Journal of Complex Networks (2021) **1**, Advance Access Publication on 24 December 2021
https://doi.org/10.1093/comnet/cnab050

The robustness of the photosynthetic system I energy transfer complex network to targeted node attack and random node failure

M. BELLINGERI[†]
Dipartimento di Scienze Matematiche, Fisiche e Informatiche, Università di Parma, via G.P. Usberti, 7/a, 43124 Parma, Italy, Dipartimento di Fisica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy and INFN, Gruppo Collegato di Parma, I-43124 Parma, Italia

D. MONTEPIETRA
Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università di Modena e Reggio Emilia, via Campi, 213/a, 41125 Modena, Italy and CNR NANO S3, Via Campi 213/A, 41125 Modena, Italy

D. CASSI
Dipartimento di Scienze Matematiche, Fisiche e Informatiche, Università di Parma, via G.P. Usberti, 7/a, 43124 Parma, Italy and INFN, Gruppo Collegato di Parma, I-43124 Parma, Italia

AND

F. SCOTOGNELLA
Dipartimento di Fisica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy and Center for Nano Science and Technology@PoliMi, Istituto Italiano di Tecnologia, Via Giovanni Pascoli 70/3, 20133 Milan, Italy
[†]Corresponding author. Email: michele.bellingeri@polimi.it

Edited by: Jose Mateos

[Received on 14 June 2021; editorial decision on 9 November 2021; accepted on 2 December 2021]

In this article, we implement and compare 10 node removal (attack) strategies from the literature over the photosystem I (PSI) complex network of the common pea plant (*Pisum sativum*), representing the FRET energy transfer among its nodes/chromophores. We measure the network robustness (functioning) with four indicators. The node attack strategies and the network robustness indicators consider both the binary-topological and the weighted structure of the network. First, we find that the well-known node betweenness centrality attack, which has proven highly effective in dismantling most real-world networks' topological connectivity, is ineffective over the PSI network. Second, the degeneracy of the node properties caused by the PSI's higher network connectivity level induces a random-like node removal even when nodes are removed according to a specific node centrality measure. This phenomenon triggers a very low decrease of the PSI network functioning even when subjected to node attack. Such an outcome would indicate that the node attack strategies based on classic node properties, such as the degree or the betweenness centrality, may show low efficacy in dismantling real-world networks with very high connectivity levels.

Last, the PSI network can be built by tuning a cut-off distance (CD) that defines the viable energy transfers among nodes/chromophores and progressively discards the lower energy transfer links among distant nodes/chromophores. This represents a 'weight thresholding' procedure allowing us to investigate the efficacy of the node attack strategies when links of lower weight are progressively pruned from the PSI network. We find that the best node attack strategies change by decreasing the CD, showing that the weight thresholding procedure affects the network response to node removal. This last outcome outlines the importance of investigating the stability of the system response for real-world weighted complex networks subjected to the weight thresholding procedure.

Keywords: complex biological networks; energy transfer; photosynthetic network; network robustness; photosystem I

1. Introduction

Network science can model a variety of real-world systems, yielding valuable insight in the fields of social network analysis [1–3], economics [4], urban and international transport [5–7], ecology [8–10], psychology [11, 12], biology [13], infrastructure [14] and finance [15, 16].

One of the main topics in network science is the investigation of the network functioning robustness to random node removal (node failure) or targeted node removal (intentional node attack) [17–24]. Implementing node attacks in complex networks helps to describe a variety of real problems [5, 8–10, 23, 24]. Given this wide range of practical applications, analysing network robustness to node attack or, conversely, finding the best node attack strategy has been an intensely investigated question in the last decade [17–22].

The approach to this problem is straightforward: nodes are removed from the network according to some properties, and meanwhile, the network functioning decrease is traced according to a specific functioning indicator [18], the most used being the largest connected component (*LCC* or giant cluster) and the network efficiency (*EFF*) [18].

Many node attack strategies based on node properties (centralities) have also been crafted to efficiently decrease the network functioning (robustness), such as node degree, betweenness centrality, closeness centralities, and many others [17–22]. A recent and exhaustive research comparison showed that, on average, the classic betweenness centrality is the most effective node attack strategy in decreasing the network functioning, with both *LCC* and the *EFF* as indicators [18]. Besides this average result, the authors showed that the strategy efficacy might vary between different systems, leaving open the problem of finding the best node attack strategies for new real-world networks. However, Wandelt *et al.*'s comparison [18] focused only on binary-topological networks, neglecting the evaluation of the effect of link weights on the network functioning. Research results revealed that ignoring link weights may change the network robustness and effectiveness of the attack strategies, outlining the necessity of testing the system robustness of weighted networks [25–27].

In this article, we implement node attack strategies over a recent and newly assembled biological network, that is, the photosynthetic system I (PSI) energy transfer complex network of the common pea plant *Pisum sativum* [28], to assess the robustness of this real-world system. Photosystem I is a membrane protein-chromophore complex functioning as a light-driven electron pump within the oxygenic photosynthetic process [29, 30]. The PSI is a weighted and directed network in which the nodes represent the chlorophyll and carotenoid molecules, while the links between the nodes model the energetic coupling between the chromophores (see Montepietra *et al.* [28] for details).

We implement 10 node attack strategies from literature, considering the binary-topological and the weighted structure of the network. We measure the system robustness using four indicators of network functioning, considering the directionality and the weight of the links.

We find that the classic node attack based on the node betweenness centrality, which is high performing in dismantling real-world networks, is ineffective over the PSI network. Further, we observe that the node attack strategies considering the link weights are slightly more effective in decreasing the indicator of network functioning, both for binary and weighted indicators.

Last, when performing the node removals over the PSI network by progressively discarding the lower energy transfer links among distant node/chromophores, we discover the changes in the PSI system response to node attack and the efficacy of the node attack strategies. This brings interesting suggestions within the 'weight thresholding' problem [31, 32] in weighted complex networks by showing that the weight thresholding procedure affects the PSI network response to node removal.

2. Methods

An unweighted network $G = (N, L)$, of N nodes and L links, can be completely described by the adjacency matrix A, a $N \times N$ square matrix whose entry $a_{ij}(i, j = 1, ..., N)$ is equal to 1 when there is a link l_{ij} from i to j, and zero otherwise. A weighted network $G^w = (N, L, W)$ can be described by a weights matrix W, a $N \times N$ matrix whose entry w_{ij} is the weight of the link connecting node i to node j; $w_{ij} = 0$ if the nodes i and j are not connected, and $w_{ij} > 0$ otherwise.

2.1 *The node removal strategies*

RAND: Random removal of nodes. It represents the possibility of node failure (error) in the network [7, 22].

DEG: Removal of nodes according to their degree, that is, the number of links to the node [17, 18, 20]. The degree k_i of node i is given by:

$$k_i = \sum_{j=1}^{N} a_{ij}, \tag{1}$$

where a_{ij} is 1 if there is a link connecting nodes i and j and is 0 otherwise, the term N means the sum is calculated over all nodes in the network. The *DEG* strategy removes nodes with higher topological connectivity in the network, usually called hubs [33, 34].

InDEG: Removal of nodes according to decreasing order of in-degree, that is, the number of ingoing (entering) links of the node [33, 34].

In formula:

$$k_i^{in} = \sum_{j=1}^{N} a_{ji} \tag{2}$$

where a_{ij} is 1 if there is a directed link from node j to node i and is 0 otherwise, the term N means the sum is calculated over all nodes in the network.

OutDEG: Removal of nodes according to decreasing order of out-degree, that is, the number of outgoing (exiting) links from the node [33, 34].

In formula:

$$k_i^{out} = \sum_{j=1}^{N} a_{ij}, \tag{3}$$

where a_{ij} is 1 if there is a directed link from node i to node j and is 0 otherwise, the term N means the sum is calculated over all nodes in the network.

STR: Removal of nodes according to decreasing order of strength, that is, the sum of the weights of the node links [22, 25, 35]. In the formula, the strength s_i of the node i is:

$$s_i = \sum_{j=1}^{N} a_{ij} \cdot w_{ij}, \tag{4}$$

where a_{ij} is 1 if there is a link joining nodes i and j and 0 otherwise, and w_{ij} is the weight of this connection between nodes i and j. The strength s_i is also named 'the weighted degree' of the node. For this reason, *STR* can be viewed as the weighted counterpart of *DEG*.

InSTR: Removal of nodes according to decreasing order of in-strength, that is, the sum of the weights of its ingoing links [33]. The *InSTR* can be viewed as the weighted counterpart of *InDEG*.

The in-strength s_i^{in} is:

$$s_i^{in} = \sum_{j=1}^{N} a_{ji} \cdot w_{ji}, \tag{5}$$

where a_{ij} is 1 if there is a directed link from node j to node i and is 0 otherwise, the term N means the sum is calculated over all nodes in the network.

OutSTR: Removal of nodes according to decreasing order of out-strength, that is, the sum of the weights of the outgoing links [33]. The *OutSTR* can be viewed as the weighted counterpart of *OutDEG*.

The out-strength s_i^{out} is:

$$s_i^{out} = \sum_{j=1}^{N} a_{ij} \cdot w_{ij}, \tag{6}$$

where a_{ij} is 1 if there is a directed link from node i to node j and is 0 otherwise, the term N means the sum is calculated over all nodes in the network.

BC: Removal of nodes according to decreasing order of betweenness centrality. The betweenness centrality is based on the shortest paths between node pairs (also called geodesic paths). The shortest path between two nodes is the minimum number of links required to travel from a node to the other [17–20]. The betweenness centrality of a node returns the number of shortest paths from every node pair of the network passing along that node [17–20]. The betweenness $g(i)$ of the node i is:

$$g(i) = \sum_{s,t=1}^{N} \frac{\sigma_{st}(i)}{\sigma_{st}}, \tag{7}$$

where σ_{st} is the total number of shortest paths from node s to node t and $\sigma_{st}(i)$ is the number of these shortest paths crossing the node i, summed over all network nodes N.

The *BC* defined here is based on the binary shortest path (also named hop distance [33, 34]), accounting for the necessary number of links to pass between nodes, neglecting the attached link weights. For this reason, it is also called binary betweenness centrality.

BCw: Removal of nodes according to decreasing order of weighted betweenness centrality. The weighted betweenness centrality is computed using the weighted shortest paths (WSP), which consider the number of links necessary to travel between nodes and the weight attached to the links. The node *BCw* counts the WSP from any node pairs passing through that node (also called weighted geodesic) [26, 33]. In this procedure, we first need to consider whether the link weights are 'flows or costs'. If the link weights are flows, we first compute the inverse of the link weights, then the WSP are computed as the minimum sum of these inverse link weights necessary to travel among nodes [35]. In the case the link weights are costs (or distances), the WSP can be computed directly by summing the original link weights. The link weights in the PSI network are flows; for this reason, we use the first procedure to compute the WSP.

Then the weighted betweenness centrality $g^w(i)$ of the node i is defined:

$$g^w(i) = \sum_{s,t=1}^{N} \frac{\sigma_{st}^w(i)}{\sigma_{st}^w}, \tag{8}$$

where σ_{st}^w is the total number of WSP between nodes s and t and $\sigma_{st}^w(i)$ is the number of these WSP passing through the node i, summed up over the total number of nodes N. The higher the *BCw* of a node, the higher is the number of WSP passing through it. The *BCw* is the weighted counterpart of *BC*.

TR: Removal of nodes according to their transitivity (or clustering). The node transitivity measures the probability that its adjacent nodes (neighbours) are connected themselves. It is calculated as the proportion of links between the node neighbours divided by the total number of possible links [36]. Equivalently, we can compute the transitivity considering the 'triangles' in the network, that is, subgraphs of three nodes. In this case, it is calculated as the ratio between the closed triangles (complete subgraphs of three nodes) connected to the node and all the possible triangles centred on the node, as defined below:

$$T_i = \frac{\lambda_i}{\frac{1}{2}k_i(k_i - 1)}, \tag{9}$$

where λ_i is the number of closed triangles among neighbours of node i and $\frac{1}{2}k_i(k_i - 1)$ is the total possible number of triangles centred on node i. The node transitivity is also called 'local transitivity' or 'node clustering coefficient' [33]. In network theory, node transitivity is a measure of the magnitude to which nodes in a network tend to cluster together. The node transitivity defined here is a topological (binary) metric of nodes clustering, not including the link weights.

For all the node attack strategies, in the case of ties, that is, nodes with equal ranking, we randomly sort their sequence. We perform 10^3 simulations for each node attack strategy. The list of the node attack strategies with reference is in Table 1.

2.2 *The network functioning indicators*

weakLCC: The weakly connected largest connected component (*weakLCC*), also called 'giant cluster' or 'spanning cluster', represents the maximum number of connected nodes in the network [17–20] and can be written as:

$$weakLCC = \max(S_j), \tag{10}$$

where S_j is the size (number of nodes) of the jth cluster. The *weakLCC* is a simple and widely used indicator of network functioning (robustness). It is a non-directed and non-weighted indicator, that is, it

TABLE 1 *List of the node properties adopted for the corresponding node attack strategies*

Strategy	Type of node removal	Key	Refs	Unweighted/ weighted	Undirected/ directed
Random	Random node removal	*RAND*	[7, 22]	Unweighted	Undirected
Degree	The degree is the number of node links.	*DEG*	[17–22]	Unweighted	Undirected
In-degree	The in-degree is the number of node ingoing links.	*InDEG*	[33, 34]	Unweighted	Directed
Out-degree	The out-degree is the number of node outgoing links.	*OutDEG*	[33, 34]	Unweighted	Directed
Strength	The strength is the sum of the node link weights.	*STR*	[22, 25]	Weighted	Undirected
In-strength	The in-strength is the sum of the node ingoing link weights.	*InSTR*	[33]	Weighted	Directed
Out-strength	The out-strength is the sum of the node outgoing link weights.	*OutSTR*	[33]	Weighted	Directed
Betweenness	The node betweenness centrality is the number of binary shortest paths passing on it.	*BC*	[17–20]	Unweighted	Undirected
Weighted betweenness	The weighted betweenness centrality of the node is the number of weighted shortest paths passing on it.	*wBC*	[35]	Weighted	Undirected
Transitivity	The node transitivity is the ratio of the closed triangles connected to the node and all the possible triangles centred on that node.	*TR*	[33, 36]	Unweighted	Undirected

does not consider the network links directionalities and weights. The *weakLCC* of a network is the largest number of nodes connected by an undirected path.

strongLCC: The *strongLCC* is the directed counterpart of the *weakLCC,* as it represents the largest number of nodes in which every node can reach any other node by a directed path [33]. The *strongLCC* can be written:

$$strongLCC = \max(\kappa_j), \tag{11}$$

where κ_j is the subset of the maximal nodes, among all the possible j subsets, where a directed path connects every node.

EFF: The undirected network efficiency is a widely used indicator to quantify the network information spreading [37]. It is computed for undirected networks or neglecting the directionality of the links in the shortest paths (SP) between nodes. A network path is a sequence of links connecting two nodes in the network. While the binary SP between a pair of nodes is the minimum number of links connecting them, the weighted shortest paths (WSP) consider the link weights to account for the path length. The higher the weight of a link, the faster the information flows between the linked nodes. Conversely, the longer the WSP among nodes, the lower the efficiency of the network. As explained above, to calculate WSP, we first must consider whether the link weights in the network are 'flows or costs.' If the weights are flows, we first compute the inverse of the link weights, and then the WSP length connecting two nodes is the minimum sum of these inverse link weights necessary to travel between them [25, 26, 37]. If the weights are costs, the length of the WSP connecting two nodes is the minimum sum of the original link weights necessary to travel between them [25, 26, 37]. The network efficiency can properly evaluate both binary and weighted networks. Given that the PSI is a weighted network, and for the sake of comparison with the *LCC* measures that are binary-topological indicators, we used the weighted network efficiency (*EFF*), defined as:

$$EFF = \frac{1}{N \cdot (N-1)} \sum_{i \neq j \in G} \frac{1}{d_{i,j}}, \tag{12}$$

where $d_{i,j}$ is the weighted shortest path length between node i and node j.

dirEFF: dirEFF is a weighted and directed indicator and the directed counterpart of the EFF, quantifying the information spreading in the network. It is computed considering the directed WSP length in the network. [28].

The list of the network functioning indicators with meaning and references is in Table 2.

2.3 *The robustness (R) of the network*

To compare the effectiveness of the node removal strategies, we compute the network robustness (R), a single numerical value corresponding to the area underlying the curve of the system functioning indicator (*weakLCC, strongLCC, EFF* and *dirEFF*) as a function of the fraction of nodes removed. The faster is the decreasing of the network functioning indicator as a function of node removal, the lower is the network robustness R. The *weakLCC, strongLCC, EFF* and *dirEFF* indicators account for $R_{weakLCC}$,$R_{strongLCC}$,R_{EFF} and R_{dirEFF}. See Bellingeri *et al.* [22] for details about the R measure.

All the simulations are performed with the R package (4.0.2 version).

2.4 *The photosystem I (PSI) complex network*

We performed the node attack strategies over the real-world *P. sativum* Photosystem I complex network. The PSI network describes the FRET energy transfer occurring among chromophores toward the reaction centre (RC). The building of the PSI complex network was recently presented in Montepietra *et al.* [28], where the PSI is modelled as a directed and weighted network with a total number of nodes/chromophores $N = 192$. For details about the PSI network construction, see Montepietra *et al.* [28].

2.5 *The cut-off distance (CD)*

Building different sub-networks starting from the complete PSI network is a solid strategy capable of providing insights into the system's different functioning regimes [28, 39]. These sub-networks are obtained from the complete one where all possible links are present (as described above) by removing links whose

TABLE 2 *List of the network functioning (robustness) indicators*

Indicator	Formula	Meaning	Refs
weakLCC	$weakLCC = \max(S_j)$ where S_j is the size (number of nodes) of the jth cluster.	Maximum nodes number in which every node can reach any other node by a undirected path	[17–22]
strongLCC	$strongLCC = \max(\kappa_j)$ where κ_j is the maximal nodes subset in which every node can reach any other through a directed path.	Maximum nodes number in which every node can reach any other node by a directed path	[28, 33]
EFF	$EFF = \frac{2}{N\cdot(N-1)} \sum_{i \neq j \in G} \frac{1}{d_{i,j}}$ where $d_{i,j}$ is the undirected weighted shortest path length between node i and node j.	Information spreading capacity through undirected weighted paths	[25, 26, 37]
dirEFF	$dirEFF = \frac{1}{N\cdot(N-1)} \sum_{i \neq j \in G} \frac{1}{d_{\overrightarrow{i,j}}}$ where $d_{\overrightarrow{i,j}}$ is the directed weighted shortest path length between node i and node j.	Information spreading capacity through directed weighted paths	[28, 33, 37]

associated physical distance is higher than a threshold value called cut-off distance (CD). A longer distance means that the link will have a lower FRET efficiency value. Physically, this procedure corresponds to the progressive removal of links associated with less efficient FRET between chromophores, leaving only the links with the more probable energy transfers. We created networks using the following cut-off distance values: NO cut-off, 90 Å, 80 Å, 70 Å, 60 Å, 50 Å, 40 Å, 30 Å, 20 Å and 10 Å.

3. Results

Figure 1 shows the *weakLCC* functioning indicator trends under different node removal strategies for different CDs. Figures 2 and 3 show the network robustness R against the different attack strategies at different CDs for all the functionality indicators (*weakLCC and strongLCC,* EFF and dirEFF, respectively). Figure 4 represents the robustness R as a function of the CD.

The rank of effectiveness of the attack strategies considering the robustness R for each CD and each indicator is in Table 3. The list of the PSI network features for each CD value is in Table 4. Figures A1–A8 in Supplementary Appendix show the node centrality value distributions for each measure of centrality. Figures A9–A17 in Supplementary Appendix show the different scatterplots of the node centrality values. Figures A18–A21 in Supplementary Appendix show bar plots of the network robustness R against the different attack strategies at all CDs for all the functionality indicators.

weakLCC: BCw is the best strategy only when CD = NO; *BC* is the best strategy only when CD = 10 Å; *TR* is the worst strategy for most CDs; most of the strategies are almost equally performing for most CDs.

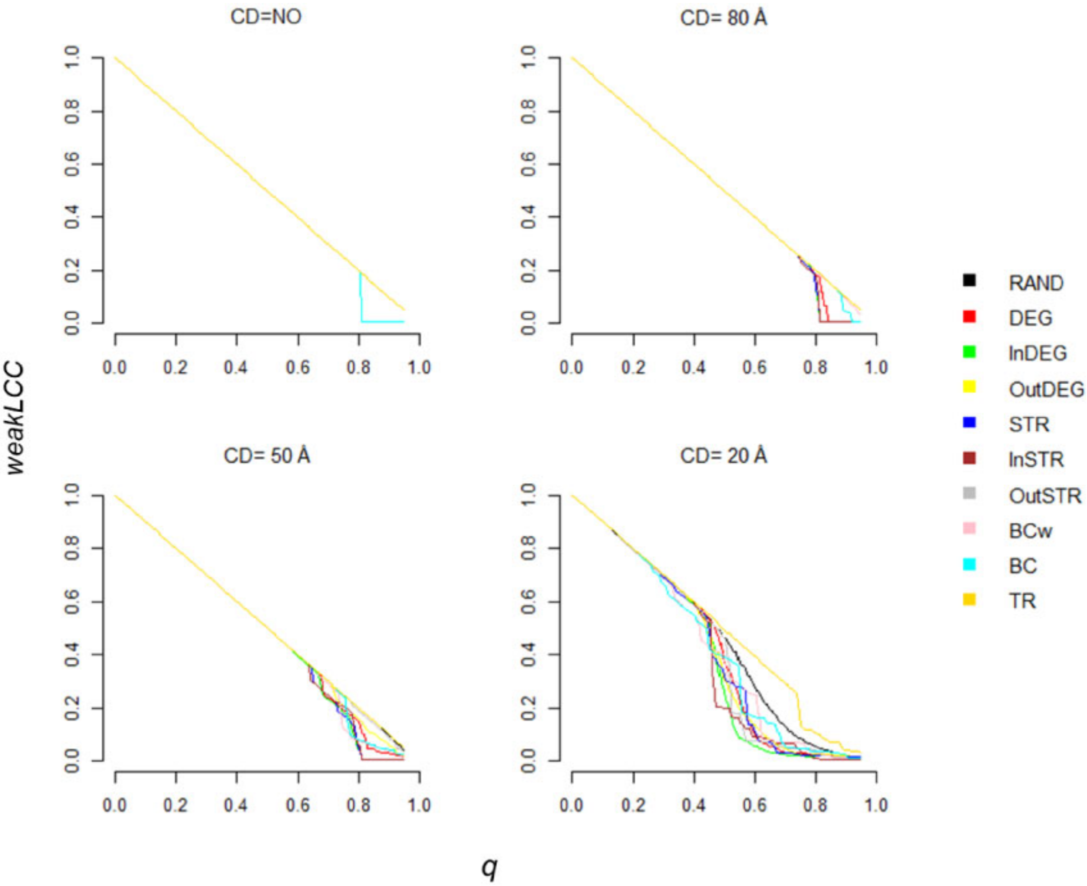

FIG. 1. Normalized *weakLCC* indicator as a function of the fraction of nodes removed q for different cut-off distance CD (Å). The *weakLCC* is normalized over the initial *weakLCC* value for that CD.

strongLCC: STR and *DEG* are the best strategies for almost all CD values; *TR* is the worst strategy for most CDs. Most of the strategies are almost equally performing for most CDs.

EFF: *TR* is the worst strategy for most CDs; *InDEG* is the best strategy for most CDs; *BC* is the best strategy for CD = 20 Å and CD = 30 Å.

dirEFF: *STR* is the best strategy for most CDs. Only for CD = 10 Å, *InDEG* is the best strategy; *TR* is the worst strategy for most CD.

Here below, we outline the most important outcomes among the many other results for each indicator.

TR for CD = 10 Å is comparable to *RAND* for every indicator. All strategies efficacy increase for every indicator when the CD value is lowered from NO cut-off to 10 Å. The different strategies generally decrease the R values of directed indicators (*strongLCC, dirEFF*) more than for the undirected indicators (*weakLCC, EFF*).

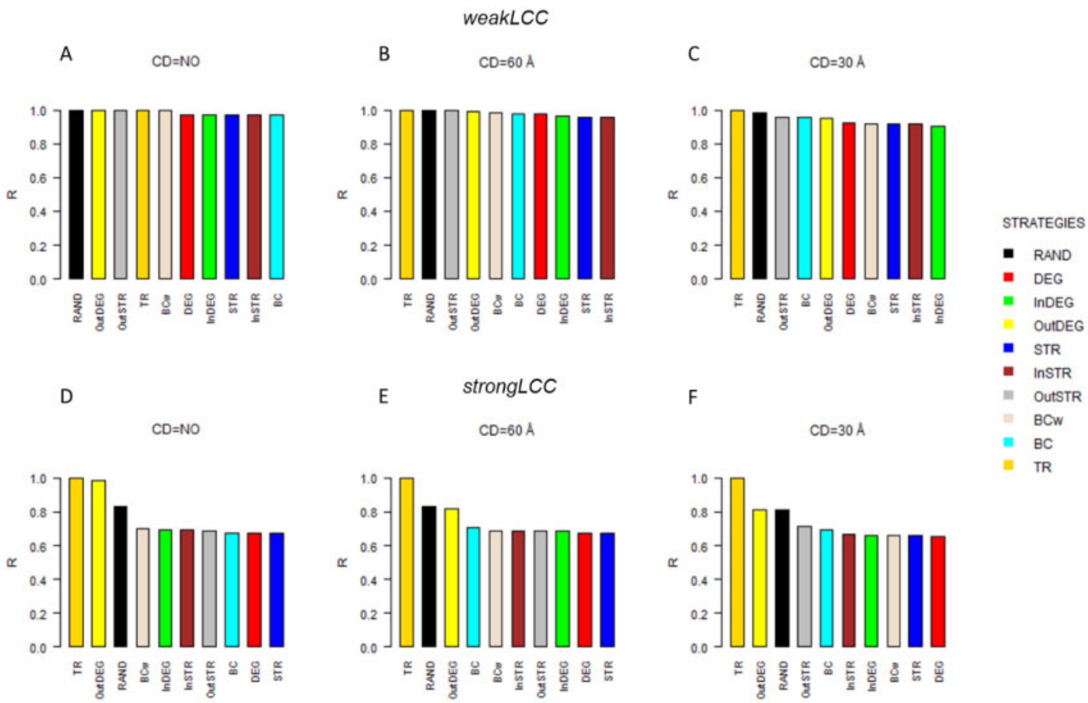

FIG. 2. Robustness measure R as a function of the cut-off distance CD (Å) for the *weakLCC* (A–C) and *strongLCC* (D–F) indicators. The robustness R is normalized over the maximum R-value for that CD.

In Fig. 4, we can observe that the robustness to the various attack strategies for both the *LCC* functioning measures remains high until CD = 30 Å, which acts as a threshold value. We see that for *EFF* and *dirEFF*, the R-value decreases continuously with the CD.

4. Discussions

4.1 *The best node attack strategies in real-world networks*

Finding the best node attack strategies is a fundamental problem of complex network theory [17–20]. A recent comprehensive comparison of node attack strategies by Wandelt *et al.* [18] showed that, on average, the old notion of node betweenness centrality (*BC*) is the best strategy to decrease the *weakLCC* in binary-topological real-world networks. Unexpectedly, we find that the PSI network is in the set of systems for which the *BC* is not the best strategy to decrease the *weakLCC* (only for CD = NO and CD = 10 Å the *BC* is among the best strategies) (Table 3). Understanding the consequences of node removal in real-world networks is a complex problem [18]. In the following, we provide some hypotheses to explain the low efficacy of the *BC* strategy in the PSI network. First, to decrease the *weakLCC* in our PSI network, many strategies perform in a very similar way, with a minimal difference among R-values (Fig. 2); for example, when CD = NO, the *BC*, *DEG*, *InDEG*, *STR* and *InSTR* strategies produced very similar R-values, and in the presence of such narrow difference, some hidden and secondary mechanisms may induce a specific strategy to prevail against the other.

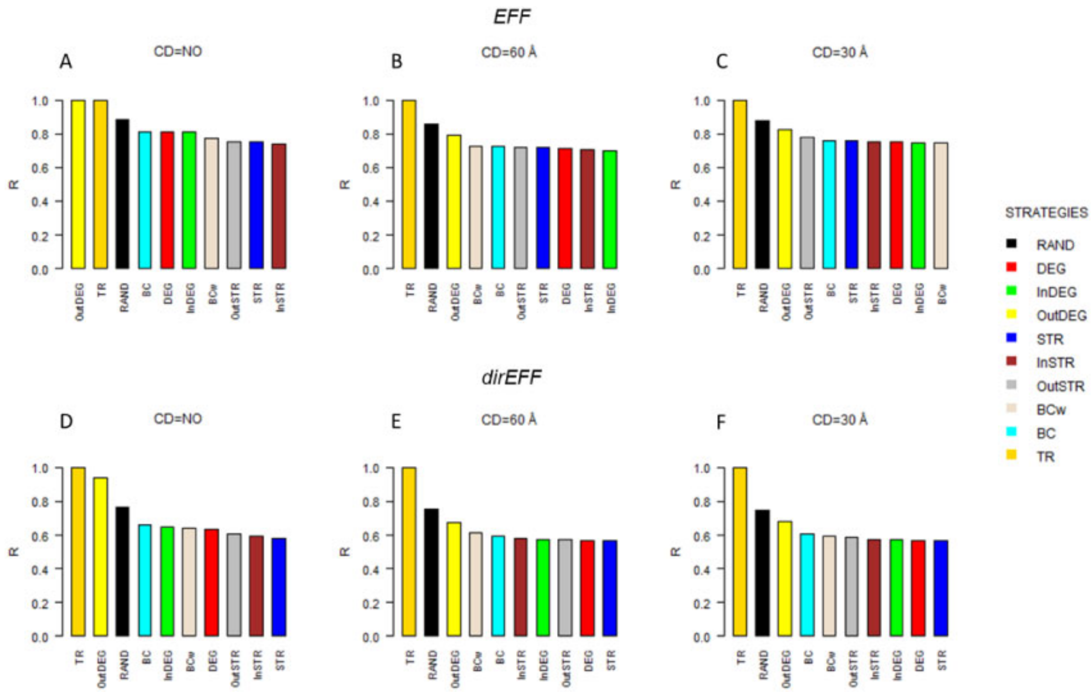

FIG. 3. Robustness measure R as a function of the cut-off distance CD (Å) for the *EFF* (A–C) and *dirEFF* indicators (D–F). The robustness R is normalized over the maximum R-value for that CD.

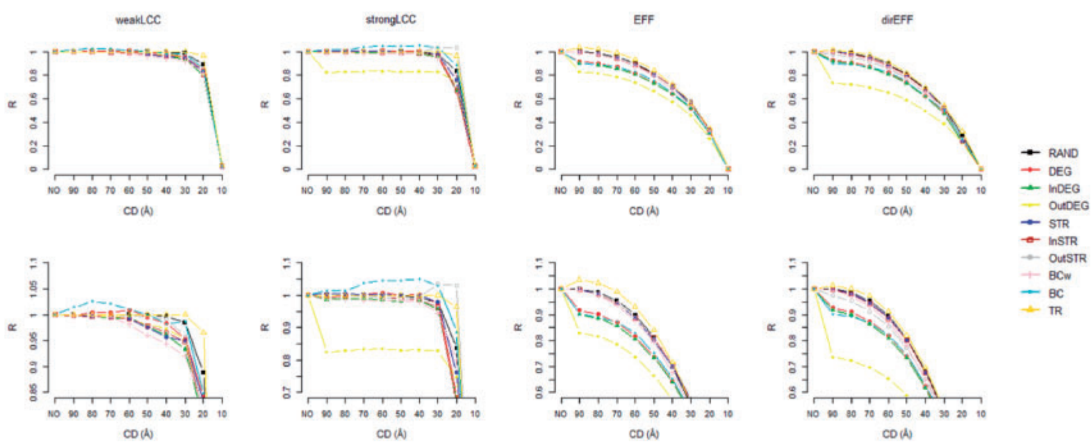

FIG. 4. Robustness measure R as a function of the cut-off distance CD (Å). The robustness R is normalized over the R-value computed for each strategy when tuning CD=NO. The bottom row plots depict each respective plot above in the top row with a reduced y-axis domain to outline the difference among curves.

TABLE 3 *Rank efficacy of the node attack strategies for each CD for each indicator. The rank is created by accounting the R-value*

weakLCC

RANK	CD NO	90 Å	80 Å	70 Å	60 Å	50 Å	40 Å	30 Å	20 Å	10 Å
1	*BC*	*InSTR*	*InSTR*	*InDEG*	*InSTR*	*STR*	*STR*	*InDEG*	*InDEG*	*BCw*
2	*InSTR*	*STR*	*STR*	*InSTR*	*STR*	*InDEG*	*InDEG*	*InSTR*	*InSTR*	*BC*
3	*STR*	*DEG*	*InDEG*	*STR*	*InDEG*	*InSTR*	*InSTR*	*STR*	*STR*	*InDEG*
4	*InDEG*	*InDEG*	*DEG*	*DEG*	*DEG*	*BCw*	*BCw*	*BCw*	*OutDEG*	*OutDEG*
5	*DEG*	*BC*	*BC*	*BC*	*BC*	*DEG*	*BC*	*DEG*	*OutSTR*	*OutSTR*
6	*BCw*	*BCw*	*BCw*	*BCw*	*BCw*	*BC*	*DEG*	*OutDEG*	*DEG*	*DEG*
7	*TR*	*TR*	*RAND*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*BC*	*BCw*	*InSTR*
8	*OutSTR*	*OutSTR*	*TR*	*RAND*	*OutSTR*	*OutSTR*	*OutSTR*	*OutSTR*	*BC*	*STR*
9	*OutDEG*	*OutDEG*	*OutSTR*	*TR*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*
10	*RAND*	*RAND*	*OutDEG*	*OutSTR*	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*

strongLCC

RANK	CD = NO	90 Å	80 Å	70 Å	60 Å	50 Å	40 Å	30 Å	20 Å	10 Å
1	*STR*	*DEG*	*DEG*	*STR*	*STR*	*STR*	*STR*	*DEG*	*InSTR*	*DEG*
2	*DEG*	*STR*	*STR*	*DEG*	*DEG*	*DEG*	*DEG*	*STR*	*DEG*	*STR*
3	*BC*	*InDEG*	*BC*	*InDEG*	*InDEG*	*InDEG*	*InDEG*	*BCw*	*InDEG*	*OutSTR*
4	*OutSTR*	*BC*	*InDEG*	*OutSTR*	*OutSTR*	*OutSTR*	*InSTR*	*InDEG*	*STR*	*InSTR*
5	*InSTR*	*OutSTR*	*OutSTR*	*BCw*	*InSTR*	*InSTR*	*BCw*	*InSTR*	*BCw*	*OutDEG*
6	*InDEG*	*BCw*	*BCw*	*InSTR*	*BCw*	*BCw*	*OutSTR*	*BC*	*BC*	*InDEG*
7	*BCw*	*InSTR*	*InSTR*	*BC*	*BC*	*BC*	*BC*	*OutSTR*	*RAND*	*BC*
8	*RAND*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*RAND*	*OutSTR*	*BCw*
9	*OutDEG*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*	*OutDEG*	*OutDEG*	*RAND*
10	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*

EFF

RANK	CD = NO	90 Å	80 Å	70 Å	60 Å	50 Å	40 Å	30 Å	20 Å	10 Å
1	*InSTR*	*OutSTR*	*InSTR*	*InSTR*	*BCw*	*BCw*	*BCw*	*BCw*	*BC*	*OutDEG*
2	*STR*	*STR*	*OutSTR*	*DEG*	*DEG*	*DEG*	*BC*	*BC*	*BCw*	*InDEG*
3	*OutSTR*	*InSTR*	*STR*	*InDEG*	*STR*	*InDEG*	*OutSTR*	*OutSTR*	*STR*	*DEG*
4	*BC*	*DEG*	*DEG*	*BCw*	*InSTR*	*InSTR*	*STR*	*InDEG*	*DEG*	*InSTR*
5	*InDEG*	*BCw*	*BCw*	*STR*	*InDEG*	*STR*	*InSTR*	*DEG*	*InSTR*	*BCw*
6	*DEG*	*InDEG*	*InDEG*	*OutSTR*	*OutSTR*	*BC*	*DEG*	*STR*	*OutSTR*	*BC*
7	*BCw*	*BC*	*BC*	*BC*	*BC*	*OutSTR*	*InDEG*	*InSTR*	*InDEG*	*STR*
8	*RAND*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*OutSTR*
9	*TR*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*
10	*OutDEG*	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*

dirEFF

RANK	CD = NO	90 Å	80 Å	70 Å	60 Å	50 Å	40 Å	30 Å	20 Å	10 Å
1	*STR*	*STR*	*STR*	*STR*	*STR*	*STR*	*DEG*	*STR*	*STR*	*InDEG*
2	*InSTR*	*DEG*	*DEG*	*DEG*	*DEG*	*DEG*	*STR*	*DEG*	*InSTR*	*DEG*
3	*OutSTR*	*BC*	*OutSTR*	*OutSTR*	*OutSTR*	*OutSTR*	*OutSTR*	*InDEG*	*DEG*	*OutDEG*
4	*DEG*	*OutSTR*	*InDEG*	*InDEG*	*InDEG*	*InDEG*	*InDEG*	*InSTR*	*InDEG*	*InSTR*
5	*BCw*	*InSTR*	*InSTR*	*InSTR*	*InSTR*	*InSTR*	*InSTR*	*OutSTR*	*BCw*	*STR*
6	*InDEG*	*InDEG*	*BC*	*BC*	*BC*	*BC*	*BC*	*BCw*	*BC*	*BCw*
7	*BC*	*BCw*	*BCw*	*BCw*	*BCw*	*BCw*	*BCw*	*BC*	*OutSTR*	*BC*
8	*RAND*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*OutDEG*	*OutSTR*
9	*OutDEG*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*	*RAND*	*TR*
10	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*	*TR*	*RAND*

TABLE 4. PSI network features for each CD value. *L* total number of links; *strongLCC* strongly largest connected component; *weakLCC* weakly largest connected component; *EFF* undirected network efficiency; *dirEFF* directed network efficiency; $<k^{in}>$ average nodes in-degree; $<k^{out}>$ average nodes out-degree; $<k>$ average nodes degree; $<S^{in}>$ average in-strength; $<S^{out}>$ average out-strength; $<S>$ average node strength; $<g(i)>$ average node betweenness centrality; $<g^{w}(i)>$ average node weighted betweenness centrality; $<T>$ average node transitivity

CD	L	*strongLCC*	*weakLCC*	*EFF*	*dirEFF*	$<k^{in}>$	$<k^{out}>$	$<k>$	$<S^{in}>$	$<S^{out}>$	$<S>$	$<g(i)>$	$<g^{w}(i)>$	$<T>$
10 Å	142	4	6	0.006	0.004	0.74	0.74	1	0.73	0.73	1.45	0.4	0.4	0
20 Å	1355	154	192	0.257	0.18	7.06	7.06	14.11	6.37	6.37	12.74	389.12	390.3	0.23
30 Å	3621	154	192	0.384	0.29	18.86	18.86	37.72	15.86	15.86	31.72	195	210.9	0.28
40 Å	6471	154	192	0.478	0.37	33.7	33.7	67.41	27.11	27.11	54.21	123	144.3	0.29
50 Å	9742	154	192	0.553	0.43	50.74	50.74	101.48	39.5	39.5	79	84.3	107.6	0.31
60 Å	13219	154	192	0.609	0.48	68.85	68.85	137.7	52.11	52.11	104.23	60.1	85.3	0.33
70 Å	16542	154	192	0.645	0.51	86.16	86.16	172.31	63.19	63.19	126.38	44.5	71	0.35
80 Å	19504	154	192	0.667	0.53	101.58	101.58	203.17	71.44	71.44	142.88	34.2	61.3	0.36
90 Å	22401	154	192	0.676	0.53	116.67	116.67	233.34	77.86	77.86	155.73	25.2	52.3	0.38
NO CD	29390	154	192	0.677	0.53	153.07	153.07	306.15	85.41	85.41	170.81	3.3	51.2	0.44

Secondly, the efficacy of the *BC* strategy to fragment the networks consists of removing 'bridge-nodes' [18, 40]. Bridge-nodes are nodes connecting different network communities, that is, sub-networks in which nodes are highly connected with other nodes in the same community and sparsely connected with nodes of different communities [18, 40]. The removal of bridge-nodes disconnects different communities, triggering a fast network dismantling and a quick *weakLCC* decrease [18]. We hypothesize that the relative efficacy of *BC* is higher when the bridge-nodes are of a low degree, that is, when the bridge-nodes share few links. In this case, the *BC* centrality and the degree-based nodes centralities (*DEG*, *InDEG*, *STR* and *InSTR*) are de-correlated. When the bridge-nodes are of low degree, they are not primarily removed by the degree-based strategies. Since the PSI network shows very high node connectivity (Table 4), and the *BC* node centrality is highly correlated with *STR* and *DEG* centrality (Figs. A9 and A10 in Supplementary Appendix), we argue that *DEG* and *STR* strategies are able to remove bridge-nodes as *BC* in the PSI network. In addition, these strategies remove nodes with higher connectivity levels than *BC*, thus producing higher (or at least similar) efficacy to *BC* for decreasing the *weakLCC*.

Further, we outline that for higher CDs (CD$>$70 Å), all node removal strategies are ineffective in decreasing the *weakLCC* (Figs. 2 and 4). Only for CD$<$70 Å, the intentional node attack strategies start to have a bit higher effectiveness to break up the *weakLCC*, that is, higher than the average represented by the random node removal (*RAND*) effectiveness. The low efficacy of the node attack strategies may be due to the very high connectivity of the PSI network for higher CDs, which triggers the 'degeneracy' of the node centrality properties, that is, most of the nodes share the same centrality values. This degeneracy of the node properties caused by the higher network connectivity level induces a random-like node sort even when nodes removed according to a specific node centrality measure (i.e. nodes of equal centrality value are randomly sorted), thus inducing a very low decrease of the *weakLCC*. This outcome would indicate that the node attack strategies based on classic node properties, such as the degree or the betweenness centrality, may show low efficacy in dismantling real-world networks with very high connectivity levels.

4.2 *Undirected vs. directed node attack strategies*

An important question in network science is understanding how the links' directionality changes the network structure [41, 42]. In node attack analyses, it refers to comprehending how the inclusion of link direction would affect the response of the network to node/link removal [43, 44]. The PSI network is genuinely directed and considering the proper direction of the FRET energy transfer link is fundamental to perform a more accurate description of the PSI system [28]. Interestingly, we find that the two best node attack strategies to decrease the directed indicators (i.e. *strongLCC* and *dirEFF*) are undirected strategies, that is, *STR* and *DEG* (Figs. 2 and 3). *strongLCC* counts the highest number of connected nodes throughout directed paths, while *dirEFF* considers the information exchange efficiency along the directed shortest paths of the PSI network. The higher efficacy of the *DEG* and *STR* strategies in decreasing these directed indicators of the network functioning may be explained by the heavy coupling between the *DEG* and*STR* measures of node centrality and their directed counterparts. In Fig. 5, we show the scatterplot of DEG vs. *InDEG/OutDEG* and *STR* vs. *InSTR/OutSTR*, and we can see how these undirected and directed measures of node centrality are strongly coupled in the PSI network. As a consequence, when removing nodes with a higher degree (*DEG*), we also remove nodes of higher in-degree (*InDEG*) and out-degree (*OutDEG*) (and analogously for *STR*), thus triggering a significant decrease in the directed indicators *strongLCC* and *dirEFF*.

Differently, we find that the *BC* strategy is not effective in reducing the *strongLCC* (Fig. 2). For this reason, our finding would denote that when considering link directionalities, the node attack strategies that are proven to be effective in undirected networks (i.e. to decrease the *weakLCC*), such as the well-known

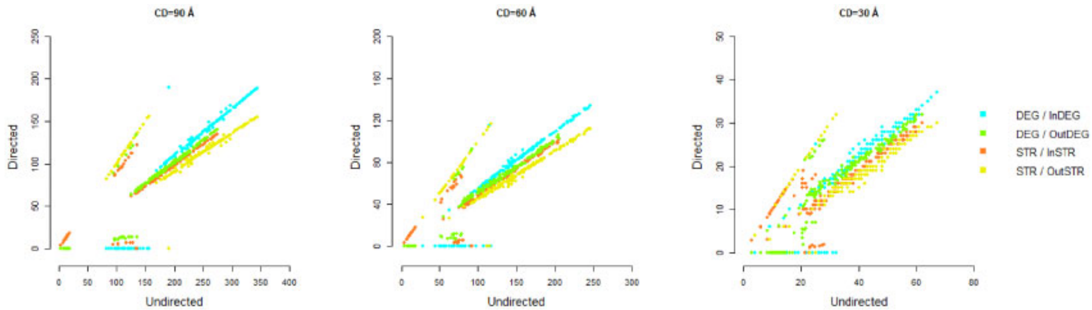

FIG. 5. Scatterplot of the undirected measures of node centrality values *vs*. their directed counterpart of node centrality values of the PSI network for three different CDs.

BC strategy [17–20], may be less effective. This outcome would suggest how the efficacy of the attack strategies may change accounting the direction of the links and outline the importance of considering link directionality to improve the modelling of the real-world systems.

4.3 *The low efficacy of the node transitivity attack*

The node transitivity attack strategy (*TR*) is clearly the worst attack strategy, performing even worse than the random removal strategy (*RAND*) (Table 3). The transitivity coefficient *TR* of a node, also known as node clustering coefficient, is the ratio of the number of triangles (closed loops of length three) over the total number of possible triangles centred on the node [33]. In other words, it is the frequency of triangles in the network, denoting the nodes' tendency to cluster together in a local community of nodes. The low efficacy of the *TR* node removal strategy can be explained by the negative coupling between the *TR* node centrality and the other node centrality measures that effectively harm the PSI network. In Figs. A12 and A13 in Supplementary Appendix, we report the scatterplot of the degree (*DEG*) and strength (*STR*) centrality against the transitivity (*TR*), and we observe how the *DEG* and *STR* are negatively correlated with the *TR* node centrality, that is, the most connected nodes are of lower transitivity and conversely. For these reasons, the *TR* node attack strategy removes nodes with low connectivity levels (both binary and weighted), resulting in shallow damage to the PSI network.

The TR node removal strategy's low efficacy provides some insights into the specific nature of higher transitivity nodes/chromophores. In Fig. 6, we can see that the higher transitivity nodes/chromophores correspond to carotenoid molecules (BCR, ZEX, XAT and LUT). Our findings indicate that the carotenoid molecules correspond to low-degree and low-strength nodes forming local network communities. Furthermore, in Table A2 in Supplementary Appendix, we see how the carotenoid nodes show very low betweenness centrality, both *BC* and *BCw*. The removal of higher transitivity nodes by the *TR* strategy would harm these peripheral communities of nodes that play a marginal contribution in routing the energy transfer within the PSI network, thus triggering a very slight decrease in all the network functioning indicators. This finding corroborates previous outcomes showing how carotenoids would not play an essential role as energy transfer hubs in the PSI network [28].

4.4 *The PSI network and the weight thresholding problem*

The 'weight thresholding' is a simple technique that aims to reduce the number of links in weighted networks that are otherwise too dense to apply standard network analysis methods [31, 32]. It consists of removing the links with weight below a given threshold. Ideally, the aim is to eliminate as many links as

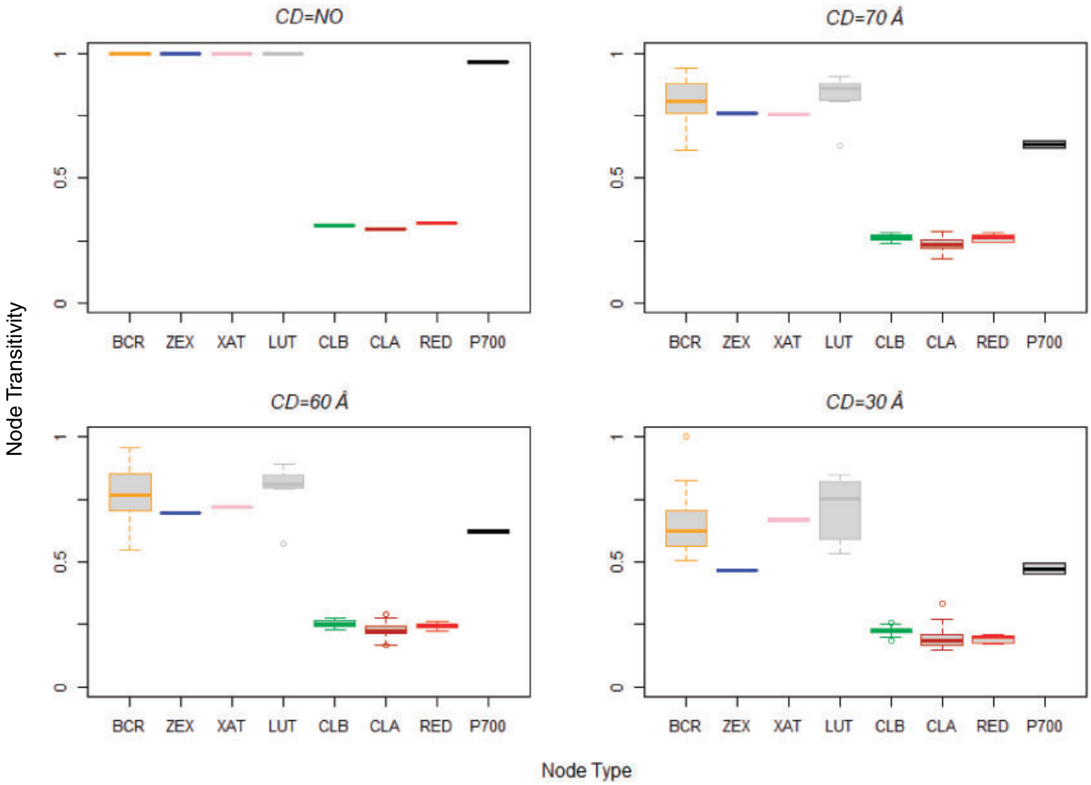

FIG. 6. Nodes transitivity (*TR*) value for each node type in the PSI network for four CDs. Node types keys are: β-carotene BCR, β-carotene derivated ZEX, violaxanthin XAT, lutein LUT, Chls b CLB, Chls a CLA, red form Chls RED, PSI reaction centre P700.

possible without drastically altering key features of the original networked system [32]. It has been shown that the brain network structural features change as a function of the threshold value, finding that many conventional network features are usually altered early on by the link deletion (pruning) procedure of the weight thresholding [45]. Investigating how network features change with weight thresholding in both model and real-world networks, Yan *et al.* [32] showed that local and global network features are often quickly lost when the network is subjected to weight thresholding, and the weight thresholding procedure does not alter the mesoscopic organization of the network, that is the groups (e.g. communities) survive even when most of the weaker links are removed.

Decreasing the cut-off distance (CD) over which links are not drawn, we progressively prune the lower-weight links (lower FRET energy transfer), acting a straightforward weight thresholding procedure over the PSI network. For these reasons, our analysis offers insights into the weight thresholding problem. First, this research shows how the PSI network response to node removal is affected by decreasing CD, that is, the best node removal strategy changes by decreasing CD (Table 3), indicating how the PSI network response to node removal may be unstable when subjected to weight thresholding. Second, we find how the distribution of classic binary-topological node centrality measures, such as the node degree (Fig. A1 in Supplementary Appendix) and the node betweenness (Fig. A6 in Supplementary Appendix), strongly change passing from CD=NO to CD=90 Å. This shows how the PSI network node properties are highly unstable in the earlier procedure of link deletion based on weight thresholding. On the other hand, the node centrality measures based on the weighted structure of the PSI networks, such as the strength and the weighted betweenness centrality of the nodes, show similar distributions shape when tuning different CD (Figs. A4 and A7 in Supplementary Appendix). This would suggest how node centralities measures considering link weights may be more stable against weight thresholding than the simple binary-topological ones.

5. Conclusion

In this article, we implemented ten node attack (removal) strategies over a recently assembled PSI complex network describing the energy transfer among nodes/chromophores of the common pea plant *P. sativum*. Unexpectedly, we discovered that the betweenness centrality node attack strategy, which is the most efficient for most real-world networks in decreasing the largest connected component, is, instead, of low efficacy for the PSI network. This outcome furnishes new insights into the field of node attack analysis, outlining how real-world networks may exhibit different and specific responses to node removals. Second, the node removal strategies based on binary-topological features of the network presented limited efficacy in damaging a highly connected network such as the PSI. Differently, the node attack strategies acting over the weighted structure of the network, thus discriminating the links according to their weight, would be more effective in harming highly connected real-world networks. This unexpected finding outlines how considering link weights may help overcome redundant information and select important nodes in highly connected networks, even when the goal is to dismantle the binary-topological structure of the network. Last, this research presents new perspectives and insights within the weight thresholding problem by demonstrating how the PSI network response to node removal and the efficacy of the node attack strategies change by progressively removing links of lower weight. This opens the question to test the robustness of the networks and the efficacy of the node attack strategies when real-world networked systems are subjected to weight thresholding.

Supplementary data

Supplementary data are available at *COMNET* online.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 816313). Many thanks to Dr. Aaron Michael Ross for the revision of the English language in the final version of the manuscript.

References

1. Otte, E. & Rousseau, R. (2002) Social network analysis: a powerful strategy, also for the information sciences. *J. Inf. Sci*. **28**, 441–453.
2. Zhu, B., Yeung, C. H., & Liem, R. P. (2021) The impact of common neighbor algorithm on individual friend choices and online social networks. *Phys. A Stat. Mech. Appl*. **566**, 125670.
3. Bellingeri, M., Bevacqua, D., Scotognella, F., Alfieri, R., Nguyen, Q., Montepietra, D., & Cassi, D. (2020) Link and node removal in real social networks: a review. *Front. Phys.* **8**, 228.
4. Smolyak, A., Levy, O., Shekhtman, L., Havlin, S. (2018) Interdependent networks in Economics and Finance—a physics approach. *Phys. A Stat. Mech. Appl*. **512**, 612–619.
5. Lordan, O., Sallan, J. M., Simo, P., & Gonzalez-Prieto, D. (2014) Robustness of the air transport network. *Transp. Res. Part E* **68**, 155–163.

6. Bellingeri, M., Lu, Z. M., Cassi, D., & Scotognella, F. (2018) Analyses of the response of a complex weighted network to nodes removal strategies considering links weight: the case of the Beijing urban road system. *Mod. Phys. Lett. B.*
7. Bellingeri, M., Bevacqua, D., Scotognella, F., Lu, Z. M., & Cassi, D. (2018) Efficacy of local attack strategies on the Beijing road complex weighted network. *Phys. A Stat. Mech. Appl.* **510**, 316–328.
8. Bellingeri, M., & Bodini, A. (2016) Food web's backbones and energy delivery in ecosystems. *Oikos* **125**, 586–594.
9. Bellingeri, M., Cassi, D., & Vincenzi, S. (2013) Increasing the extinction risk of highly connected species causes a sharp robust-to-fragile transition in empirical food webs. *Ecol. Modell.* **251**, 1–8.
10. Bellingeri, M., & Vincenzi, S. (2013) Robustness of empirical food webs with varying consumer's sensitivities to loss of resources. *J. Theor. Biol.*, **333**, 18–26.
11. Stella, M., Beckage, N. M., & Brede, M. (2017) Multiplex lexical networks reveal patterns in early word acquisition in children. *Sci. Rep.* **7**, 1–10.
12. Fiori, K. L., Smith, J., & Antonucci, T. C. (2007) Social network types among older adults: a multidimensional approach. *J. Gerontol. - Ser. B Psychol. Sci. Soc. Sci*. **62**, 322–330.
13. De Domenico, M., Sasai, S., & Arenas A. (2016) Mapping multiplex hubs in human functional brain networks. *Front. Neurosci.* **10**, 1–14.
14. Cuadra, L., Salcedo-Sanz, S., Del Ser, J., Jiménez-Fernández, S., & Geem, Z. W. (2015) A critical review of robustness in power grids using complex networks concepts. *Energies* **8**, 9211–9265.
15. Martinazzi, S., & Flori A. (2020) The evolving topology of the lightning network: centralization, efficiency, robustness, synchronization, and anonymity. *PLoS One* **15**, 1–18.
16. Garas, A., Argyrakis, P., & Havlin, S. (2008) The structural role of weak and strong links in a financial market network. *Eur. Phys. J. B* **63**, 265–271.
17. Iyer, S., Killingback, T., Sundaram, B., & Wang, Z. (2013) Attack robustness and centrality of complex networks. *PLoS One* **8**, e59613.
18. Wandelt, S., Sun, X., Feng, D., Zanin, M., & Havlin, S. (2018) A comparative analysis of approaches to network-dismantling. *Sci. Rep.* **8**, 1–15.
19. Nguyen, Q., Pham, H.D., Cassi, D., & Bellingeri, M. (2019) Conditional attack strategy for real-world complex networks. *Phys. A Stat. Mech. its Appl*. **530**, 121561.
20. Bellingeri, M., Cassi, D., & Vincenzi, S. (2014) Efficiency of attack strategies on complex model and real-world networks. *Phys. A Stat. Mech. Appl*. **414**, 174–180.
21. Lekha, D. S., & Balakrishnan, K. (2020) Central attacks in complex networks: a revisit with new fallback strategy. *Phys. A Stat. Mech. Appl*. **549**, 124347.
22. Bellingeri, M., Bevacqua, D., Scotognella, F., & Cassi, D. (2019) The heterogeneity in link weights may decrease the robustness of real-world complex weighted network. *Sci. Rep.*, **9**, 10692.
23. Hadidjojo, J., & Cheong, S.A. (2011) Equal graph partitioning on estimated infection network as an effective epidemic mitigation measure. *PLoS One* **6**, e22124.
24. Wang, Z., Zhao, D.W., Wang, L., Sun, G.Q., & Jin, Z. (2015) Immunity of multiplex networks via acquaintance vaccination. *EPL* **112**, 48002.
25. Bellingeri, M., & Cassi, D. (2018) Robustness of weighted networks. *Phys. A Stat. Mech. Appl*. **489**, 47–55.
26. Bellingeri, M., Bevacqua, D., Scotognella, F., Alfieri, R., & Cassi D. (2020) A comparative analysis of link removal strategies in real complex weighted networks. *Sci. Rep*. **10**, 1–15.
27. Dall'Asta, L., Barrat, A., Barthélemy, M., & Vespignani, A. (2006) Vulnerability of weighted networks. *J. Stat. Mech. Theory Exp.* **4**, 04006.
28. Montepietra, D., Bellingeri, M., Ross, A. M., Scotognella, F., & Cassi, D. (2020) Modelling photosystem i as a complex interacting network: modelling the photosynthetic system I as complex interacting network. *J. R. Soc. Interface* **17**, 295.
29. Fromme, P., Jordan, P., & Krauß, N. (2001) Structure of photosystem I. *Biochim. Biophys. Acta - Bioenerg*. **1507**, 5–31.

30. Golbeck, J.H. (1987) Structure, function and organization of the photosystem I reaction center complex. *BBA Rev. Bioenerg*. **895**, 167–204.

31. Tumminello, M., Aste, T., Di Matteo, T., & Mantegna, R.N. (2005) A tool for filtering information in complex systems. *Proc. Natl. Acad. Sci. USA* **102**, 10421–10426.

32. Yan, X., Jeub, L.G.S., Flammini, A., Radicchi, F., & Fortunato, S. (2018) Weight thresholding on complex networks. *Phys. Rev. E* **98**, 1–9.

33. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. (2006) Complex networks: structure and dynamics. *Phys. Rep*. **424**, 175–308.

34. Da Mata AS. 2020 Complex networks: a mini-review. *Brazilian J. Phys*. **50**, 658–672.

35. Barrat, A., Barthélemy, M., Pastor-Satorras, R., & Vespignani, A. 2004 The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. USA* **101**, 3747–3752.

36. Watts, D.J., & Strogatz, S.H. (1998) Collective dynamics of 'small-world' networks. *Nature* **393**, 440–2.

37. Latora, V., & Marchiori, M. (2001) Efficient behavior of small-world networks. *Phys. Rev. Lett.* **87**, 198701.

38. Mazor, Y., Borovikova, A., Caspy, I., & Nelson, N. (2017) Structure of the plant photosystem i supercomplex at 2.6 Å resolution. *Nat. Plants* **3**, 1–9.

39. Croce, R., & Van Amerongen, H. (2014) Natural strategies for photosynthetic light harvesting. *Nat. Chem. Biol*. **10**, 492–501.

40. Nguyen, Q., Vu, T., Dinh, H.-D., Cassi, D., Scotognella, F., Alfieri, R., & Bellingeri, M. (2021) Modularity affects the robustness of scale-free model and real-world social networks under betweenness and degree-based node attack. *Appl. Network Sci.* **6**, 1–21.

41. Bütün, E. & Kaya, M. 2019 A pattern based supervised link prediction in directed complex networks. *Phys. A Stat. Mech. Appl*. **525**, 1136–1145.

42. Squartini, T., Picciolo, F., Ruzzenenti, F., & Garlaschelli, D. (2013) Reciprocity of weighted networks. *Sci. Rep.* **3**, 2729.

43. Kashyap, G., & Ambika, G. (2019) Link deletion in directed complex networks. *Phys. A Stat. Mech. Appl*. **514**, 631–643.

44. Yu, Y., Deng, Y., Tan, S.Y., & Wu, J. (2018) Efficient disintegration strategy in directed networks based on tabu search. *Phys. A Stat. Mech. Appl*. **507**, 435–442.

45. Garrison, K.A., Scheinost, D., Finn, E.S., Shen, X., Constable, R.T. (2015) The (in)stability of functional brain network measures across thresholds. *Neuroimage* **118**, 651–661.