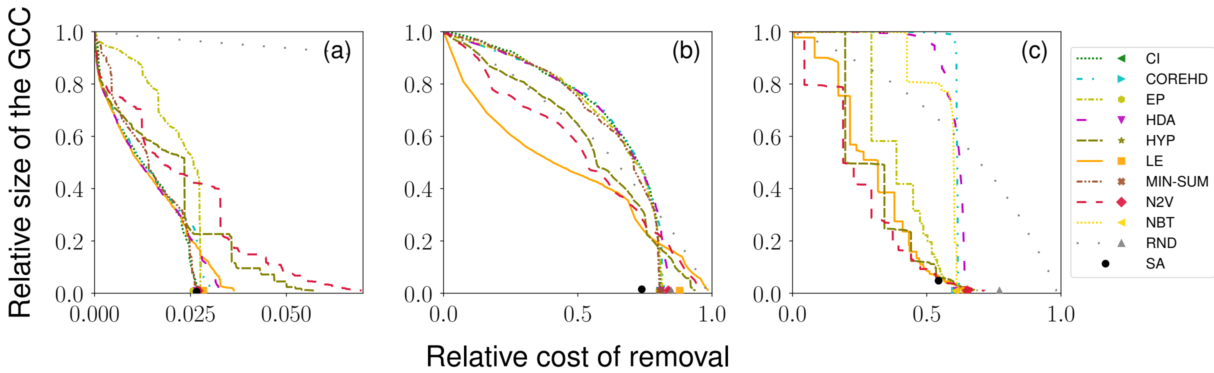

FIG. S41. Same as in Fig. S13, but for the network `IPv6`.

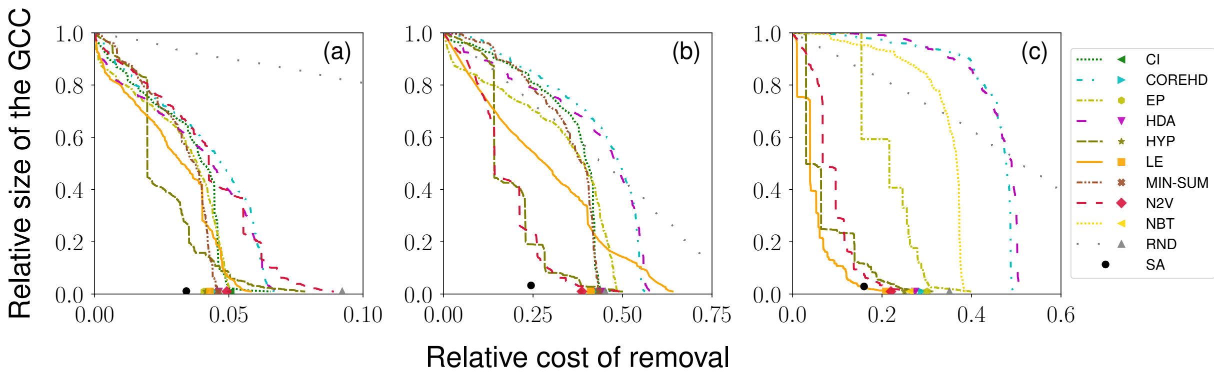

FIG. S42. Same as in Fig. S13, but for the network `PGP`.

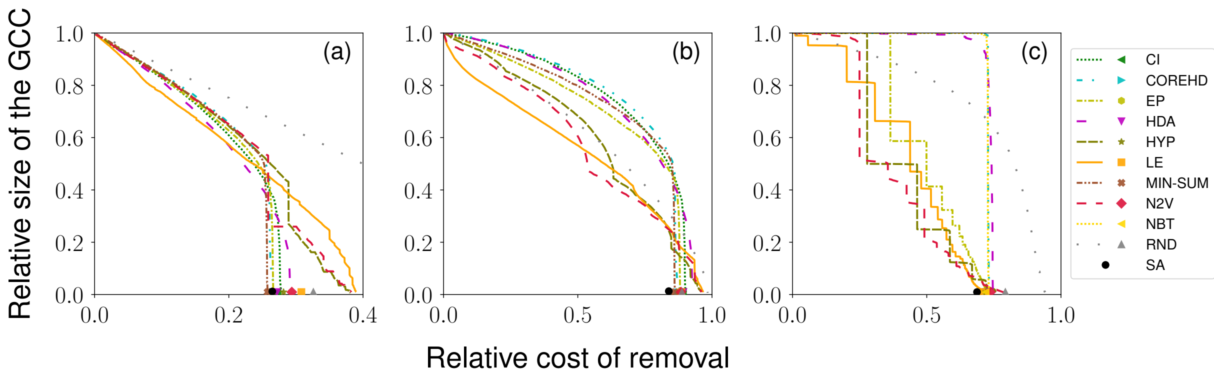

FIG. S43. Same as in Fig. S13, but for the network `Gnutella, August 4 2002`.

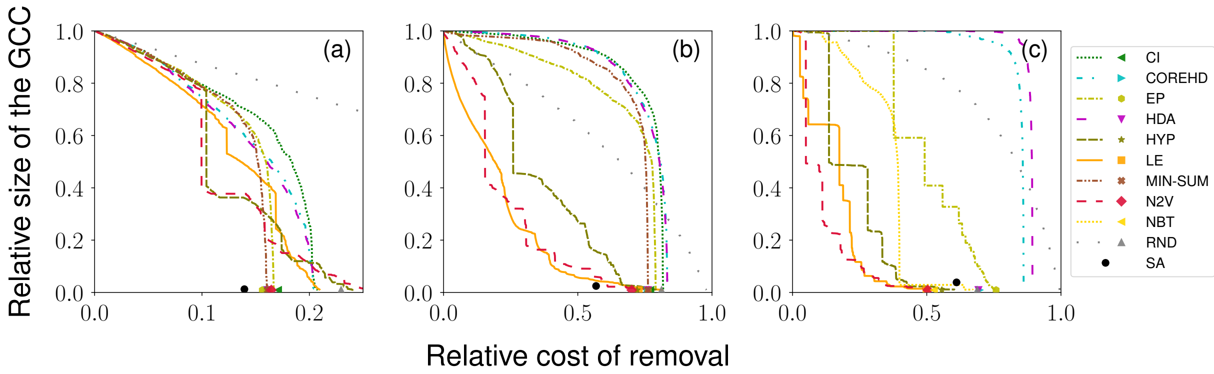

FIG. S44. Same as in Fig. S13, but for the network `Hep-Ph, 1993-2003`.

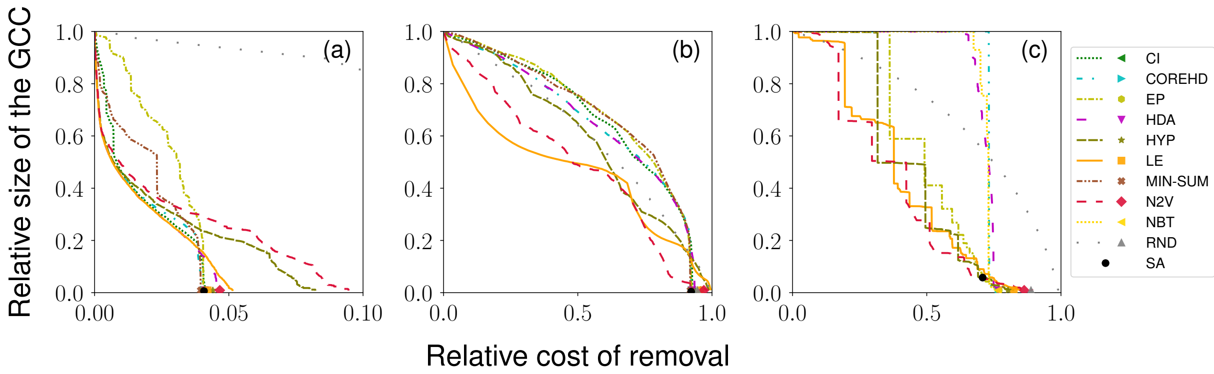

FIG. S45. Same as in Fig. S13, but for the network `Spanish`.

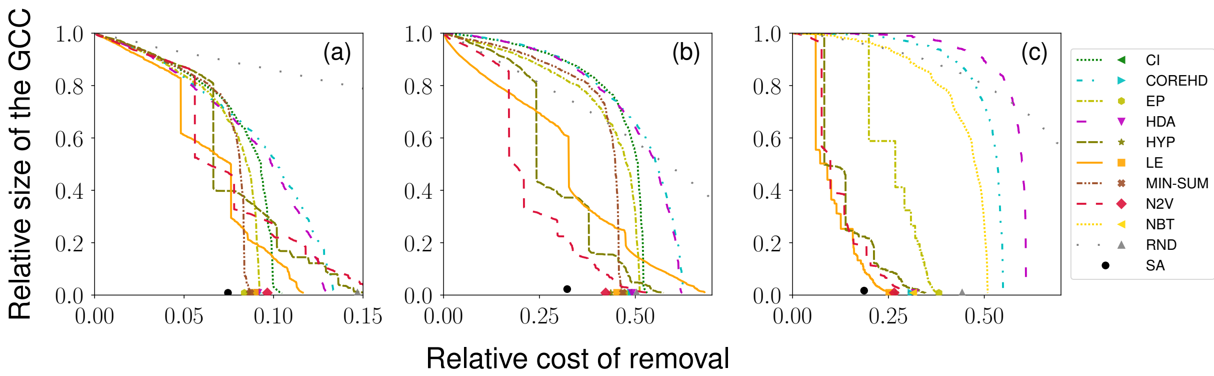

FIG. S46. Same as in Fig. S13, but for the network `Cond-Mat, 1995-1999`.

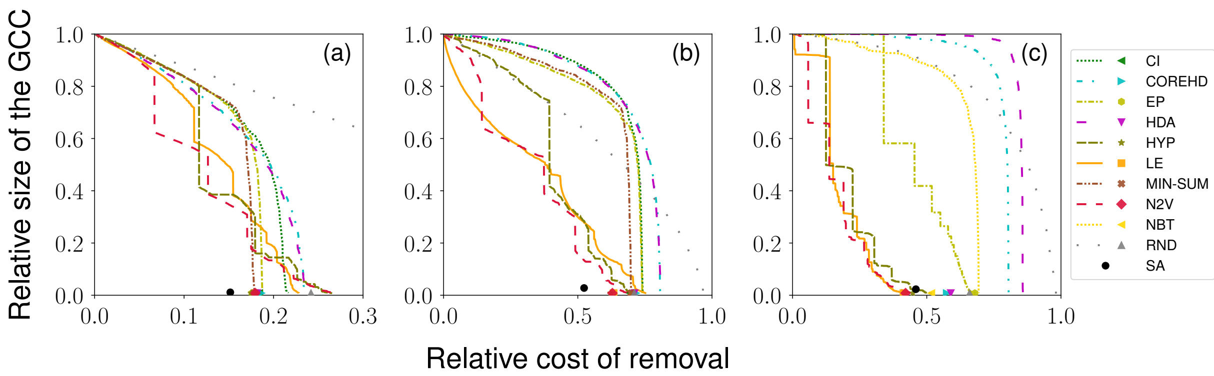

FIG. S47. Same as in Fig. S13, but for the network `Astrophysics`.

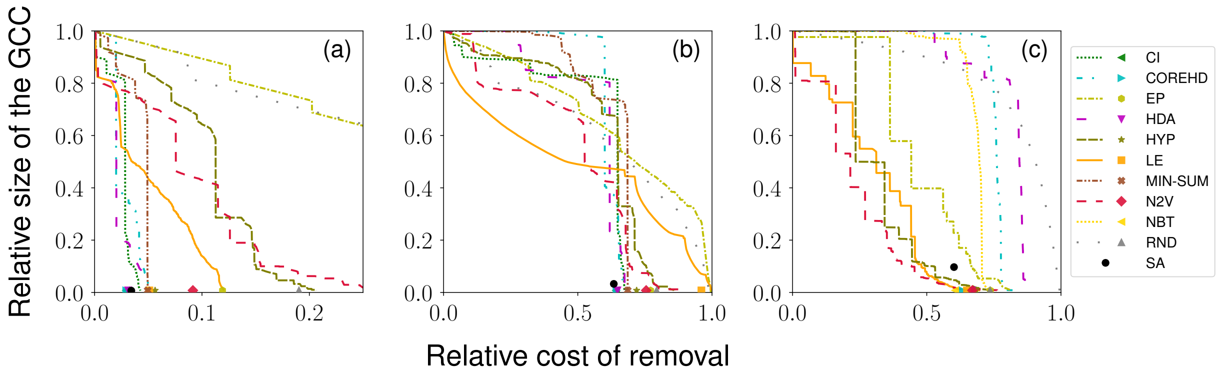

FIG. S48. Same as in Fig. S13, but for the network `Google`.

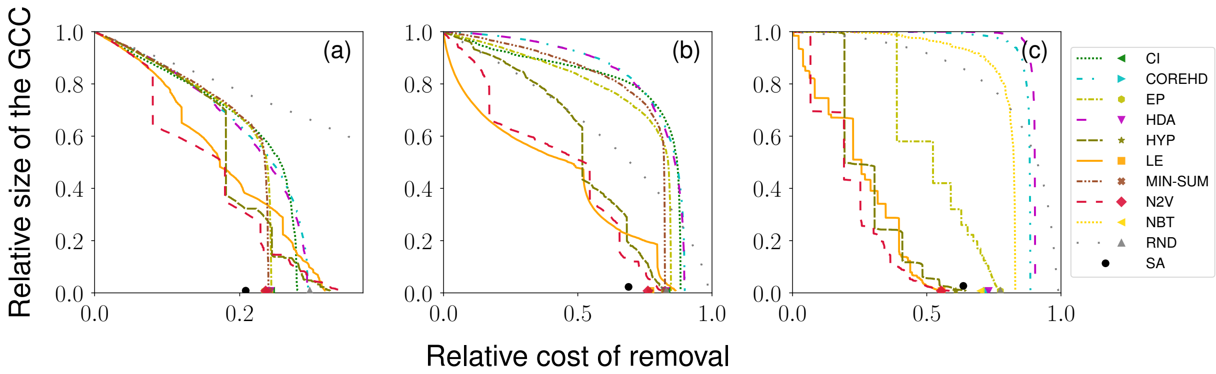

FIG. S49. Same as in Fig. S13, but for the network `ca-AstroPh`.

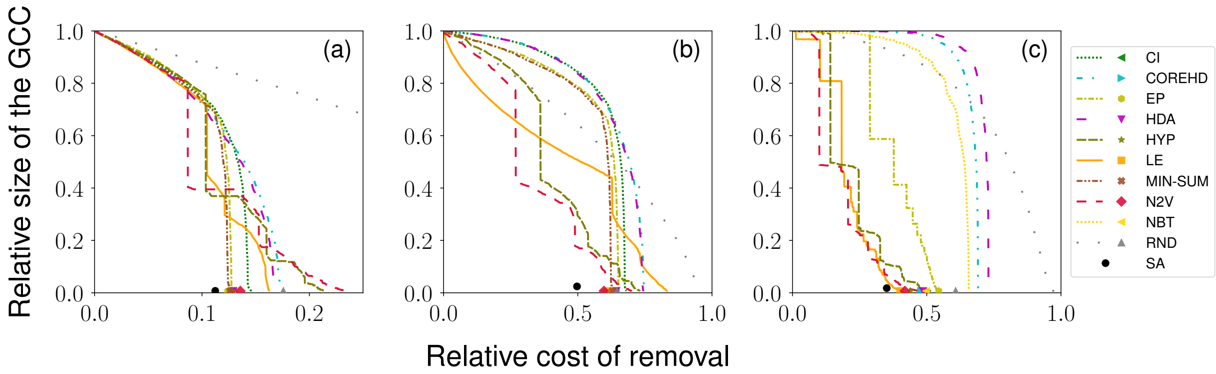

FIG. S50. Same as in Fig. S13, but for the network `Authors`.

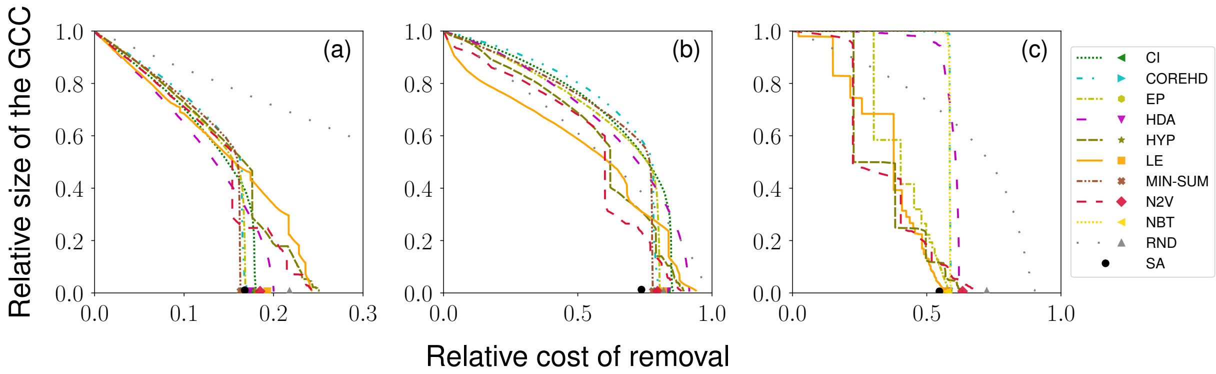

FIG. S51. Same as in Fig. S13, but for the network `Gnutella, Aug. 25, 2002`.

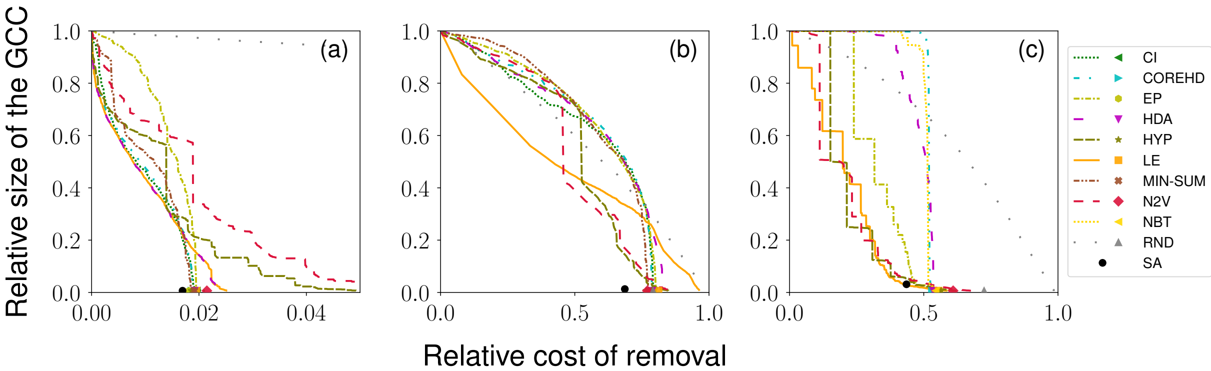

FIG. S52. Same as in Fig. S13, but for the network `Internet`.

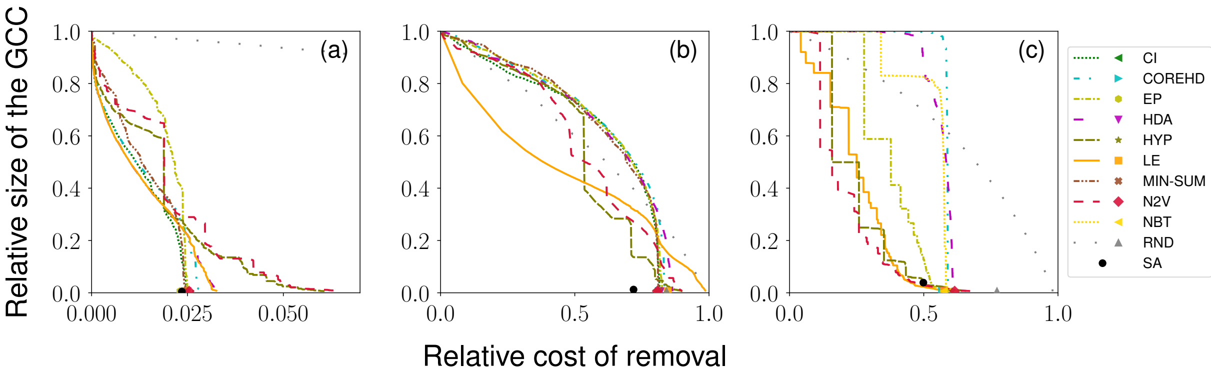

FIG. S53. Same as in Fig. S13, but for the network `Internet`.

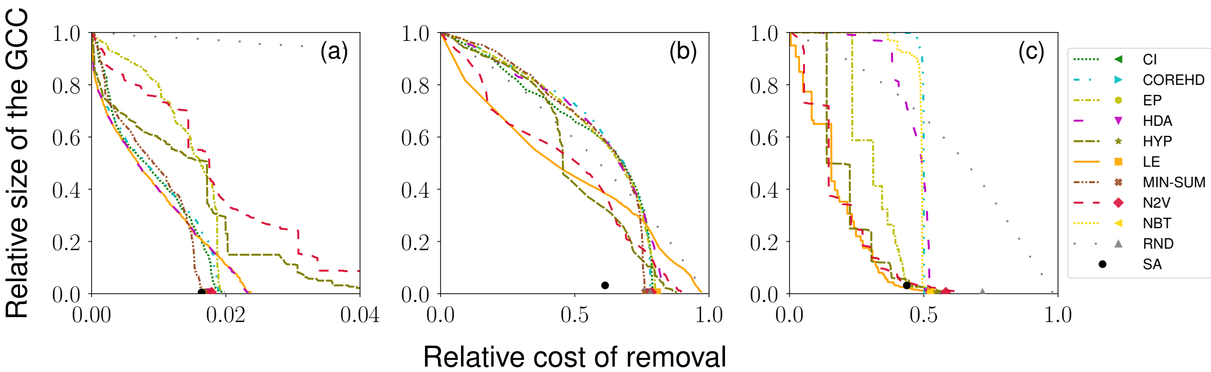

FIG. S54. Same as in Fig. S13, but for the network `caida`.

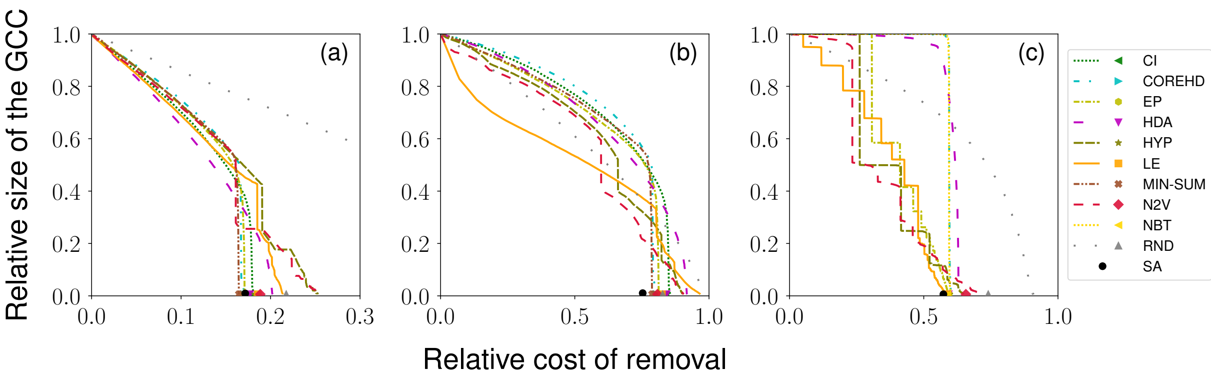

FIG. S55. Same as in Fig. S13, but for the network `Gnutella, Aug. 24, 2002`.

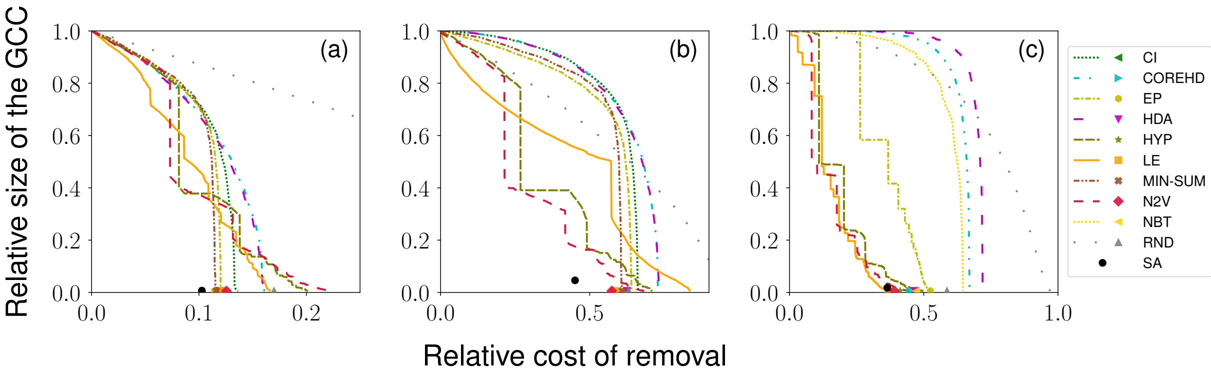

FIG. S56. Same as in Fig. S13, but for the network `Cond-Mat, 1995-2003`.

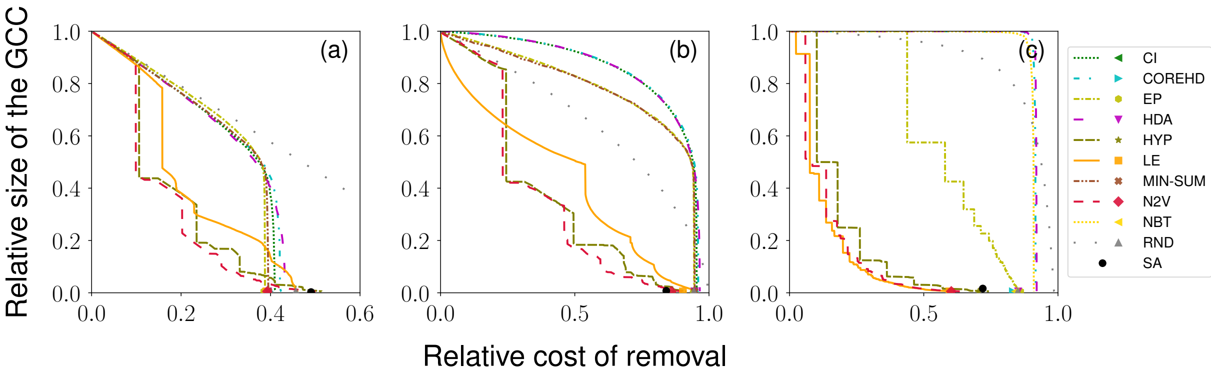

FIG. S57. Same as in Fig. S13, but for the network `Libimseti`.

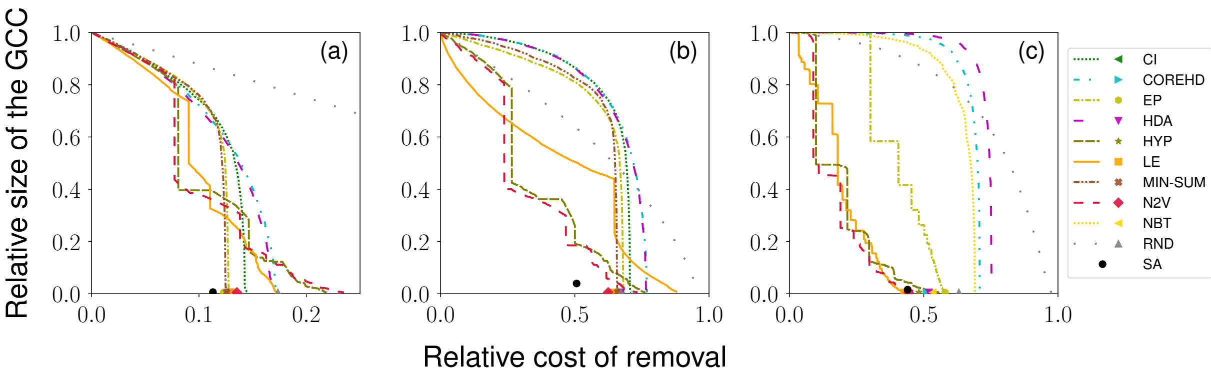

FIG. S58. Same as in Fig. S13, but for the network `Cond-Mat, 1995-2005`.

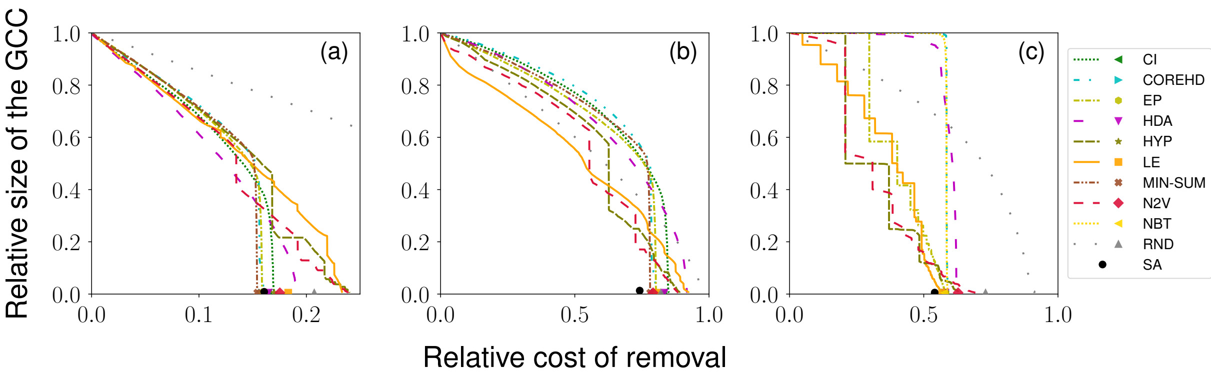

FIG. S59. Same as in Fig. S13, but for the network `Gnutella, Aug. 30, 2002`.

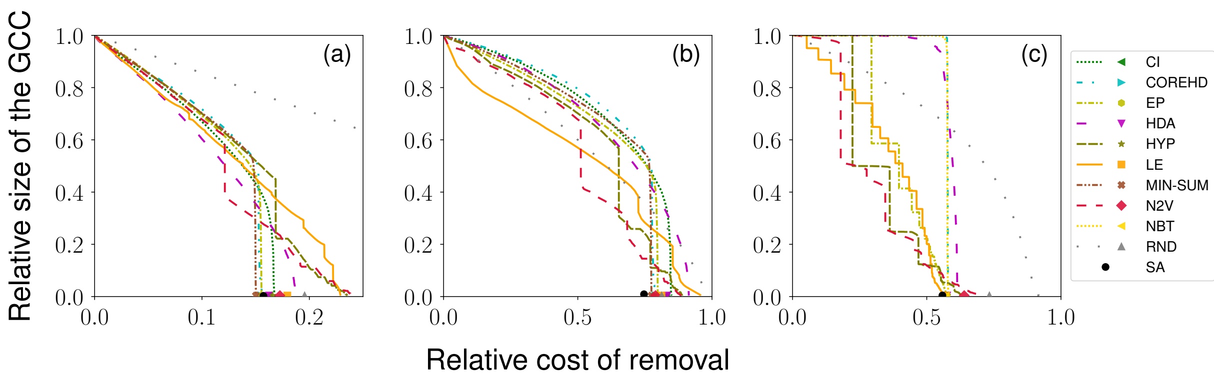

FIG. S60. Same as in Fig. S13, but for the network `Gnutella, Aug. 31, 2002`.

[1] Jérôme Kunegis, "KONECT – The Koblenz Network Collection," in *Proc. Int. Conf. on World Wide Web Companion* (2013) pp. 1343–1350.
[2] Alfredo Braunstein, Luca Dall'Asta, Guilhem Semerjian, and Lenka Zdeborová, "Network dismantling," Proceedings of the National Academy of Sciences **113**, 12368–12373 (2016), https://www.pnas.org/content/113/44/12368.full.pdf.
[3] Haroldo V Ribeiro, Luiz GA Alves, Alvaro F Martins, Ervin K Lenzi, and Matjaž Perc, "The dynamical structure of political corruption networks," Journal of Complex Networks **6**, 989–1003 (2018).
[4] Vittoria Colizza, Romualdo Pastor-Satorras, and Alessandro Vespignani, "Reaction–diffusion processes and metapopulation models in heterogeneous networks," Nature Physics **3**, 276–282 (2007).
[5] Ron Milo, Shalev Itzkovitz, Nadav Kashtan, Reuven Levitt, Shai Shen-Orr, Inbal Ayzenshtat, Michal Sheffer, and Uri Alon, "Superfamilies of evolved and designed networks," Science **303**, 1538–1542 (2004).
[6] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon, "Network motifs: simple building blocks of complex networks," Science **298**, 824–827 (2002).
[7] L. Šubelj and M. Bajec, "Robust network community detection using balanced propagation," The European Physical Journal B **81**, 353–362 (2011).
[8] Roger Guimera, Leon Danon, Albert Diaz-Guilera, Francesc Giralt, and Alex Arenas, "Self-similar community structure in a network of human interactions," Physical review E **68**, 065103 (2003).
[9] Lada A. Adamic and Natalie Glance, "The political blogosphere and the 2004 U.S. election: Divided they blog," in *Proceedings of the 3rd International Workshop on Link Discovery*, LinkKDD '05 (Association for Computing Machinery, New York, NY, USA, 2005) p. 36–43.
[10] M. Ángeles Serrano, Marián Boguñá, and Francesc Sagués, "Uncovering the hidden geometry behind metabolic networks," Mol. BioSyst. **8**, 843–850 (2012).
[11] Hawoong Jeong, Sean P Mason, A-L Barabási, and Zoltan N Oltvai, "Lethality and centrality in protein networks," Nature **411**, 41–42 (2001).
[12] Shin-ya Takemura, Arjun Bharioke, Zhiyuan Lu, Aljoscha Nern, Shiv Vitaladevuni, Patricia K. Rivlin, William T. Katz, Donald J. Olbris, Stephen M. Plaza, Philip Winston, Ting Zhao, Jane Anne Horne, Richard D. Fetter, Satoko Takemura, Katerina Blazek, Lei-Ann Chang, Omotara Ogundeyi, Mathew A. Saunders, Victor Shapiro, Christopher Sigmund, Gerald M. Rubin, Louis K. Scheffer, Ian A. Meinertzhagen, and Dmitri B. Chklovskii, "A visual motion detection circuit suggested by drosophila connectomics," Nature **500**, 175–181 (2013).
[13] Dongbo Bu, Yi Zhao, Lun Cai, Hong Xue, Xiaopeng Zhu, Hongchao Lu, Jingfen Zhang, Shiwei Sun, Lunjiang Ling, Nan Zhang, *et al.*, "Topological structure analysis of the protein–protein interaction network in budding yeast," Nucleic acids research **31**, 2443–2450 (2003).
[14] Joan Serrà, Álvaro Corral, Marián Boguñá, Martín Haro, and Josep Ll. Arcos, "Measuring the evolution of contemporary western popular music," Scientific Reports **2**, 521 (2012).
[15] Thomas Rolland, Murat Taşan, Benoit Charloteaux, Samuel J. Pevzner, Quan Zhong, Nidhi Sahni, Song Yi, Irma Lemmens, Celia Fontanillo, Roberto Mosca, Atanas Kamburov, Susan D. Ghiassian, Xinping Yang, Lila Ghamsari, Dawit Balcha, Bridget E. Begg, Pascal Braun, Marc Brehme, Martin P. Broly, Anne-Ruxandra Carvunis, Dan Convery-Zupan, Roser Corominas, Jasmin Coulombe-Huntington, Elizabeth Dann, Matija Dreze, Amélie Dricot, Changyu Fan, Eric Franzosa, Fana Gebreab, Bryan J. Gutierrez, Madeleine F. Hardy, Mike Jin, Shuli Kang, Ruth Kiros, Guan Ning Lin, Katja Luck, Andrew MacWilliams, Jörg Menche, Ryan R. Murray, Alexandre Palagi, Matthew M. Poulin, Xavier Rambout, John Rasla, Patrick Reichert, Viviana Romero, Elien Ruyssinck, Julie M. Sahalie, Annemarie Scholz, Akash A. Shah, Amitabh Sharma, Yun Shen, Kerstin Spirohn, Stanley Tam, Alexander O. Tejeda, Shelly A. Trigg, Jean-Claude Twizere, Kerwin Vega, Jennifer Walsh, Michael E. Cusick, Yu Xia, Albert-László Barabási, Lilia M. Iakoucheva, Patrick Aloy, Javier De Las Rivas, Jan Tavernier, Michael A. Calderwood, David E. Hill, Tong Hao, Frederick P. Roth, and Marc Vidal, "A proteome-scale map of the human interactome network," Cell **159**, 1212–1226 (2014).
[16] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos, "Graph evolution: Densification and shrinking diameters," ACM Transactions on Knowledge Discovery from Data (TKDD) **1**, 2 (2007).
[17] Filippo Radicchi, "Who is the best player ever? a complex network analysis of the history of professional tennis," PloS one **6**, e17249 (2011).
[18] Duncan J Watts and Steven H Strogatz, "Collective dynamics of 'small-world'networks," nature **393**, 440–442 (1998).
[19] Lorenzo Isella, Juliette Stehlé, Alain Barrat, Ciro Cattuto, Jean-François Pinton, and Wouter Van den Broeck, "What's in a crowd? analysis of face-to-face behavioral networks," Journal of theoretical biology **271**, 166–180 (2011).
[20] Mark EJ Newman, "The structure of scientific collaboration networks," Proceedings of the National Academy of Sciences **98**, 404–409 (2001).
[21] Matei Ripeanu, Ian Foster, and Adriana Iamnitchi, "Mapping the gnutella network: Properties of large-scale peer-to-peer systems and implications for system design," arXiv preprint cs/0209028 (2002).
[22] K Claffy, Y Hyun, K Keys, M Fomenkov, and D Krioukov, "Catch," (2009).
[23] Kimberly Claffy, Young Hyun, Ken Keys, Marina Fomenkov, and Dmitri Krioukov, "Internet mapping: from art to science," in *IEEE DHS CATCH* (IEEE, 2009).
[24] Marián Boguñá, Romualdo Pastor-Satorras, Albert Díaz-Guilera, and Alex Arenas, "Models of social networks based on social distance attachment," Physical Review E **70**, 056122 (2004).

[25] Gergely Palla, Illes J Farkas, Peter Pollner, Imre Derenyi, and Tamás Vicsek, "Directed network modules," New Journal of Physics **9**, 186 (2007).

[26] Lukas Brozovsky and Vaclav Petricek, "Recommender system for online dating service," arXiv preprint cs/0703042 (2007).

[27] Jérôme Kunegis, Gerd Gröner, and Thomas Gottron, "Online dating recommender systems: The split-complex number approach," in *Proceedings of the 4th ACM RecSys workshop on Recommender systems and the social web* (ACM, 2012) pp. 37–44.