J. Phys. A: Math. Theor. **43** (2010) 395001 (28pp) doi:10.1088/1751-8113/43/39/395001

Network resilience against intelligent attacks constrained by the degree-dependent node removal cost

A Annibale[1], **A C C Coolen**[1,2] **and G Bianconi**[3]

[1] Department of Mathematics, King's College London, The Strand, London WC2R 2LS, UK
[2] Randall Division of Cell and Molecular Biophysics, King's College London, New Hunt's House, London SE1 1UL, UK
[3] Northeastern University, Boston, MA, USA

E-mail: alessia.annibale@kcl.ac.uk, ton.coolen@kcl.ac.uk and g.bianconi@neu.edu

Received 25 May 2010, in final form 27 July 2010
Published 23 August 2010
Online at stacks.iop.org/JPhysA/43/395001

Abstract

We study the resilience of complex networks against attacks in which nodes are targeted intelligently, but where disabling a node has a cost to the attacker which depends on its degree. Attackers have to meet these costs with limited resources, which constrains their actions. A network's integrity is quantified in terms of the efficacy of the process that it supports. We calculate how the optimal attack strategy and the most attack-resistant network degree statistics depend on the node removal cost function and the attack resources. The resilience of networks against intelligent attacks is found to depend strongly on the node removal cost function faced by the attacker. In particular, if node removal costs increase sufficiently fast with the node degree, power law networks are found to be more resilient than Poissonian ones, even against optimized intelligent attacks. For cost functions increasing quadratically in the node degrees, intelligent attackers cannot damage the network more than random damages would.

PACS numbers: 75.10.Nr, 05.20.−y, 64.60.Cn

1. Introduction

In recent years there have been several studies into the resilience of complex networks against random failures and targeted attacks, in which a fraction of the nodes or of the bonds is removed. It was found that scale-free networks (with degree distributions that decay slowly via power laws, as in preferential attachment models) are more robust against random node removal than Poissonnian (or Erdös–Rényi) graphs which may explain why many real-world complex systems involve networks with power-law distributed degrees. However, scale-free networks were found to be very vulnerable to intelligent attackers that target high-degree nodes [1, 2]. Against edge removal, Poissonian and power law networks turned out to produce similar responses [3]. There are two reasons why we aim to study network resilience further. First, while the motivation behind such studies is that networks provide the infrastructure for some process (with interacting 'agents' or processors occupying the nodes), and that process disruption is the true goal of an attacker, most authors measure the impact of attacks indirectly, via topological properties that serve as proxies for the integrity of the process (e.g. the overall connectivity and path-length statistics [4], or percolation characteristics [1, 2, 5–7]). Here we seek to quantify the damage inflicted by attacks directly in terms of the process which the network is meant to support, similar to [8]. This requires solving stochastic processes on complex networks with arbitrary degree distributions, which is what statistical mechanics enable us to do. Our second and most important reason is that network resilience has so far been studied strictly in the context of random or intelligent removal of a fixed *fraction* of sites or bonds. This seems unrealistic. In most real-world scenarios (attacks on computer networks, viruses attacking cellular networks, etc) attacking a highly connected node demands more effort on behalf of the attacker than removing a weakly connected one. Similarly, any sensible defender of a network would devote more resources to the protection of 'hubs' than to the protection of 'outpost' nodes. The study of network resilience against attack or dilution calls for more appropriate and realistic definitions that include the inevitable resource constraints faced by attackers and defenders alike.

Turning to a formulation where attackers have *finite resources*, to be deployed intelligently when the cost of removing a network node depends on the degree of that node, changes the game drastically. It introduces a trade-off between the merit in terms of inflicted damage of targeting high-degree nodes versus the disadvantage of associated cost (attacking many 'hubs' may be unaffordable). One would like to know the maximum amount of damage that can be inflicted (by e.g. a virus to a biological network), given the limited resources available to the attacker (e.g. food, lifetime) and given the network's degree-dependent node removal costs. Similarly one would like to identify the most resilient network degree statistics to withstand an optimal attack. The answers to these questions may not only aid our understanding of structural properties of biological (e.g. proteomic) signalling networks, where competition and natural selection act as driving forces towards attack resistance, but also the design of attack-resistant synthetic real-world (e.g. communication) networks.

Here we develop a framework for the study of network resilience that includes limited attack resources, degree-dependent node removal costs and resilience measures based on process integrity. We consider two types of processes where structurally different interacting variables are placed on the nodes of networks with arbitrary degree distributions: interacting Ising spins (where the global order is ferromagnetic or of the spin-glass type), and coupled Kuramoto oscillators (where the global order is measured by synchronization). Both are solvable using finite connectivity replica theory, which enables us to quantify their integrity by the critical temperature of the ordered state. An attacker with finite resources seeks to destabilize these processes by removing or disrupting selected network nodes using his knowledge of the network's degrees. The attacker is also allowed to disable nodes partially (with a proportional reduction in attack costs). We identify the most damaging attack strategy, given a network's degree distribution and given the degree dependence of the node removal costs and the attack resources available. We then determine the optimal network topology from the point of view of the defender, i.e. that degree distribution for which the integrity of the process is preserved best when attacked by a foe who employs the most damaging attack strategy. The optimal attack strategy and the optimally attack-resistant network topology are found to be universal across the types of microscopic variables and types of global order considered. As expected, the resilience of network processes against intelligent attacks depends strongly on the node removal cost function faced by the attacker. Moreover, in sharp contrast to the traditional setup where attackers are allowed to remove a fixed *fraction* of the nodes (and hence can simply target the 'hubs'), we find that if node removal costs increase sufficiently fast with the node degree, and if attackers have finite resources to meet these costs, power law networks are more resilient than Poissonian ones, even against optimized intelligent attacks.

2. Definitions

2.1. Processes, supporting networks and constrained attack variables

We study two systems in which interacting stochastic variables are placed on the N nodes of a complex network. The network is defined via variables $c_{ij} \in \{0, 1\}$, with $c_{ij} = 1$ if and only if the nodes i and j are connected. We define $c_{ij} = c_{ji}$ and $c_{ii} = 0$ for all (i, j) and abbreviate $\mathbf{c} = \{c_{ij}\}$. The first system (A) consists of N Ising spins $\sigma_i \in \{-1, 1\}$, in thermal equilibrium, characterized by the following Hamiltonian:

$$\text{A:}\quad H(\boldsymbol{\sigma}) = -\sum_{i<j} c_{ij} J_{ij} \xi_i \xi_j \sigma_i \sigma_j \tag{1}$$

with $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)$. The second system (B) consists of N Kuramoto oscillators, with phases $\theta_i \in [-\pi, \pi]$, again in equilibrium but now with the Hamiltonian:

$$\text{B:}\quad H(\boldsymbol{\theta}) = -\sum_{i<j} c_{ij} J_{ij} \xi_i \xi_j \cos(\theta_i - \theta_j) \tag{2}$$

with $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$. The bonds $J_{ij} \in \mathbb{R}$ are drawn randomly and independently from a distribution $P(J)$. The variables $\xi_i \in \Xi \subseteq [0, 1]$ in (1), (2) represent the impact of attacks, with $\xi_i = 0$ if node i is removed completely and $\xi_i = 1$ if it is left alone. We demand that $1 \in \Xi$, so leaving a node intact is always an option, and for simplicity we take Ξ to be discrete and finite. We define the node degrees $k_i(\mathbf{c}) = \sum_j c_{ij}$, so the degree distribution and the average connectivity of $\mathbf{c}$ are $p(k|\mathbf{c}) = N^{-1} \sum_i \delta_{k,k_i(\mathbf{c})}$ and $\langle k \rangle = \sum_{k \geqslant 0} k p(k|\mathbf{c})$, respectively.

We assume that the cost to the attacker of setting $\xi_i = \xi$ at a node of degree $k_i = k$ is $\psi(\xi, k) \geqslant 0$, where $\psi(1, k) = 0$ and $\partial\psi(\xi, k)/\partial\xi \leqslant 0$ for all $\xi \in \Xi$. If attackers have limited resources they can only disrupt a subset of the nodes, since the $\{\xi_i\}$ will now be subject to a constraint of the form $\sum_i \psi(\xi_i, k_i) \leqslant C$. A natural choice for ψ is

$$\psi(\xi, k) = \kappa(1 - \xi)\phi(k), \tag{3}$$

where $\phi(k)$ is a non-decreasing function, with $\phi(0) = 0$ and $\phi(k > 0) > 0$. The attack cost for a node increases with the number of links to/from it; disconnected nodes can be attacked for free. The normalization factor κ is chosen such that the resource constraint takes the simple form $N^{-1} \sum_i \psi(\xi_i, k_i) \leqslant 1$. The attacker is assumed to act intelligently, using knowledge of the network's degrees, so the degrees $\{k_i\}$ and the attack variables $\{\xi_i\}$ will generally be correlated. Finally we draw the network $\mathbf{c}$ randomly from a maximum-entropy ensemble defined by a probability distribution in which the degrees are constrained to take prescribed values $\mathbf{k} = (k_1, \ldots, k_N)$:

$$\text{Prob}(\mathbf{c}) = Z^{-1}[\mathbf{k}] \prod_i \delta_{k_i, k_i(\mathbf{c})}, \qquad Z[\mathbf{k}] = \sum_{\mathbf{c}} \prod_i \delta_{k_i, k_i(\mathbf{c})}. \tag{4}$$

We abbreviate $p(\xi, k) = N^{-1} \sum_i \delta_{\xi_i, \xi} \delta_{k_i, k}$ and define $q(\xi|k)$ via $p(\xi, k) = q(\xi|k)p(k)$. The resource constraint on the attack variables then translates into $\sum_{\xi k} \psi(\xi, k) q(\xi|k) p(k) \leqslant 1$.

The attacker is assumed to know the degree sequence $\mathbf{k}$ of the network to be attacked and can adapt accordingly the conditional likelihood $q(\xi|k)$ to maximize the impact of his actions; $q(\xi|k)$ constitutes his attack strategy. The realistic regime is that where $\psi(\xi,k)$ obeys $\sum_k \psi(0,k)p(k) > 1$, so that the trivial 'destroy-all' attack strategy $q(\xi|k) = \delta_{\xi,0}\,\forall k$ is not feasible (i.e. too costly).

2.2. Quantifying process integrity and optimal attack and defence strategies

With each process (1), (2) running on the network, each associated ordered phase (ferromagnetic, spin-glass or synchronized) corresponds a critical temperature T_c, which will for large N depend on the network and attack characteristics $\mathbf{k}$ and $\boldsymbol{\xi}$ only via $q(\xi|k)$ and $p(k)$. The larger the T_c, the more robust is the ordered phase against local noise, so we can quantify the integrity of the process by the value of $T_c[p,q]$. The attacker wants to destroy the ordered phase of the process, whereas the defender seeks to protect it. This allows us to give precise definitions for the optimal attack strategy and the optimally resistant degree distribution in terms of process integrity. The optimal attack strategy $q^\star[p]$ is the conditional distribution $q(\xi|k)$ for which $T_c[p,q]$ is minimal, given the degree distribution p and given the resource constraint:

$$q^\star[p] = \text{argmin}_{\{q,\sum_{\xi k}\psi(\xi,k)q(\xi|k)p(k)\leqslant 1\}} T_c[p,q]. \tag{5}$$

The optimal (most resistant) degree distribution $p^\star$ to be chosen by the defender, given the average connectivity c (finite network resources) and attack cost function ψ is then that $p(k)$ which subsequently maximizes this q-minimized critical temperature:

$$\begin{aligned} p^\star &= \text{argmax}_{\{p,\sum_k kp(k)=c\}} T_c[p,q^\star[p]] \\ &= \text{argmax}_{\{p,\sum_k kp(k)=c\}}\, \text{argmin}_{\{q,\sum_{\xi k}\psi(\xi,k)q(\xi|k)p(k)\leqslant 1\}} T_c[p,q]. \end{aligned} \tag{6}$$

The end result is a situation where the defender, by choosing an appropriate degree distribution, maintains the highest achievable critical temperature $T_c[p^\star,q^\star[p^\star]]$, given he is subjected to the most damaging attack. However, within this scenario one could in fact ask many more interesting questions, such as what would be the effect of misinformation, a situation where a defender optimizes the network on the basis of an anticipated attack $q^\star[p]$ (so he chooses degree distribution $p^\star$) but is then faced with an attack with strategy $q' \neq q^\star[p]$, so that the actual critical temperature is $T_c[p^\star,q']$.

We see that the problem of identifying the optimal attack and defence strategies (5), (6) splits automatically into two distinct parts. The first part is calculating the critical temperature(s) $T_c[p,q]$ of the relevant phases. This is done by evaluating for the systems (1), (2) the asymptotic disorder-averaged free energy per spin $\overline{f}$, from which one extracts the phase diagrams for systems on typical graphs from (4):

$$\overline{f}_A = -\lim_{N\to\infty}\frac{1}{\beta N}\overline{\log\sum_{\boldsymbol{\sigma}} e^{-\beta H(\boldsymbol{\sigma})}} \tag{7}$$

$$\overline{f}_B = -\lim_{N\to\infty}\frac{1}{\beta N}\overline{\log\int_{-\pi}^{\pi} d\boldsymbol{\theta}\, e^{-\beta H(\boldsymbol{\theta})}} \tag{8}$$

in which $\beta = T^{-1}$ (where T denotes the temperature) and where $\overline{\cdots}$ denotes averaging over the disorder in the problem, namely the randomly drawn graphs with statistics (4) and the random bonds $\{J_{ij}\}$. The calculation of (7) and (8) is done with the finite connectivity replica method, based on the identity $\overline{\log Z} = \lim_{n\to 0} n^{-1}\log\overline{Z^n}$, and details are relegated to appendix A and appendix B in order not to disrupt the flow of the paper. The second part of the problem, to be tackled once the formulae for $T_c[p,q]$ have been derived (which, expectedly and fortunately, turn out to be simple and very similar across models and ordered phases), is to carry out the constrained optimizations in (5), (6), by a combination of analytical and numerical techniques.

3. The process integrity measure

We show in the appendices of this paper that the critical temperatures $T_c[p,q]$ for the emergence of global ferromagnetic (F) or spin-glass (SG) order, given we choose the bond distribution $P(J) = \frac{1}{2}(1+\eta)\delta(J-J_0) + \frac{1}{2}(1-\eta)\delta(J+J_0)$ (with $J_0 \geqslant 0$), follow for both Ising spins and coupled oscillators from formulae of the following form:

$$\text{F:}\quad \lambda^{(1)}_{\max}(\beta)=1,\qquad \lambda^{(1)}:\quad \text{eigenvalues of } M^{(1)}_{\xi\xi'}(\beta)=\eta K(\beta J_0\xi\xi')\gamma(\xi') \tag{9}$$

$$\text{SG:}\quad \lambda^{(2)}_{\max}(\beta)=1,\qquad \lambda^{(2)}:\quad \text{eigenvalues of } M^{(2)}_{\xi\xi'}(\beta)=K^2(\beta J_0\xi\xi')\gamma(\xi') \tag{10}$$

in which $\beta = 1/T$ and

$$\gamma(\xi) = \langle k\rangle^{-1}\sum_k q(\xi|k)p(k)k(k-1). \tag{11}$$

Here $K(z) = \tanh(z)$ for interacting Ising spins and $K(z) = I_1(z)/I_0(z)$ for coupled oscillators. In both cases $K(-z) = -K(z)$, $\frac{\mathrm{d}}{\mathrm{d}z}K(z) \geqslant 0$ and $\lim_{z\to\infty}K(z) = 1$. There is no F phase if $\eta \leqslant 0$, so we take $\eta > 0$ from now on. The structure of the above formulae is in agreement with results from percolation theory and spreading phenomena, which show that the threshold characterizing the percolation transition or an epidemic outbreak in a network depends on the ratio $\langle k^2\rangle/\langle k\rangle$ of the first two moments of its degree distribution [1, 2, 5–7, 9–11]. The approach followed here is closer to the envisaged picture of interacting agents or processors on network nodes and has the benefit of applying to the whole interval $\Xi = [0,1]$, as opposed to $\Xi = \{0,1\}$ which can be accessed by the percolation theory.

3.1. Tests and bounds for critical temperatures

Before any attack one has $\xi \in \{1\}$, so $\gamma(\xi) = \gamma(1) = \langle k^2\rangle/\langle k\rangle - 1$ and the above formulae would have reproduced the known results for the unperturbed system, namely

$$\text{F:}\quad \eta K(\beta J_0)[\langle k^2\rangle/\langle k\rangle - 1] = 1 \tag{12}$$

$$\text{SG:}\quad K^2(\beta J_0)[\langle k^2\rangle/\langle k\rangle - 1] = 1. \tag{13}$$

Another simple test is to consider $\Xi = \{0,1\}$. Here each node is either unaffected or removed completely, leaving a new network identical to an unperturbed network as described by (12), (13), but with reduced size $N' = \sum_i \xi_i$, and with degrees $k_i' = \sum_j c_{ij}\xi_j$. We would find, in the case of random attacks $q(\xi|k) = \zeta\delta_{\xi,0} + (1-\zeta)\delta_{\xi,1}$:

$$\langle k\rangle' = \lim_{N\to\infty}\frac{1}{(1-\zeta)N}\sum_{ij}\xi_i c_{ij}\xi_j = (1-\zeta)\langle k\rangle \tag{14}$$

$$\langle k^2\rangle' = \lim_{N\to\infty}\frac{1}{(1-\zeta)N}\sum_{ij\ell}\xi_i c_{ij}c_{i\ell}\xi_j\xi_\ell = (1-\zeta)^2\langle k^2\rangle + \zeta(1-\zeta)\langle k\rangle \tag{15}$$

giving the following transparent formulae for the post-attack transition points:

$$\text{F:}\quad \eta K(\beta J_0)(1-\zeta)[\langle k^2\rangle/\langle k\rangle - 1] = 1 \tag{16}$$

$$\text{SG:}\quad K^2(\beta J_0)(1-\zeta)[\langle k^2\rangle/\langle k\rangle - 1] = 1. \tag{17}$$

If, alternatively, we apply to this scenario the result (9), (10), we find $\gamma(\xi) = q(\xi)[\langle k^2\rangle/\langle k\rangle - 1]$ and $K(\beta J_0\xi\xi') = K(\beta J_0)\delta_{\xi,1}\delta_{\xi',1}$, and the relevant matrices reduce to $M^{(1)}_{\xi\xi'}(\beta) = \eta K(\beta J_0)(1-\zeta)[\langle k^2\rangle/\langle k\rangle - 1]\delta_{\xi,1}\delta_{\xi',1}$ and $M^{(2)}_{\xi\xi'}(\beta) = K^2(\beta J_0)(1-\zeta)[\langle k^2\rangle/\langle k\rangle - 1]\delta_{\xi,1}\delta_{\xi',1}$. One solves the eigenvalue problems trivially and indeed recovers (16), (17). A final trivial test is to consider $q(\xi|k) = \delta_{\xi,\xi_0}$, where $\xi_0 \in (0,1)$, an attack equivalent to replacing $J_0 \to \xi_0^2 J_0$. Upon substituting this choice into (9), (10) one confirms, via $\gamma(\xi) = \delta_{\xi,\xi_0}[\langle k^2\rangle/\langle k\rangle - 1]$, that our general theory indeed reduces to (12), (13) with the correctly reduced coupling strength.

Solving the eigenvalue problems (9), (10) analytically is not always possible, but eigenvalue bounds are obtained easily. Our matrices are of the form $M_{\xi\xi'} = L(\xi\xi')\gamma(\xi')$, where $L(u) = \eta K(\beta J_0 u)$ for the F transition (so $L(u)$ is anti-symmetric) and $L(u) = K^2(\beta J_0 u)$ for the SG transition (so $L(u)$ is symmetric), and where $\gamma(\xi) \geqslant 0$ for all ξ. We symmetrize the eigenvalue problem $\lambda x(\xi) = \sum_{\xi'} M_{\xi\xi'}x(\xi')$ by defining $y(\xi) = x(\xi)\sqrt{\gamma(\xi)}$, giving $\lambda y(\xi) = \sum_{\xi'}[\sqrt{\gamma(\xi)}L(\xi\xi')\sqrt{\gamma(\xi')}]y(\xi')$. This implies that

$$\lambda_{\max} = \max_{\mathbf{y}} \frac{\sum_{\xi\xi'} y(\xi)\sqrt{\gamma(\xi)}L(\xi\xi')\sqrt{\gamma(\xi')}y(\xi')}{\sum_\xi y^2(\xi)} \tag{18}$$

which can be simplified to

$$\lambda_{\max} = \max_{\mathbf{y}} \frac{\sum_{\xi\xi'} y(\xi)\sqrt{\gamma(\xi)}L(|\xi\xi'|)\sqrt{\gamma(\xi')}y(\xi')}{\sum_\xi y^2(\xi)}. \tag{19}$$

Variational arguments can now be applied in order to get lower bounds. In particular, upon substituting $y(\xi) = \delta_{\xi,\hat{\xi}}$ and varying $\hat{\xi}$ one derives the statement

$$\lambda_{\max} \geqslant \max_\xi \{\gamma(\xi)L(\xi^2)\}. \tag{20}$$

To find upper bounds, we use the fact that the maximum in (19) will have $y(\xi) \geqslant 0$ for all ξ. We then use the inequalities $L(|u|) \leqslant \alpha\eta\beta J_0|u|$ (for F) and $L(|u|) \leqslant (\alpha\beta J_0)^2|u|^2$ (for SG), where $\alpha = 1$ for Ising spins and $\alpha = \frac{1}{2}$ for coupled oscillators, to get

$$\lambda^{(1)}_{\max} \leqslant \alpha\eta\beta J_0 \max_{\mathbf{y}} \left\{ \frac{\sum_{\xi\xi'} y(\xi)\gamma^{\frac{1}{2}}(\xi)|\xi||\xi'|\gamma^{\frac{1}{2}}(\xi')y(\xi')}{\sum_\xi y^2(\xi)} \right\} \tag{21}$$

$$\lambda^{(2)}_{\max} \leqslant (\alpha\beta J_0)^2 \max_{\mathbf{y}} \left\{ \frac{\sum_{\xi\xi'} y(\xi)\gamma^{\frac{1}{2}}(\xi)|\xi|^2|\xi'|^2\gamma^{\frac{1}{2}}(\xi')y(\xi')}{\sum_\xi y^2(\xi)} \right\}. \tag{22}$$

The last two maxima are calculated easily, leading us to

$$\text{F:}\quad T_c[p,q] \leqslant \eta\alpha J_0 \sum_\xi \xi^2\gamma(\xi) \tag{23}$$

$$\text{SG:}\quad T_c[p,q] \leqslant \alpha J_0 \left(\sum_\xi \xi^4\gamma(\xi)\right)^{\frac{1}{2}}. \tag{24}$$

3.2. Explicit simple form for a process integrity measure $\Gamma[p, q]$

The inequalities (23), (24) become equalities for large c, where the critical temperatures diverge and hence $\beta \to 0$ in (12), (13); the right-hand sides of (23), (24) then become the true integrity measures of the process. Moreover, for certain natural choices of the set Ξ the latter statement is in fact true for *any* connectivity c. For instance, if $\Xi \subseteq \{0, 1\}$ (all nodes are either fully disabled or left alone) one may use $K(\beta J_0\xi\xi') = \xi\xi' K(\beta J_0)$ to diagonalize the matrices in (9), (10) and find

$$\text{F:} \quad \frac{1}{K(J_0/T_c[p,q])} = \eta \sum_{\xi} \xi^2 \gamma(\xi) \tag{25}$$

$$\text{SG:} \quad \frac{1}{K(J_0/T_c[p,q])} = \left(\sum_{\xi} \xi^4 \gamma(\xi)\right)^{\frac{1}{2}} \tag{26}$$

which reveals that the critical temperatures are monotonically increasing functions of the sums $\sum_\xi \xi^2\gamma(\xi)$ for the F-type order and $\sum_\xi \xi^4\gamma(\xi)$ for the SG-type order (for $\Xi = \{0, 1\}$ the two sums are in fact identical). In view of these properties, and in view of the minor differences between the F and SG cases, in the remainder of this study we adopt the quantity $\sum_\xi \xi^2\gamma(\xi)$ as our integrity measure, giving

$$\Gamma[p, q] = \frac{1}{\langle k\rangle} \sum_{\xi k} \xi^2 q(\xi|k) p(k) k(k-1). \tag{27}$$

We define the set of relevant degrees k as $S = \{k > 1 | p(k) > 0\}$. The optimal attack strategy is then the choice $q^\star[p]$ which solves the following optimization problem:

$$\text{minimize:} \quad \Gamma[p, q] \tag{28}$$

$$\text{subject to:} \quad q(\xi|k) \geqslant 0\, \forall (\xi, k), \qquad \sum_{\xi\in\Xi} q(\xi|k) = 1\, \forall k \in S \tag{29}$$

$$\sum_{\xi\in\Xi}\sum_{k\in S} (1-\xi)\phi(k) q(\xi|k) p(k) \leqslant \kappa^{-1}. \tag{30}$$

To avoid trivial pathologies we assume that $\exists k \geqslant 2$ with $p(k) > 0$ (if untrue we would not have an ordered state in the first place, as it would have given $T_c[p, q] = 0$), and that $\sum_k p(k)k(k-1) < \infty$ (if untrue there would not be a finite critical temperature before the attack). Clearly $q^\star(\xi|k) = \delta_{\xi,1}$ for $k \notin S$; any other choice would sacrifice attack resources without benefit. The best defence against optimal attacks is the choice for the degree distribution $p(k)$ such that the above minimum over q is maximized.

3.3. Bounds on the process integrity measure

To judge the quality of attack strategies it will prove useful to have bounds on the value $\Gamma[p, q^\star[p]]$ corresponding to the optimal attack $q^\star[p]$. An upper bound is easily obtained by inspecting the result of non-intelligent random attacks of the type $q(\xi|k) = (1-Q)\delta_{\xi,0} + Q\delta_{\xi,1}$, with $0 \leqslant Q \leqslant 1$:

$$\Gamma[p, q] = [\langle k^2\rangle/\langle k\rangle - 1]Q \tag{31}$$

$$Q \geqslant 1 - 1/\kappa\langle\phi(k)\rangle. \tag{32}$$

The sharpest bound of this form follows when seeking equality in the last line, giving[4]

$$\Gamma[p,q^\star[p]] \leqslant \Gamma[p,q^\star_{\text{random I}}] = \left(1-\frac{1}{\kappa\langle\phi(k)\rangle}\right)\langle k(k-1)\rangle/\langle k\rangle. \tag{33}$$

If $\Xi = \{0,1\}$ then (33) is the best possible upper bound based on random attacks. If $\Xi = [0,1]$ we can improve upon (33) by investigating random attacks of the form $q(\xi|k) = \delta[\xi - \hat{\xi}]$. The optimal choice turns out to be $\hat{\xi} = 1 - 1/\kappa\langle\phi\rangle$, giving

$$\Gamma[p,q^\star[p]] \leqslant \Gamma[p,q^\star_{\text{random II}}] = \left(1-\frac{1}{\kappa\langle\phi(k)\rangle}\right)^2\langle k(k-1)\rangle/\langle k\rangle. \tag{34}$$

To find lower bounds for $\Gamma[p,q^\star[p]]$ we first define modified probabilities $\pi(k)\in[0,1]$:

$$\pi(k) = \frac{\sum_{\xi\in\Xi}(1-\xi)q(\xi|k)\phi(k)p(k)}{\left\langle\sum_{\xi\in\Xi}(1-\xi)q(\xi|k)\phi(k)\right\rangle}, \tag{35}$$

with associated averages written as $\langle\cdots\rangle_\pi$. Note that the denominator of (35) is bounded from above by κ^{-1}, via the resource constraint. We can now write

$$\begin{aligned}
\Gamma[p,q] &= \frac{\langle k(k-1)\rangle}{\langle k\rangle} - \frac{1}{\langle k\rangle}\sum_k p(k)k(k-1)\sum_{\xi\in\Xi}(1-\xi^2)q(\xi|k)\\
&= \frac{\langle k(k-1)\rangle}{\langle k\rangle} - \frac{1}{\langle k\rangle}\left\langle\sum_{\xi\in\Xi}(1-\xi)q(\xi|k)\phi(k)\right\rangle\left\langle\frac{k(k-1)}{\phi(k)}\frac{\sum_{\xi\in\Xi}(1-\xi^2)q(\xi|k)}{\sum_{\xi\in\Xi}(1-\xi)q(\xi|k)}\right\rangle_\pi\\
&\geqslant \frac{\langle k(k-1)\rangle}{\langle k\rangle} - \frac{1}{\kappa\langle k\rangle}\left\langle\frac{k(k-1)}{\phi(k)}\frac{\sum_{\xi\in\Xi}(1-\xi^2)q(\xi|k)}{\sum_{\xi\in\Xi}(1-\xi)q(\xi|k)}\right\rangle_\pi\\
&\geqslant \frac{\langle k(k-1)\rangle}{\langle k\rangle} - \frac{C_\Xi}{\kappa\langle k\rangle}\left\langle\frac{k(k-1)}{\phi(k)}\right\rangle_\pi
\end{aligned} \tag{36}$$

in which the factor $C_\Xi \geqslant 0$ depends only on the choice made for the value set Ξ:

$$C_\Xi = \max_w\left\{\frac{1-\langle\xi^2\rangle_w}{1-\langle\xi\rangle_w}\right\} \quad \text{with} \quad \langle f(\xi)\rangle_w = \sum_{\xi\in\Xi}w(\xi)f(\xi) \quad \text{and} \quad \sum_{\xi\in\Xi}w(\xi)=1. \tag{37}$$

One easily proves using $\Xi\subseteq[0,1]$ that $C_\Xi\in[1,2]$, $C_{\{0,1\}}=1$ and $C_{[0,1]}=2$. We conclude, in combination with (33), (34), that

$$\Xi=\{0,1\}:\quad 1-\frac{1}{\kappa\langle k(k-1)\rangle}\left\langle\frac{k(k-1)}{\phi(k)}\right\rangle_\pi \leqslant \frac{\langle k\rangle\Gamma[p,q^\star[p]]}{\langle k(k-1)\rangle} \leqslant 1-\frac{1}{\kappa\langle\phi(k)\rangle} \tag{38}$$

$$\Xi=[0,1]:\quad 1-\frac{2}{\kappa\langle k(k-1)\rangle}\left\langle\frac{k(k-1)}{\phi(k)}\right\rangle_\pi \leqslant \frac{\langle k\rangle\Gamma[p,q^\star[p]]}{\langle k(k-1)\rangle} \leqslant \left(1-\frac{1}{\kappa\langle\phi(k)\rangle}\right)^2. \tag{39}$$

The lower bounds are satisfied with equality if the attack resources are exhausted and if $\langle(1-\xi^2)\rangle_q = \langle(1-\xi)\rangle_q$ for each $q(\xi|k)$ with $k\in S$; the last condition is always met if $\Xi=\{0,1\}$. However, the lower bounds still depend on the attack strategy via the measure π. From (38), (39) and the general property $\Gamma[p,q]\geqslant 0$, which follows from the definition of $\Gamma[p,q]$, we finally obtain the strategy-independent bounds

$$\Xi=\{0,1\}:\quad \max\left\{0, 1-\frac{R/\kappa}{\langle k(k-1)\rangle}\right\} \leqslant \frac{\langle k\rangle\Gamma[p,q^\star[p]]}{\langle k(k-1)\rangle} \leqslant 1-\frac{1}{\kappa\langle\phi(k)\rangle} \tag{40}$$

[4] Note that $\kappa\langle\phi(k)\rangle > 1$ due to our earlier ruling out of the trivial attack strategy $q(\xi|k)=\delta_{\xi,0}$.

$$\Xi = [0, 1]: \quad \max\left\{0, 1 - \frac{2R/\kappa}{\langle k(k-1)\rangle}\right\} \leqslant \frac{\langle k\rangle\Gamma[p, q^\star[p]]}{\langle k(k-1)\rangle} \leqslant \left(1 - \frac{1}{\kappa\langle\phi(k)\rangle}\right)^2 \tag{41}$$

with

$$R = \max_{k\in S}\left\{k(k-1)/\phi(k)\right\}. \tag{42}$$

The latter bounds reveal immediately two distinct situations where it is not possible for *any* intelligent attack to improve on the damage done by random attacks: the case $\phi(k) = k(k-1)$ for all $k \in S$ (here the benefit of degree knowledge exactly balances the cost to the attacker of using it), and the case of regular random graphs, namely $p(k) = \delta_{k,\langle k\rangle}$, where there is no degree knowledge to be exploited in the first place.

4. Optimal attack and optimal defence for $\Xi = \{0, 1\}$

The attacker's objective is to minimize $\Gamma[p, q]$. We have seen that for $\Xi = \{0, 1\}$, where nodes are either fully disabled or left alone and $C_{\{0,1\}} = 1$, the lower bound in (40) could in principle be realized. This will serve as an efficient guide in finding $q^\star[p]$. Attack strategies for $\Xi = \{0, 1\}$ are of the form $q(\xi|k) = q(0|k)\delta_{\xi,0} + [1 - q(0|k)]\delta_{\xi,1}$, so we need to determine $q(0|k)$ for all $k \in S$.

4.1. Construction of the optimal attack strategy

We first define the attacker's 'target' degree set $\mathcal{A} \subseteq S$, with R as defined in (42):

$$\mathcal{A} = \{k \in S|k(k-1)/\phi(k) = R\}. \tag{43}$$

The inequality $\langle k(k-1)/\phi(k)\rangle_\pi \leqslant R$ used in the final step of our derivation of (40) is satisfied with *equality* only if $\pi(k) = 0$ for all $k \notin \mathcal{A}$. According to (35) this requires $q(\xi|k) = \delta_{\xi,1}$ for all $k \notin \mathcal{A}$. The only remaining requirement for satisfying the lower bound in (40) is that we satisfy the resource constraint with equality. Hence, the set of optimal attack strategies is defined strictly by the following demands:

$$\forall k \notin \mathcal{A}: \quad q(0|k) = 0 \tag{44}$$

$$\forall k \in \mathcal{A}: \quad q(0|k) \in [0, 1], \qquad \sum_{k\in\mathcal{A}} q(0|k)\phi(k)p(k) = 1/\kappa. \tag{45}$$

It is straightforward to verify directly, using $\phi(k) = k(k-1)/R$ for all $k \in \mathcal{A}$, that strategies satisfying these conditions indeed give the lowest possible value for $\Gamma[p, q]$ according to our bounds and satisfy the resource constraint with equality. By construction, the set $\mathcal{A}$ cannot be empty.

At this stage in our argument we must distinguish between two distinct cases. In the first case the attacker need not look beyond nodes in the target set $\mathcal{A}$ (43), since removing those will already exhaust or exceed his resources; he will simply remove as many of those as can be afforded. In the second case, the removal of all nodes in $\mathcal{A}$ does not exhaust the attack resources, and new target sets need to be identified.

- The target set $\mathcal{A}$ is exhausting, $\sum_{k\in\mathcal{A}}\phi(k)p(k) \geqslant 1/\kappa$:

  Here it is immediately clear that optimal attacks will indeed exist, i.e. the conditions (44), (45) can be met. Only nodes from $\mathcal{A}$ will be removed. If there is at least one $k^\star \in \mathcal{A}$ with $\phi(k^\star)p(k^\star) \geqslant 1/\kappa$, the attacker can simply execute

$$k^\star = \underset{k\in\mathcal{A}}{\operatorname{argmax}}\{\phi(k)p(k)\} \tag{46}$$

$$q(0|k^\star) = 1/\kappa p(k^\star)\phi(k^\star), \qquad \forall k \neq k^\star : q(0|k) = 0. \tag{47}$$

If instead $\phi(k)p(k) < 1/\kappa$ for all $k \in \mathcal{A}$ there is no target degree in $\mathcal{A}$ which would on its own exhaust the attacker's resources. The attacker will first remove all nodes with degree $k_1^\star = \operatorname{argmax}_{k\in\mathcal{A}}\{\phi(k)p(k)\}$ by setting $q(0|k_1^\star) = 1$. He will next direct attention to the reduced set $\mathcal{A}/\{k_1^\star\}$ and remove nodes with degree $k_2^\star = \operatorname{argmax}_{k\in\mathcal{A}/\{k_1^\star\}}\{\phi(k)p(k)\}$, etc until the resources are exhausted. At the end of this iterative process the attacker will have removed a sequence of degrees $\{k_1^\star, \ldots, k_L^\star\} \subseteq \mathcal{A}$ (where nodes with degree $k_L^\star$ will generally be only partially removed, as allowed by remaining resources). In words the attacker first determines the target set $\mathcal{A}$ of those degrees with $p(k) > 0$ for which the ratio $k(k-1)/\phi(k)$ is maximal. He then ranks the degrees in $\mathcal{A}$ according to the value of $\phi(k)p(k)$ and proceeds to remove degrees iteratively according to this ranking until his resources are exhausted. This strategy will always lead to $\langle k(k-1)/\phi(k)\rangle_\pi = R$ and satisfy the lower bound in (40) with equality.

- The target set $\mathcal{A}$ is non-exhausting, $\sum_{k\in\mathcal{A}}\phi(k)p(k) < 1/\kappa$:

  Here the attacker can afford to remove completely all degrees in the set $\mathcal{A}$, but setting $q(0|k) = 1$ for all $k \in \mathcal{A}$ does not exhaust his resources. He should subsequently direct attention to those nodes in the reduced set $S/\mathcal{A}$ for which the ratio $k(k-1)/\phi(k)$ is maximal, and so on. The result is again an iteration, at the end of which the attacker will have removed a set of degrees $\{k_1^\star, \ldots, k_L^\star\} \supset \mathcal{A}$ (where nodes with degree $k_L^\star$ will generally be only partially removed). In this case $\langle k(k-1)/\phi(k)\rangle < R$ and the lower bound in (40) is no longer satisfied with equality; however, this does not imply that the strategy is non-optimal, since it might be that the bound is no longer tight. Here it is therefore difficult to prove rigorously that the identified strategy always constitutes the optimal attack, but it is the logical continuation of the optimal attack identified earlier and its optimality is consistent with numerical experiments (to be shown later).

We can combine both cases above in a transparent iterative attack protocol. We define at each step ℓ: the target set $\mathcal{A}_\ell$, the set S_ℓ of nodes that have not yet been targeted and the resource remainder $\Delta_\ell = \kappa^{-1} - \sum_k q(0|k)\phi(k)p(k)$. The process is initialized according to $S_0 = S$ and $\Delta_0 = \kappa^{-1}$, and starts with $q(0|k) = 0$ for all k. It is iterated until $\Delta_\ell = 0$, according to

step 1: calculate the new ratio $R_\ell = \max_{k\in S_{\ell-1}}\{k(k-1)/\phi(k)\}$
step 2: identify the target set $\mathcal{A}_\ell = \{k \in S_{\ell-1} | k(k-1)/\phi(k) = R_\ell\}$
step 3: choose (any) $k_\ell^\star \in \mathcal{A}_\ell$ for which $\phi(k_\ell^\star)p(k_\ell^\star) = \max_{k\in\mathcal{A}_\ell}\{\phi(k)p(k)\}$
step 4: check whether attack resources can be exhausted:

$\phi(k_\ell^\star)p(k_\ell^\star) \geqslant \Delta_{\ell-1}$: yes,
remove as many degree $k_\ell^\star$ nodes as possible
set $q(0|k_\ell^\star) = \Delta_{\ell-1}/\phi(k_\ell^\star)p(k_\ell^\star)$
$\Delta_\ell = 0$, attack terminates

$\phi(k_\ell^\star)p(k_\ell^\star) < \Delta_{\ell-1}$: no,
remove all degree $k_\ell^\star$ nodes
set $q(0|k_\ell^\star) = 1$

step 5: define $S_\ell = S_{\ell-1}/k_\ell^\star$ and $\Delta_\ell = \Delta_{\ell-1} - \phi(k_\ell^\star)p(k_\ell^\star)q(0|k_\ell^\star)$.

Always the end result is a sequence $\{k_1^\star,\ldots,k_{L-1}^\star\}$ of target degrees that are fully removed, possibly supplemented by a further degree $k_L^\star$ of which a fraction will be removed (to exhaust fully the attack resources).

4.2. *Properties of the optimal attack strategy*

We next evaluate the impact of the above attack strategy $q^\star[p]$ on our process integrity measure. We define the set $\mathcal{A}^\star=\{k_1^\star,\ldots,k_{L-1}^\star\}$ of fully removed degrees, and write $k_L^\star$ simply as $k^\star$. The post-attack value of the process integrity measure will be

$$\begin{aligned}\Gamma[p,q^\star[p]]&=\frac{\langle k(k-1)\rangle}{\langle k\rangle}-\frac{1}{\langle k\rangle}\sum_{k\in\mathcal{A}^\star}p(k)k(k-1)-\frac{1}{\langle k\rangle}q(0|k^\star)k^\star(k^\star-1)p(k^\star)\\&=\frac{\langle k(k-1)\rangle}{\langle k\rangle}-\frac{1}{\langle k\rangle}\sum_{k\in\mathcal{A}^\star}p(k)k(k-1)-\frac{\Delta_{L-1}}{\langle k\rangle}\frac{k^\star(k^\star-1)}{\phi(k^\star)}.\end{aligned}\tag{48}$$

The attack $q^\star[p]$ exhausts all resources, so $\Delta_{L-1}=\kappa^{-1}-\sum_{k\in\mathcal{A}^\star}\phi(k)p(k)$. Hence,

$$\Gamma[p,q^\star[p]]=\frac{\langle k(k-1)\rangle}{\langle k\rangle}-\frac{1}{\langle k\rangle}\sum_{k\in\mathcal{A}^\star}p(k)\phi(k)\left[\frac{k(k-1)}{\phi(k)}-\frac{k^\star(k^\star-1)}{\phi(k^\star)}\right]-\frac{k^\star(k^\star-1)}{\kappa\langle k\rangle\phi(k^\star)}.\tag{49}$$

Since by definition $k(k-1)/\phi(k)>k^\star(k^\star-1)/\phi(k^\star)$ for all $k\in\mathcal{A}^\star$, both the second and the third term of (49) are strictly non-positive.

The result (49) can be compared to that of random attack (where no degree information is used), namely to (33). The benefit $\Delta\Gamma=\Gamma[p,q^\star[p]]-\Gamma[p,q^\star_{\rm random}]$ to the attacker of using optimal attacks as opposed to random attacks then takes the form

$$\begin{aligned}\Delta\Gamma&=-\frac{1}{\langle k\rangle}\sum_{k\in\mathcal{A}^\star}p(k)\phi(k)\left[\frac{k(k-1)}{\phi(k)}-\frac{k^\star(k^\star-1)}{\phi(k^\star)}\right]-\frac{1}{\kappa\langle k\rangle}\left[\frac{k^\star(k^\star-1)}{\phi(k^\star)}-\frac{\langle k(k-1)\rangle}{\langle\phi(k)\rangle}\right]\\&=-\frac{1}{\langle k\rangle}\sum_{k\in\mathcal{A}^\star}p(k)\phi(k)\left[\frac{k(k-1)}{\phi(k)}-\frac{k^\star(k^\star-1)}{\phi(k^\star)}\right]\\&\quad-\frac{1}{\kappa\langle k\rangle\langle\phi(k)\rangle}\sum_k p(k)\phi(k)\left[\frac{k^\star(k^\star-1)}{\phi(k^\star)}-\frac{k(k-1)}{\phi(k)}\right]\\&=-\frac{1}{\langle k\rangle}\sum_{k\in\mathcal{A}^\star}p(k)\phi(k)\left[\frac{k(k-1)}{\phi(k)}-\frac{k^\star(k^\star-1)}{\phi(k^\star)}\right]\left(1-\frac{1}{\kappa\langle\phi(k)\rangle}\right)\\&\quad+\frac{1}{\kappa\langle k\rangle\langle\phi(k)\rangle}\sum_{k\notin\mathcal{A}^\star}p(k)\phi(k)\left[\frac{k(k-1)}{\phi(k)}-\frac{k^\star(k^\star-1)}{\phi(k^\star)}\right].\end{aligned}\tag{50}$$

Since the set $\mathcal{A}^\star\subseteq S$ is constructed specifically from those degrees for which $k(k-1)/\phi(k)$ is maximal, and since $\langle\phi(k)\rangle>\kappa^{-1}$, both terms of $\Delta\Gamma$ are strictly non-positive. One will thus generally have $\Delta\Gamma<0$. Again we also recognize the two special cases where there will be no gain in intelligent attacks, namely $\phi(k)=k(k-1)$ (with any degree distribution) and $p(k)=\delta_{k,\langle k\rangle}$ (with any cost function $\phi(k)$).

In terms of the dependence of our results on the cost function $\phi(k)$ it is clear that everything evolves around the dependence on k of the ratio $\phi(k)/k(k-1)$. This ratio represents for each k the balance between the cost of removing degree-k nodes versus the benefits in terms of damage achieved. If for simplicity we choose $\phi(k)=k^\zeta(k-1)$, then for $\zeta<1$ the intelligent attack will be to take out first the nodes with the *largest* degrees $k\in S$ that can be removed without violating the resource constraint (i.e. the 'greedy' attack strategy is optimal), whereas

Table 1. Size N, average connectivity $\langle k\rangle$, maximum degree $k_{\max}$, experimental detection method or source, and reference for the biological (protein interaction) network data sets used in our numerical experiments. The detection methods or sources are abbreviated as follows: Y2H, yeast two-hybrid; PMS, purification-mass spectrometry; HPRD, the human protein reference database.

Species	N	$\langle k\rangle$	$k_{\max}$	Method	Reference
C. elegans	3512	3.72	524	Y2H	[12]
C. jejuni	1324	17.52	207	Y2H	[13]
E. coli	2457	7.05	641	PMS	[14]
H. sapiens	9306	7.53	247	HPRD	[15]
S. cerevisiae	3241	2.69	279	Y2H	[16]

for $\zeta > 1$ the intelligent attack will target first the nodes with the *smallest* degrees $k \in S$ that can be removed without violating the constraint (here attacking hubs is too expensive to be efficient). Furthermore, it is not at all *a priori* clear what would be the most resistant degree distribution against such attacks in the presence of resource constraints. Naively one could perhaps have expected that for small ζ (where the attacker will target hubs) the best strategy for the defender could be to choose a narrowly distributed degree distribution, so there are no hubs to be exploited. Interestingly, we will see below that that is not the case, and the optimal degree distribution can be more subtle.

5. Numerical results

5.1. General methods

In this section we illustrate, apply and extend via numerical experimentation the results derived above. We determine by numerical maximization the most resistant degree distribution against optimal intelligent attacks, and we compare for typical biological networks the effects of optimal intelligent attacks in terms of process integrity against random attacks and against the bounds established earlier. The biological networks used are experimentally determined protein interaction networks (PINs) of different species, namely *Caenorhabditis elegans*, *Campylobacter jejuni*, *Escherichia coli*, *Homo sapiens* and *Saccharomyces cerevisiae*; see table 1 for characteristics and references. For each biological network we also generate several synthetic alternatives with the same size N and average connectivity $\langle k\rangle$ as the biological one, but with different degree distributions: Poissonian, the optimally resistant degree distribution, or a distribution generated via preferential attachment with a fat tail similar to the biological network. In all cases we choose node attack cost functions of the form $\phi(k) = k^{\zeta}(k-1)$, with $\zeta = 0, 1, 2$, and we set the attack resource limit to $\kappa^{-1} = \langle\phi(k)\rangle_{\mathcal{P}}/q$, where $q > 1$ is a control parameter and the average $\langle\cdots\rangle_{\mathcal{P}}$ is calculated over the Poissonian distribution $\mathcal{P}(k) = \langle k\rangle^{k} e^{\langle k\rangle}/k!$ For each PIN and each synthetically generated counterpart the average cost function $\langle\phi(k)\rangle$ is found to be always larger than or equal to the one calculated over the equivalent Poissonian distribution (if $\zeta = 0$, where $\phi(k) = k-1$, equality of course holds trivially for *all* distributions since they share by construction the value of $\langle k\rangle$). This ensures that for $q > 1$ the attacker's resources will in all our experiments be in the relevant regime $\langle\phi(k)\rangle > \kappa^{-1}$. The optimally attack resistant networks are found via a stochastic graph dynamics, starting from a biological protein interaction network, in which at each step a bond is selected at random and is moved to another location if this move increases the post-attack integrity measure $\Gamma[p, q^{\star}[p]]$. Bond relocations are the minimal moves that preserve the average degree of the network. After each move, the optimal attack strategy $q^{\star}[p]$ defined in the previous section is applied to the new network. In order to prevent the graph dynamics

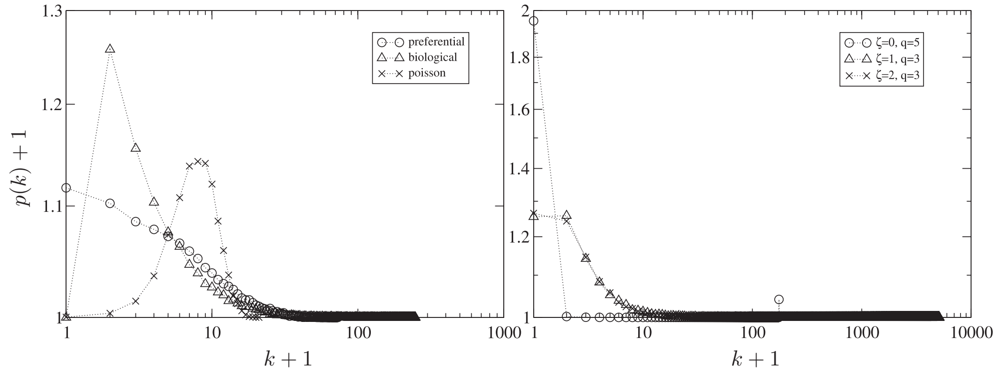

Figure 1. Left: log–log plot of the degree distribution of the *H. sapiens* PIN, synthetically generated networks with the same size *N* and average connectivity $\langle k\rangle$ as *H. sapiens* PIN but with Poissonian and preferential attachment degree distributions. Right: log–log plot of the degree distribution of the network that has the same size and average connectivity as the *H. sapiens* PIN, but that has been constructed to be optimally resistant against optimal intelligent attacks, given the node removal cost function $\phi(k) = k^{\zeta}(k-1)$, with $\zeta = 0, 1, 2$, and given available attack resources, characterized by $\kappa^{-1} = \frac{1}{q}\langle\phi(k)\rangle_{\mathcal{P}}$, with $q = 3, 5$ (see the legend). We observe that upon decreasing ζ, where the optimal attack strategy starts targeting the high-degree nodes, the optimally resistant degree distribution takes a binary form, describing a module of high-degree nodes in a sea of unconnected nodes.

from getting stuck in suboptimal configurations, we allow initially for groups of bonds to be moved and as the algorithm proceeds, the size of these groups is reduced, in the spirit of [17].

5.2. Degree statistics before and after network optimization

Figure 1 shows the results of applying the above procedures to the *H. sapiens* PIN for $\zeta = 0, 1, 2$ and $q = 3, 5$. For $\zeta = 1, 2$ (where it is not advantageous to the attacker to target high-degree nodes) the optimally resistant degree distribution $p^{\star}[q]$ is seen to exhibit a smooth dependence on the degree *k*. For $\zeta = 0$, where the degree dependence of node removal costs is modest and the optimal attack strategy is to target high degree nodes, one could expect the optimal network to become regular, in order to disallow attackers to benefit from degree information. Instead, we observe an entirely different solution. Here, the optimal defender produces as many hubs as possible, so that the attacker is unable to remove all of these. The result is a distribution of the form

$$p^{\star}(k) = \left(1 - \frac{\langle k\rangle}{K}\right)\delta_{k,0} + \frac{\langle k\rangle}{K}\delta_{k,K}, \qquad K \geqslant \langle k\rangle \tag{51}$$

with $K = 192$ for a network with size and average connectivity identical to the *H. sapiens* PIN. In a situation where attackers can and will target nodes with maximal degree first, it appears that the optimal defender chooses a network with a 'modular' configuration, with a core of nodes highly connected to each other, in a sea of disconnected nodes. The attacker is prevented by resource limitations from removing more than a (tiny) fraction of the core. The strategy of the optimal defender is to sacrifice a few highly connected nodes to save many.

For distributions of the form (51) the attack resources can be exhausted, and the optimal attack is $q^{\star}(0|k) = \delta_{k,K}K/\kappa\phi(K)\langle k\rangle$. This results in

$$\Gamma[p, q^{\star}[p]] = K - 1 - K^{1-\zeta}/\kappa\langle k\rangle. \tag{52}$$

Insertion into (40) shows that both bounds are now satisfied with equality. For $\zeta > 0$ the defender would wish to choose K as large as possible, but for a finite network there is a limit. There are just $N\langle k\rangle/K$ connected nodes; if each of these is to have K neighbours we must demand $N\langle k\rangle/K - 1 \geqslant K$, i.e. $K \leqslant K_c = \sqrt{N\langle k\rangle} + \mathcal{O}(N^0)$. The same result follows from general entropic arguments [18, 19]. For large N the number of graphs with degree distribution $p(k)$ equals $\exp[NS]$ where $S = \frac{1}{2}\langle k\rangle[\log(N/\langle k\rangle)+1] - \sum_k p(k)\log[p(k)/\pi(k)]$ with $\pi(k) = \mathrm{e}^{-\langle k\rangle}\langle k\rangle^k/k!$. For (51) one obtains

$$S = \frac{1}{2}\langle k\rangle\left\{1 - \log\left(\frac{K^2}{N\langle k\rangle}\right)\right\} + \mathcal{O}(K^{-1}\log K). \tag{53}$$

Again we obtain the cut-off point $K \leqslant K_c \approx \sqrt{N\langle k\rangle}$ for graphs with (51) to exist. The value $K = 192$ found numerically, see figure 1, is consistent with this bound (for *H. sapiens* one has $K_c = 264$), but not identical to it. This is expected to reflect finite size corrections to our theory, and the fact that the theory requires all relevant k to be finite relative to N, whereas close to K_c one has $k = \mathcal{O}(\sqrt{N})$.

Note, however, that the distribution chosen by the optimal defender is not always exactly of the form (51). In some situations (depending on the amount of resources available to the attacker) the peak at $k = 0$ is not strictly δ-shaped, so that K is no longer subjected to the previously identified cut-off K_c, and one indeed observes the second peak to move to higher values of K (albeit with a reduced height). This results in a bimodal distribution with a δ-peak at some $K > K_c$ and a broader peak at $k = 0$, corresponding to a strongly disassortative network configuration (for the notion of assortativity see e.g. [20]) where a small number of hubs are connected with an extremely large number of low degree nodes (reminiscent of results derived in [21]). For this later distribution, as was the case with the bimodal distribution (51) with two strictly δ-shaped peaks, the attacker will again exhaust his resources upon removal of just a tiny fraction, $q(0|K) = 1/\kappa\phi(K)p(K)$, of hubs.

The actual distribution $p^\star(k)$ selected by the optimal defender when the attacker is bound, by resource limitations, to play the strategy $q^\star(0|k) = \delta_{k,K}\, 1/\kappa\phi(k)p(k)$ is the one which maximizes the minimal integrity measure

$$\Gamma(p, q^\star[p]) = \frac{1}{\langle k\rangle}\left(\langle k(k-1)\rangle - \sum_k p(k)k(k-1)q(0|k)\right) \tag{54}$$

$$= \frac{1}{\langle k\rangle}(\langle k(k-1)\rangle - \kappa^{-1}K^{1-\zeta}) \tag{55}$$

achieved by the attacker. It is clear that the shape of the optimally resistant distribution will depend on the interplay between $\langle k^2\rangle$ and K, which is controlled by the resource limit κ^{-1}. For $\zeta = 0$, numerical studies show that for q sufficiently small (large amount of resources) $p^\star(k)$ assumes the shape (51), whereas for large q (small amount of resources) the width of the peak at $k = 0$ increases and the second peak moves to $K \gg K_c$.

5.3. Values of process integrity measures before and after attacks

In figure 2 we plot the integrity measure $\Gamma[p, q]$, for the different network distributions considered, before and after an optimal intelligent attack. We also show the values that the integrity measure would take after a random attack (where sites are picked up at random and removed until resources are exhausted), with the same attack resource limit. The node attack cost function and resource limit chosen are $\phi(k) = k^\zeta(k-1)$ and $\kappa^{-1} = \langle\phi(k)\rangle_{\mathcal{P}}/q$, respectively, with $\zeta = 0, 1, 2$ and $q > 1$. In addition we show the lower and upper bounds

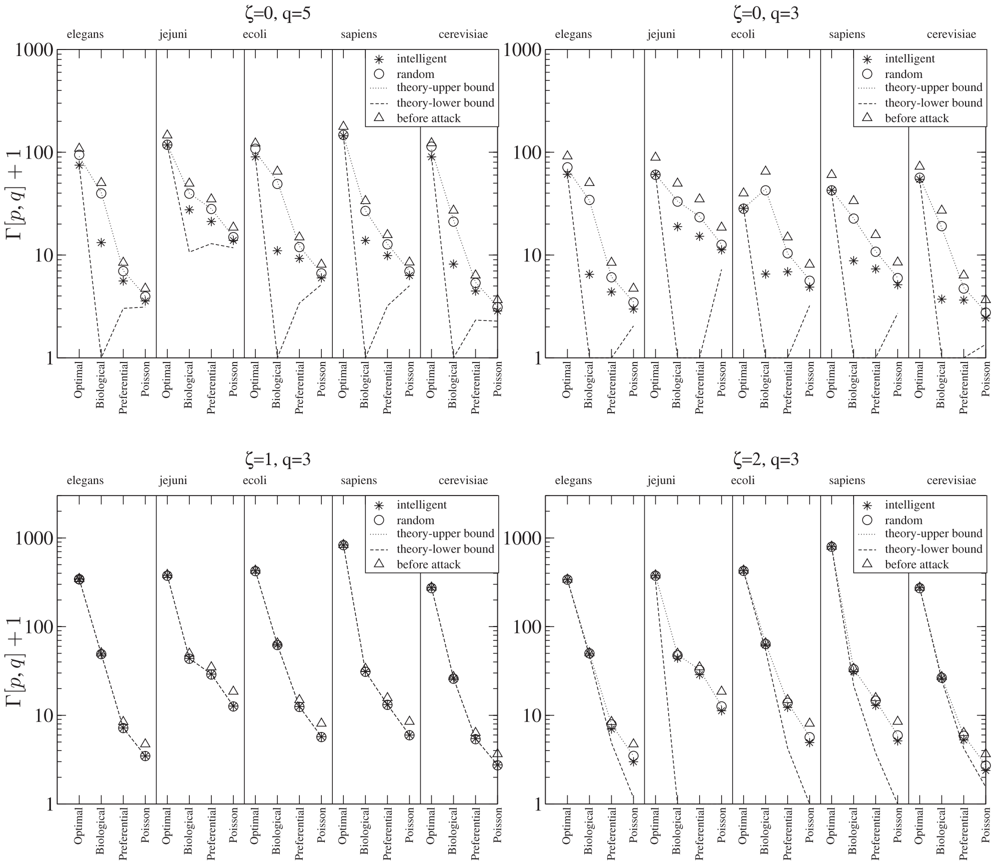

Figure 2. Values of the process integrity measure $\Gamma[p, q]$ before attack ($\triangle$), after optimal intelligent attacks ($*$) and after optimal random attacks ($\circ$), for different networks. The specific networks considered are experimentally determined PINs of different species (*C. elegans*, *C. jejuni*, *E. coli*, *H. sapiens* and *S. cerevisiae*) and their synthetically generated counterparts with the same size and average connectivity, but different degree distributions (Poissonian, preferential attachment and optimally attack resistant degree distribution following the attack $q^{\star}[p]$). The node attack cost function is $\phi(k) = k^{\zeta}(k-1)$ and the available attack resources are characterized by $\kappa^{-1} = \langle\phi(k)\rangle_{\mathcal{P}}/q$ with $\zeta = 0, 1, 2$ and $q = 3, 5$ (see the legends). The theoretical upper and lower bounds (40) are shown as dotted and dashed lines, respectively. All results consistently reproduce the built-in order $\Gamma_{\text{before}} \geqslant \Gamma_{\text{random}} \geqslant \Gamma_{\text{intelligent}}$ (namely $\triangle \geqslant \circ \geqslant *$). Furthermore, the network realizations are consistently ranked, with the optimally resistant network (as expected) always outperforming the others, but with also the biological and preferential attachment network outperforming their Poissonian counterparts. In fact, the degree of resistance of the optimally resistant network is quite remarkable.

(40) on the process integrity measure after an optimal intelligent attack $q^{\star}[p]$, as described by the protocol in the previous section, as dashed and dotted lines. As expected, the data points for random attacks always coincide with the dotted line of the theoretical upper bound (since such strategies formed the basis from which the upper bound was derived). Optimally attack-resistant degree distributions are expected and indeed seen to be the ones for which the integrity measure of the network process after optimal intelligent attack is the highest, but when quantified via $\Gamma[p, q]$ as in the figure one is struck by how well they perform, i.e.

by the remarkably small reduction in the process integrity measure which they exhibit. For $\zeta = 1$, the theoretical lower and upper bounds coincide with each other, and with the integrity measure values for random attacks and optimal intelligent attacks. Here $\phi(k) = k(k-1)$, so costs and benefit for the attacker of degree knowledge balance each other out, and we have already shown that there is then no scope for the intelligent attacker to improve on the damage inflicted by random attacks. One can often understand the actual values obtained for $\Gamma[p, q]$. The relative reduction of the integrity measure before and after random attacks, for instance, can be calculated from (40), and for our resource limit $\kappa^{-1} = \langle\phi(k)\rangle_{\mathcal{P}}/q$, this gives $(\Gamma_{\text{before}} - \Gamma_{\text{after}})/\Gamma_{\text{before}} = \langle\phi(k)\rangle_{\mathcal{P}}/q\langle\phi(k)\rangle$. For $\zeta = 0$ this is always equal to $1/q$; for $\zeta = 1, 2$ it is small for degree distributions with large second and third moments, much larger than for a Poissonian distribution. Similarly, for optimal intelligent attacks equation (40) yields an upper bound on the relative attack-induced reduction of the process integrity measure: $(\Gamma_{\text{before}} - \Gamma_{\text{after}})/\Gamma_{\text{before}} \leqslant [\langle k^{\zeta}(k-1)\rangle_{\mathcal{P}}/q\langle k(k-1)\rangle]\max_{k\in S}\{k^{1-\zeta}\}$. For $\zeta = 1, 2$, this change is again small for optimally resistant and biological networks, as a result of their large degree variance (except for *C. jejuni*, which is distinct due to an unusually large average connectivity).

Our results re-confirm that random and hub-targeted attacks have similar effects on Poissonian graphs, as often remarked in the literature, due to the large homogeneity of the degrees. For regular graphs they would have produced identical results. However, one should be careful in concluding from this that processes running on Poissonian networks are hence the most resistant ones against hub removal. In contrast, figure 2 shows that they are the most vulnerable ones, as their post-attack integrity measure is the smallest. Interestingly, we find that processes running on networks produced by a preferential attachment mechanism are more resistant than those running on Poissonian networks, against both random attacks *and* optimal intelligent attacks. All this is due to the profound impact of resource constraints on the network resilience problem. Moreover, the degree distributions found in biological PINs generally exhibit, in turn, higher values for the post-attack process integrity measure than both Poissonian and preferential attachment networks, for random attacks and optimal intelligent attacks. A final feature emerging from figure 2 is that, while overall more robust compared to their preferential attachment and Poisonnian counterparts, biological networks seem significantly more resilient against random attacks than against hub-targeted attacks (see the top two panels with $\zeta = 0$, where the optimal attack indeed targets hubs).

5.4. Connection with results of previous studies—fraction of removed nodes

Previous studies of network resilience, based on the analysis of static topological properties of networks under attacks, had shown that power law networks (such as the ones produced by a preferential attachment mechanism) are more resistant than Poissonian ones against random removal of a *fixed fraction* of nodes (see e.g. [22, 23]), but are very vulnerable against hub removal [1, 24]. In the light of our new results, one may wonder how the fraction f of nodes removed varies among different degree distributions, when considering attacks constrained by degree-dependent node removal costs, with limited attack resources. The results of numerical explorations for $\zeta = 0$ (where optimal attacks will target hubs) and $q = 3$ are shown in figure 3, in the form of scatter plots of the relative variation $\Delta\Gamma/\Gamma = (\Gamma_{\text{before}} - \Gamma_{\text{after}})/\Gamma_{\text{before}}$ of the integrity measure under optimal attacks versus the fraction of sites removed (left) and under random attacks versus the fraction of nodes removed (right). The dotted line in the latter plot shows the theoretically predicted (constant) value of the relative variation of the process integrity measure under random attacks for $\zeta = 0$. Figure 3 reveals that biological PINs are indeed affected by a dramatic drop in the integrity measure under hub removal, even for tiny fractions of removed nodes; this confirms our intuition that their observed resilience depends

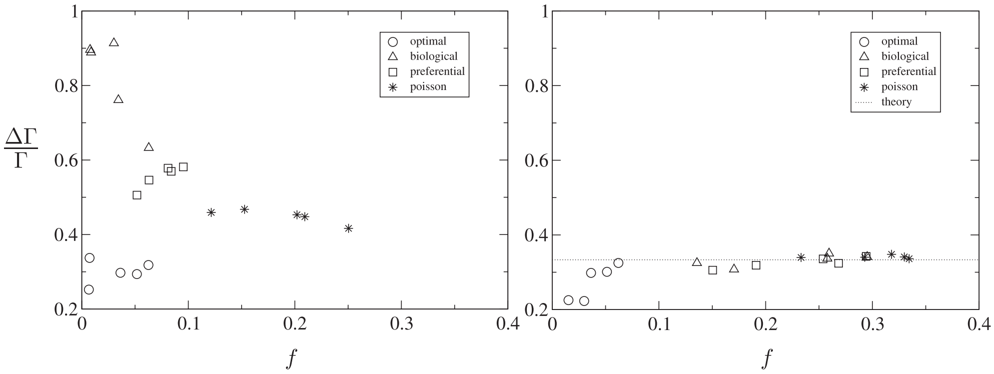

Figure 3. Scatter plots of the relative variation $\Delta\Gamma/\Gamma$ of the integrity measure versus the fraction f of sites removed under optimal intelligent (left) and random (right) attacks, with $\phi(k) = k - 1$ and $\kappa^{-1} = \langle\phi(k)\rangle/3$. Different markers correspond to different network families. The biological family is composed of the five experimentally determined PINs of table 1. The other families are the synthetically generated counterparts of the biological PINs, with the same size and average connectivity, but different degree distributions (Poissonian, preferential attachment and optimally resistant against intelligent attack).

crucially on having degree statistics such that attack costs prevent intelligent attackers from removing significant number of hubs. The same statement is expected to apply to any random graph drawn from the ensemble (4), where the imposed local degrees are those of the biological networks. Finally, figure 3 also shows that, as expected, the fraction of sites removed during hub-targeted attacks in a Poissonian graph, with fixed attack resources and when the node removal cost function is monotonically increasing with k, is considerably larger than in power law graphs. We conclude that the often claimed superiority of Poissonian networks over power law graphs for hub-targeted attacks is strictly a consequence of the decision to keep the fraction of removed sites fixed. This is consistent with the findings in [3], where it was argued that power law networks are no longer more fragile than Poissonian graphs against hub-targeted attacks when one looks at the number of removed *links*, and that the efficiency of hub removal in power law graphs would mainly lie in the fact that this removes many more links than it would have in Poissonian graphs.

In order to make contact with earlier results in the literature, we consider below optimal attacks calculated for the constant cost function $\phi(k) = \phi$. Here, the effects of attack costs should vanish from the problem, and our attacks should reduce to those where the fraction f of degrees to be removed is kept fixed. In fact, from the resource constraint one has $f = \sum_k p(k)q(0|k) = 1/\kappa\phi$. We plot in figure 4 (left) the integrity measure $\Gamma[p, q]$ for the different networks considered so far, before and after optimal intelligent and random attacks, with the constant cost function $\phi = \langle k\rangle - 1$ and resources $\kappa^{-1} = \phi/5$. In the right panel we show a scatter plot of the relative variation of the process integrity measure under optimal intelligent attack versus the fraction of removed sites, similar to figure 3. The fraction of removed sites is now constant, as expected, and indeed equals 1/5 for our choice of the resource limit. We see that for the constant node removal cost function the dependence of the post-attack integrity measure on the degree distribution is drastically different from that in the case of monotonically increasing cost functions, and we retrieve the old results known from the literature: power law networks are now more resilient than Poissonian ones against random attacks, but are extremely sensitive to hub removal.

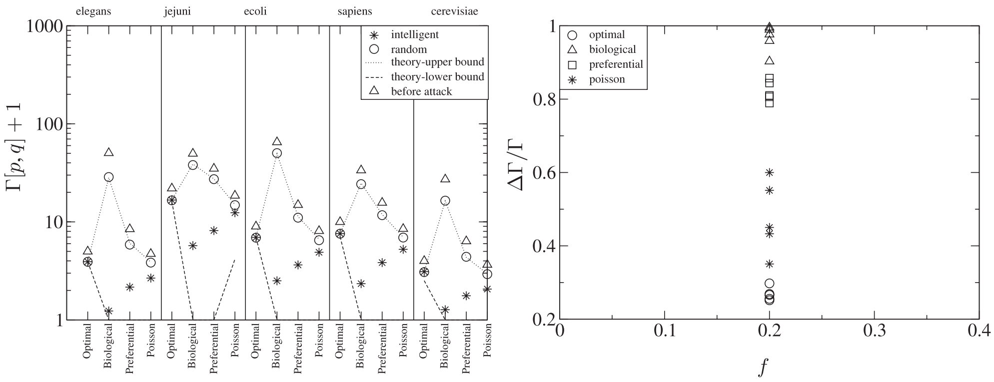

Figure 4. Left: process integrity measure $\Gamma[p, q]$ before and after optimal intelligent and random attacks, when the node removal cost function is a constant, $\phi(k) = \langle k\rangle - 1$, and attack resources are constrained according to $\kappa^{-1} = \langle\phi(k)\rangle/5$. The theoretical upper and lower bounds of (40) are shown as dotted and dashed lines, respectively. Right: a scatter plot of the relative variation $\Delta\Gamma/\Gamma$ of the integrity measure versus the fraction of sites removed under intelligent attacks, with attack resources constrained according to $\kappa^{-1} = \langle\phi(k)\rangle/5$ and with node attack cost function $\phi(k) = \langle k\rangle - 1$. Here different markers correspond to different network families, similar to figure 3.

5.5. *Misinformation*

We finally illustrate briefly the possible effects on the network resilience problem of misinformation, i.e. a situation where a network is designed to be optimally resistant against an optimal intelligent attack on the basis of a node removal cost function $\phi_d(k)$ and a resource limit κ_d^{-1}, but where in fact it faces an optimal intelligent attack constrained by an actual cost function $\phi_a(k)$ and with resource limit κ_a^{-1}. Here the cost functions $\phi_\ell(k)$ and resource limits κ_ℓ^{-1} are defined as follows (with integer $\ell \geqslant 1$): $\phi_1(k) = \langle k\rangle - 1$, $\phi_{\ell>1}(k) = k^{\ell-2}(k-1)$ and $\kappa_\ell^{-1} = \langle\phi_\ell(k)\rangle_{\mathcal{P}}/q_\ell$ with $q_1 = 5$ and $q_{\ell>1} = 3$. Note that for $a = 3$ the intelligent attacks in fact reduce to random ones. The results of our numerical explorations are shown in figure 5. For every attack, the distribution for which the post-attack process integrity measure is the largest is indeed seen to be the one which is optimally resistant to the actual attack, i.e. the choice $d = a$ (for $d = 3, 4$ the optimally resistant degree distributions are very similar, and their behaviour is almost identical). Figure 5 suggests that, as long as the node removal cost functions are monotonically increasing with the node degree (i.e. for $a \geqslant 2$), networks which are optimally resistant against non-hub attacks ($a = 3, 4$) are reasonably resistant against hub-targeted attacks ($a = 2$), whereas networks which are well prepared against hub-targeted attacks behave quite poorly when subjected to non-hub attacks ($a = 3, 4$). In other words, degree statistics designed to be optimally resistant against hub removal appear to be quite sensitive to misinformation, whereas those optimally resistant to random attacks suffer less from misinformation, at least as long as the node removal cost function is monotonically increasing with the node degree.

6. Discussion

Many research papers have been devoted recently to the resilience of networks under attacks. Most study resilience in terms of the behaviour of static properties of networks under random and intelligent removal of a *fixed fraction* of sites or bonds. Results obtained empirically

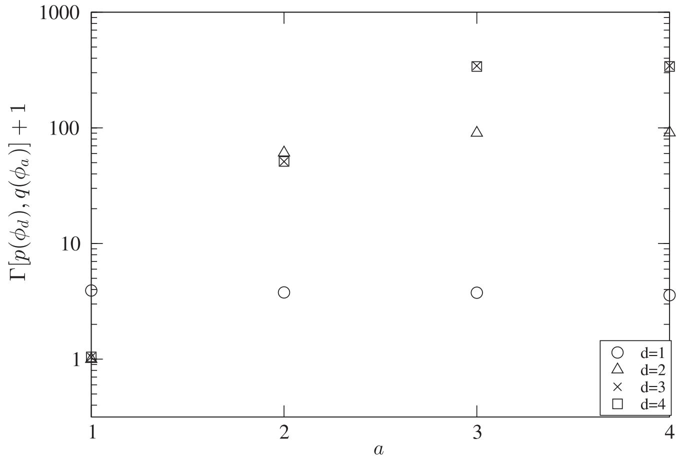

Figure 5. Log–plot of the post-attack integrity measure for the case where assumed and actual node removal costs and resource constraints need not be identical. Our network has the size and average connectivity of the *C. elegans* PIN. The node removal cost function is taken from a family $\phi_\ell(k)$ and the resource constraint is $\kappa_\ell^{-1} = \langle\phi_\ell(k)\rangle_\mathcal{P}/q_\ell$ with $\ell \in \{1, 2, 3, 4\}$ (see the main text for further details). The defender assumes that $\ell = d$ and chooses the associated optimally resistant degree distribution, whereas the actual value is $\ell = a$, and the attacker bases his strategy on the latter. The defender is optimally prepared only for $a = d$.

[1, 25] or analytically (within mean-field and asymptotic approximations) [2, 6, 7, 26, 27] have shown that power law networks are more resistant than Poissonian ones against random attacks (see e.g. [22, 23]), but are very vulnerable against hub removal [1, 24]. In contrast, more recent studies [3] suggest that power law networks are not more fragile than Poissonian graphs against hub-targeted attacks when one looks at the number of removed *links* (as opposed to nodes).

In this paper, we have sought to study network resilience in a more realistic setting, where attackers have *fixed resources* and where removing or disrupting a node carries a cost for the attacker which depends on the degree of the disrupted node. We quantify the resilience of the system in terms of the process for which the network acts as infrastructure, based on determining the critical temperatures for the onset of various types of global orders that could be envisaged (the resulting network integrity measure is only weakly dependent upon the specific choices made). This formulation also allows for attacks involving partial disruption of individual nodes, which would have been inaccessible to the techniques normally used when studying network resilience, such as percolation theory. We can define precisely the most damaging attack strategy, given knowledge of the degree sequence of a network, and for any given node-removal cost function. In addition we could subsequently define the optimal network topology, i.e. the degree distribution for which the integrity of the collective process is preserved best when attacked by a foe who employs the most damaging attack strategy.

A network's resilience against attacks is extremely sensitive to the dependence of the node removal cost function on the degree of the targeted node. This dependence determines the crucial outcome of the competition in such scenarios between the benefit and the cost of attacking high-degree nodes. If we choose a trivial constant cost function, we retrieve results from the literature on network resilience under random and targeted removal of fixed fractions of sites or bonds. However, as soon as one chooses more realistic node removal cost functions that increase sufficiently fast with the node's degree, power law networks are found to be more resistant than Poissonian ones, even against optimized intelligent attacks. Our results show that 'modular' configurations with a core of nodes highly connected to each other, in a sea of disconnected nodes, and strongly disassortative configurations, are, depending on the attacker's resources, the most resistant ones against hub-targeted attacks, respectively. Broad distributions with fat tails are the best defence against random and low degree targeted attacks. We also touched briefly upon the effects of misinformation, where a network is designed to be optimally resistant to a certain attack, whereas it actually faces a different one. Results suggest that for monotonically increasing cost functions, degree distributions with fat tails are much less sensitive to misinformation effects.

Upon comparing real protein interaction networks with random networks of the same size and average degree, we found that the attack resilience of the biological networks is superior to that of power law and Poissonian ones, even against optimized intelligent attacks. It may be that topological properties beyond the degree sequence play an important role here, and this deserves further investigation. In particular, one could calculate the integrity measure for processes supported by networks drawn from ensembles tailored to the production of graphs with built-in structure beyond that imposed by the degree distribution, along the lines of [18, 28]. Another direction for future work may be to consider graph ensembles in which both the network topologies and the node removal cost functions involve hidden variables.

Our paper emphasizes the importance of distinguishing between different classes of network attacks on the basis of the node removal cost function and resource limitations imposed upon the attacker, and of studying and quantifying network resilience strictly within a given class of attacks. Previously proposed conclusions about the vulnerability of power law networks against intelligent attacks should be moderated in all cases where there is no compelling reason to assume that the cost to attackers of node removal is independent of the node degrees.

Acknowledgment

One of the authors (ACCC) would like to thank the Engineering and Physical Sciences Research Council (UK) for support in the form of a Springboard Fellowship.

Appendix A. Equilibrium analysis for model A

Most of the derivations in both appendices follow the lines of similar calculations in e.g. [28, 29], and we will hence be brief and highlight only crucial steps to indicate the changes generated by the introduction of the attack variables $\{\xi_i\}$. Following [18] we use the property that with $\langle k\rangle = N^{-1}\sum_i k_i$ the ensemble (4) is identical to

$$\text{Prob}(\mathbf{c}) = \frac{\delta_{\mathbf{k},\mathbf{k}(\mathbf{c})}}{\mathcal{Z}} \prod_{i<j}\left[\frac{\langle k\rangle}{N}\delta_{c_{ij},1} + \left(1-\frac{\langle k\rangle}{N}\right)\delta_{c_{ij},0}\right] \tag{A.1}$$

$$\mathcal{Z} = \sum_{\mathbf{c}} \delta_{\mathbf{k},\mathbf{k}(\mathbf{c})} \prod_{i<j}\left[\frac{\langle k\rangle}{N}\delta_{c_{ij},1} + \left(1-\frac{\langle k\rangle}{N}\right)\delta_{c_{ij},0}\right]. \tag{A.2}$$

A.1. Derivation of saddle-point equations

We write the Kronecker δ's of the degree constraints in the integral form, and we introduce the short-hands $\boldsymbol{\sigma}_i = (\sigma_i^1, \ldots, \sigma_i^n) \in \{-1, 1\}^n$ so that

$$\begin{aligned}\overline{f}_A &= \lim_{N\to\infty}\lim_{n\to 0}\frac{1}{\beta n N}\left\{\log \mathcal{Z} - \log \sum_{\boldsymbol{\sigma}_1\ldots\boldsymbol{\sigma}_N}\int_{-\pi}^{\pi}\prod_i\left[\frac{\mathrm{d}\omega_i}{2\pi}\,\mathrm{e}^{\mathrm{i}\omega_i k_i}\right]\right.\\ &\quad\left.\times\prod_{i<j}\left(1+\frac{\langle k\rangle}{N}\left[\int \mathrm{d}J\,P(J)\,\mathrm{e}^{\beta J\xi_i\xi_j\boldsymbol{\sigma}_i\cdot\boldsymbol{\sigma}_j-\mathrm{i}(\omega_i+\omega_j)}-1\right]\right)\right\}\\ &= \lim_{N\to\infty}\lim_{n\to 0}\frac{1}{\beta n N}\left\{\log \mathcal{Z} - \log \sum_{\boldsymbol{\sigma}_1\ldots\boldsymbol{\sigma}_N}\int_{-\pi}^{\pi}\prod_i\left[\frac{\mathrm{d}\omega_i}{2\pi}\,\mathrm{e}^{\mathrm{i}\omega_i k_i}\right]\right.\\ &\quad\left.\times\exp\left[\frac{\langle k\rangle}{2N}\sum_{ij}\left[\int \mathrm{d}J\,P(J)\,\mathrm{e}^{\beta J\xi_i\xi_j\boldsymbol{\sigma}_i\cdot\boldsymbol{\sigma}_j-\mathrm{i}(\omega_i+\omega_j)}-1\right]+\mathcal{O}(N^0)\right]\right\}. \qquad \text{(A.3)}\end{aligned}$$

We proceed by introducing for $\boldsymbol{\sigma} \in \{-1, 1\}^n$ and $\xi \in \Xi$ the functions $D(\xi, \boldsymbol{\sigma}|\{\boldsymbol{\sigma}_i, \omega_i, \xi_i\}) = N^{-1}\sum_i \delta_{\xi,\xi_i}\delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}_i}\mathrm{e}^{-\mathrm{i}\omega_i}$. They are introduced via the substitution of integrals over appropriate δ-distributions, written in the integral form:

$$1 = \int \frac{\mathrm{d}D(\xi,\boldsymbol{\sigma})\,\mathrm{d}\hat{D}(\xi,\boldsymbol{\sigma})}{2\pi/N}\,\mathrm{e}^{\mathrm{i}N\hat{D}(\xi,\boldsymbol{\sigma})[D(\xi,\boldsymbol{\sigma})-D(\xi,\boldsymbol{\sigma}|\{\boldsymbol{\sigma}_i,\omega_i,\xi_i\})]}. \qquad \text{(A.4)}$$

Upon using the short hand $\{\mathrm{d}D\,\mathrm{d}\hat{D}\} = \prod_{\xi,\boldsymbol{\sigma}} D(\xi,\boldsymbol{\sigma})\,\mathrm{d}\hat{D}(\xi,\boldsymbol{\sigma})$ we then obtain

$$\begin{aligned}\overline{f}_A &= \lim_{N\to\infty}\lim_{n\to 0}\frac{1}{\beta n N}\left\{\log\mathcal{Z} - \log\int\{\mathrm{d}D\,\mathrm{d}\hat{D}\}\,\mathrm{e}^{\mathrm{i}N\sum_{\xi\boldsymbol{\sigma}}\hat{D}(\xi,\boldsymbol{\sigma})D(\xi,\boldsymbol{\sigma})-\frac{1}{2}N\langle k\rangle+\mathcal{O}(\log N)}\right.\\ &\quad\times\exp\left[\frac{1}{2}\langle k\rangle N\sum_{\xi\xi'}\sum_{\boldsymbol{\sigma}\boldsymbol{\sigma}'}D(\xi,\boldsymbol{\sigma})D(\xi',\boldsymbol{\sigma}')\int\mathrm{d}J\,P(J)\,\mathrm{e}^{\beta J\xi\xi'\boldsymbol{\sigma}\cdot\boldsymbol{\sigma}'}\right]\\ &\quad\left.\times\exp\left[N\sum_{\xi k}p(\xi,k)\log\sum_{\boldsymbol{\sigma}}\int_{-\pi}^{\pi}\frac{\mathrm{d}\omega}{2\pi}\,\mathrm{e}^{\mathrm{i}\omega k-\mathrm{i}\hat{D}(\xi,\boldsymbol{\sigma})\mathrm{e}^{-\mathrm{i}\omega}}\right]\right\}. \qquad \text{(A.5)}\end{aligned}$$

We next define $z = \lim_{N\to\infty} N^{-1}\log\mathcal{Z}$ (anticipating this limit to exist), which allows us to evaluate $\overline{f}$ by the steepest descent:

$$\overline{f}_A = \lim_{n\to 0}\frac{1}{n}\,\mathrm{extr}_{\{D,\hat{D}\}} f_{n,A}[\{D,\hat{D}\}] \qquad \text{(A.6)}$$

$$\begin{aligned}f_{n,A}[\ldots] = -\frac{1}{\beta}&\left\{\mathrm{i}\sum_{\xi\boldsymbol{\sigma}}\hat{D}(\xi,\boldsymbol{\sigma})D(\xi,\boldsymbol{\sigma}) - \frac{1}{2}\langle k\rangle - z\right.\\ &+\frac{1}{2}\langle k\rangle\sum_{\xi\xi'}\sum_{\boldsymbol{\sigma}\boldsymbol{\sigma}'}D(\xi,\boldsymbol{\sigma})D(\xi',\boldsymbol{\sigma}')\int\mathrm{d}J\,P(J)\,\mathrm{e}^{\beta J\xi\xi'\boldsymbol{\sigma}\cdot\boldsymbol{\sigma}'}\\ &\left.+\sum_{\xi k}p(\xi,k)\log\sum_{\boldsymbol{\sigma}}\int_{-\pi}^{\pi}\frac{\mathrm{d}\omega}{2\pi}\,\mathrm{e}^{\mathrm{i}\omega k-\mathrm{i}\hat{D}(\xi,\boldsymbol{\sigma})\mathrm{e}^{-\mathrm{i}\omega}}\right\}. \qquad \text{(A.7)}\end{aligned}$$

Extremization (A.7) with respect to $\{D, \hat{D}\}$ gives the saddle-point equations

$$\hat{D}(\xi,\boldsymbol{\sigma}) = \mathrm{i}\langle k\rangle\sum_{\xi'}\sum_{\boldsymbol{\sigma}'}D(\xi',\boldsymbol{\sigma}')\int\mathrm{d}J\,P(J)\,\mathrm{e}^{\beta J\xi\xi'\boldsymbol{\sigma}\cdot\boldsymbol{\sigma}'} \qquad \text{(A.8)}$$

$$D(\xi,\boldsymbol{\sigma}) = \sum_k p(\xi,k)\frac{\int_{-\pi}^{\pi}\mathrm{d}\omega\,\mathrm{e}^{\mathrm{i}\omega(k-1)-\mathrm{i}\hat{D}(\xi,\boldsymbol{\sigma})\mathrm{e}^{-\mathrm{i}\omega}}}{\sum_{\boldsymbol{\sigma}'}\int_{-\pi}^{\pi}\mathrm{d}\omega\,\mathrm{e}^{\mathrm{i}\omega k-\mathrm{i}\hat{D}(\xi,\boldsymbol{\sigma}')\mathrm{e}^{-\mathrm{i}\omega}}}. \tag{A.9}$$

The second of these equations is simplified using the identity

$$\int_{-\pi}^{\pi}\mathrm{d}\omega\,\mathrm{e}^{\mathrm{i}\omega\ell-\mathrm{i}\hat{D}(\xi,\boldsymbol{\sigma})\mathrm{e}^{-\mathrm{i}\omega}} = \begin{cases}2\pi[-\mathrm{i}\hat{D}(\xi,\boldsymbol{\sigma})]^\ell/\ell! & \text{if}\quad \ell\geqslant 0\\ 0 & \text{if}\quad \ell<0.\end{cases} \tag{A.10}$$

So, if we also re-define $\hat{D}(\xi,\boldsymbol{\sigma}) = \mathrm{i}\langle k\rangle F(\xi,\boldsymbol{\sigma})$, we arrive at

$$F(\xi,\boldsymbol{\sigma}) = \sum_{\xi'}\sum_{\boldsymbol{\sigma}'} D(\xi',\boldsymbol{\sigma}')\int\mathrm{d}J\,P(J)\,\mathrm{e}^{\beta J\xi\xi'\boldsymbol{\sigma}\cdot\boldsymbol{\sigma}'} \tag{A.11}$$

$$D(\xi,\boldsymbol{\sigma}) = \sum_{k>0} p(\xi,k)\frac{k}{\langle k\rangle}\frac{F^{k-1}(\xi,\boldsymbol{\sigma})}{\sum_{\boldsymbol{\sigma}'}F^k(\xi,\boldsymbol{\sigma}')}. \tag{A.12}$$

We note that $\sum_\xi\sum_{\boldsymbol{\sigma}} D(\xi,\boldsymbol{\sigma})F(\xi,\boldsymbol{\sigma}) = 1$ at the saddle-point. The term $z = \lim_{N\to\infty}N^{-1}\log\mathcal{Z}$ measures the number of graphs in the ensemble. It follows from $\lim_{\beta\to 0}(\beta\overline{f}) = -\log 2$, giving $z = \langle k\rangle\log\langle k\rangle - \langle k\rangle - \sum_k p(k)\log k!$, and hence

$$\overline{f}_A = -\lim_{n\to 0}\frac{1}{\beta n}\sum_{\xi k} p(\xi,k)\log\left[\sum_{\boldsymbol{\sigma}} F^k(\xi,\boldsymbol{\sigma})\right]. \tag{A.13}$$

A.2. Replica symmetric theory

To take the required limit $n\to 0$ in our formulae we make the replica-symmetric (RS) ansatz. The order parameter $D(\xi,\boldsymbol{\sigma})$ must now be invariant under all replica permutations, and thus have the following form:

$$D(\xi,\boldsymbol{\sigma}) = \int\mathrm{d}h\,D(\xi,h)\frac{\mathrm{e}^{\beta h\sum_\alpha\sigma_\alpha}}{[2\cosh(\beta h)]^n}. \tag{A.14}$$

Via equations (A.11), (A.12) one then finds a similar structure for $F(k,\boldsymbol{\sigma})$

$$F(\xi,\boldsymbol{\sigma}) = \int\mathrm{d}h\,F(\xi,h)\mathrm{e}^{\beta h\sum_\alpha\sigma_\alpha} \tag{A.15}$$

and in the limit $n\to 0$, after some standard manipulations, a closed set of transparent equations for the RS order parameters $D(\xi,h)$ and $F(\xi,h)$:

$$F(\xi,h) = \sum_{\xi'}\int\mathrm{d}h'\,\mathrm{d}J\,D(\xi',h')P(J)\delta\left[h-\frac{1}{\beta}\mathrm{atanh}[\tanh(\beta J\xi\xi')\tanh(\beta h')]\right] \tag{A.16}$$

$$D(\xi,h) = \sum_k p(\xi,k)\frac{k}{\langle k\rangle}\frac{\int\prod_{\ell<k}[\mathrm{d}h_\ell\,F(\xi,h_\ell)]\delta\big[h-\sum_{\ell<k}h_\ell\big]}{\big[\int\mathrm{d}h' F(\xi,h')\big]^k}. \tag{A.17}$$

We note upon integrating and combining these equations that $\int\mathrm{d}h\,F(\xi,h) = \sum_{\xi'}\int\mathrm{d}h\,D(\xi',h) = 1$. This enables us to write $F(\xi,h) = F(h|\xi)$ with $\int\mathrm{d}h\,F(h|\xi) = 1$, which gives immediate probabilistic interpretations of the functions $F(h|\xi)$. Upon eliminating $D(\xi,h)$ the RS saddle-point equations then take the new form

$$F(h|\xi) = \sum_{k\xi'} p(\xi',k)\frac{k}{\langle k\rangle}\int\mathrm{d}J\,P(J)\int\prod_{\ell<k}[\mathrm{d}h_\ell\,F(h_\ell|\xi')] \times\delta\left[h-\frac{1}{\beta}\mathrm{atanh}\left[\tanh(\beta J\xi\xi')\tanh\left(\beta\sum_{\ell<k}h_\ell\right)\right]\right]. \tag{A.18}$$

Clearly $F(h|0) = \delta(h)$. To identify the relevant observables and calculate for $\boldsymbol{\sigma} \in \{-1, 1\}^n$ the quantity $P(\xi, k, \boldsymbol{\sigma}) = \lim_{N\to\infty} N^{-1} \sum_i \overline{\langle \delta_{\xi,\xi_i} \delta_{k,k_i} \delta_{\boldsymbol{\sigma},\boldsymbol{\sigma}_i} \rangle}$ one uses the alternative form of the replica identity, namely

$$\overline{\langle g(\boldsymbol{\sigma})\rangle} = \overline{\left[\frac{\sum_{\boldsymbol{\sigma}} g(\boldsymbol{\sigma}) e^{-\beta H(\boldsymbol{\sigma})}}{\sum_{\boldsymbol{\sigma}} e^{-\beta H(\boldsymbol{\sigma})}}\right]} = \lim_{n\to 0} \sum_{\boldsymbol{\sigma}^1 \ldots \boldsymbol{\sigma}^n} \overline{g(\sigma^1) e^{-\beta \sum_{\alpha=1}^n H(\boldsymbol{\sigma}^\alpha)}}. \tag{A.19}$$

Upon also making the RS ansatz this results in

$$P_{\mathrm{RS}}(\xi, k, \boldsymbol{\sigma}) = p(\xi, k) \int \mathrm{d}h\, W(h|\xi, k) \frac{\mathrm{e}^{\beta h \sum_\alpha \sigma_\alpha}}{[2\cosh(\beta h)]^n} \tag{A.20}$$

$$W(h|\xi, k) = \int \prod_{\ell \leqslant k} [\mathrm{d}h_\ell\, F(h_\ell|\xi)] \delta\left[h - \sum_{\ell\leqslant k} h_\ell\right]. \tag{A.21}$$

The measure $W(h|\xi, k)$ is the effective field distribution for those sites where $(\xi_i, k_i) = (\xi, k)$. We note that $W(h|0, k) = \delta(h)$. With $W(h) = \sum_{\xi k} p(\xi, k) W(h|\xi, k)$ we can write the conventional scalar order parameters $m = \lim_{N\to\infty} N^{-1} \sum_i \overline{\langle \sigma_i \rangle}$ and $q = \lim_{N\to\infty} N^{-1} \sum_i \overline{\langle \sigma_i \rangle^2}$ in their familiar forms

$$m = \int \mathrm{d}h\, W(h) \tanh(\beta h), \qquad q = \int \mathrm{d}h\, W(h) \tanh^2(\beta h). \tag{A.22}$$

The subset of sites with $(\xi_i, k_i) = (\xi, k)$ can be regarded as sublattices in the sense of [30], and we can define sublattice magnetizations $m(\xi, k)$ via $m(\xi, k) = \int \mathrm{d}h\, W(h|\xi, k) \tanh(\beta h)$, such that $m = \sum_{\xi k} p(\xi, k) m(\xi, k)$. In the limit $T \to \infty$ (i.e. $\beta \to 0$) the only solution of (A.18) is as always the trivial paramagnetic (P) one: $F(h|\xi) = \delta(h)$. This is a saddle-point at any temperature, but can become unstable in favour of ferromagnetic (F) or spin-glass (SG) states as T is lowered.

A.3. Continuous phase transitions away from the paramagnetic state

Continuous bifurcations away from the trivial state are found as usual by expanding (A.18) in moments of $F(h|\xi)$, assuming the existence of a small parameter ϵ with $0 < |\epsilon| \ll 1$ such that $\int \mathrm{d}h h^\ell F(h|\xi) = \mathcal{O}(\epsilon^\ell)$. With some foresight we define a function $\gamma(\xi)$ and two $|\Xi|\times|\Xi|$ matrices $M^{(\ell)}(\beta)$ with entries $M^{(\ell)}_{\xi\xi'}(\beta)$ for $\ell \in \{1, 2\}$:

$$\gamma(\xi) = \langle k\rangle^{-1} \sum_k p(\xi, k) k(k-1) \tag{A.23}$$

$$M^{(\ell)}_{\xi\xi'}(\beta) = \gamma(\xi') \int \mathrm{d}J\, P(J) \tanh^\ell(\beta J \xi\xi'). \tag{A.24}$$

We define $\lambda^{(\ell)}_{\max}(\beta)$ as the largest eigenvalue of $M^{(\ell)}(\beta)$. If the first order to bifurcate away from $F(h|\xi) = \delta(h)$ is ϵ^1, the bifurcation is towards a state where $m \neq 0$, i.e. describing a P→F transition. Upon multiplying both sides of (A.18) by h and integrating over h, the bifurcation condition for this is found to be

$$\mathrm{P} \to \mathrm{F}: \quad \lambda^{(1)}_{\max}(\beta) = 1. \tag{A.25}$$

If instead the first order to bifurcate is ϵ^2, the bifurcating new state has $m = 0$ and $q > 0$, describing a P→SG transition. Upon multiplying both sides of (A.18) by h^2 and integrating over h, the bifurcation condition for this is found to be

$$\mathrm{P} \to \mathrm{SG}: \quad \lambda^{(2)}_{\max}(\beta) = 1. \tag{A.26}$$

We focus on a specific simple bond distribution, the binary $P(J) = \frac{1}{2}(1+\eta)\delta(J-J_0) + \frac{1}{2}(1-\eta)\delta(J+J_0)$ (with $J_0 \geqslant 0$), where the matrices $M^{(\ell)}(\beta)$ take the simple form:

$$M^{(1)}_{\xi\xi'}(\beta) = \eta \tanh(\beta J_0 \xi\xi')\gamma(\xi') \qquad M^{(2)}_{\xi\xi'}(\beta) = \tanh^2(\beta J_0\xi\xi')\gamma(\xi'). \tag{A.27}$$

Appendix B. Equilibrium analysis for model B

B.1. Derivation of saddle-point equations

The calculation for coupled oscillators is initially very similar to the previous one, with summations replaced by integrations. The main differences start at the introduction of the replica-symmetry ansatz; from then onwards, we have to implement appropriate adaptations of the calculation for XY spins in [29] (an alternative route would be to adapt the cavity-based analysis in [31]). As before we write degree constraints in the integral form, and we introduce the short-hands $\boldsymbol{\theta}_i = (\theta_i^1, \ldots, \theta_i^n) \in [-\pi, \pi]^n$ so that

$$\overline{f}_B = \lim_{N\to\infty}\lim_{n\to 0}\frac{1}{\beta n N}\left\{\log\mathcal{Z} - \log\int_{-\pi}^{\pi} \mathrm{d}\boldsymbol{\theta}_1\ldots\mathrm{d}\boldsymbol{\theta}_N\int_{-\pi}^{\pi}\prod_i\left[\frac{\mathrm{d}\omega_i}{2\pi}\,\mathrm{e}^{\mathrm{i}\omega_i k_i}\right]\right.$$
$$\left.\times\exp\left[\frac{\langle k\rangle}{2N}\sum_{ij}\left[\int\mathrm{d}J\,P(J)\,\mathrm{e}^{\beta J\xi_i\xi_j\sum_\alpha\cos(\theta_i^\alpha-\theta_j^\alpha)-\mathrm{i}(\omega_i+\omega_j)} - 1\right] + \mathcal{O}(N^0)\right]\right\}. \tag{B.1}$$

We next introduce for $\boldsymbol{\theta} \in [-\pi,\pi]^n$ and $\xi \in \{0,1\}$ the functions $D(\xi,\boldsymbol{\theta}|\{\boldsymbol{\theta}_i,\omega_i,\xi_i\}) = N^{-1}\sum_i \delta_{\xi,\xi_i}\delta[\boldsymbol{\theta},\boldsymbol{\theta}_i]\,\mathrm{e}^{-\mathrm{i}\omega_i}$, via the substitution of functional integrals over appropriate δ-distributions, written in the integral form. With the short hand $\{\mathrm{d}D\,\mathrm{d}\hat{D}\} = \prod_{\xi,\boldsymbol{\theta}} D(\xi,\boldsymbol{\theta})\,\mathrm{d}\hat{D}(\xi,\boldsymbol{\theta})$, we then obtain an expression in the form of the path integral:

$$\overline{f}_B = \lim_{N\to\infty}\lim_{n\to 0}\frac{1}{\beta n N}\left\{\log\mathcal{Z} - \log\int\{\mathrm{d}D\,\mathrm{d}\hat{D}\}\,\mathrm{e}^{\mathrm{i}N\sum_\xi\int\mathrm{d}\boldsymbol{\theta}\hat{D}(\xi,\boldsymbol{\theta})D(\xi,\boldsymbol{\theta})-\frac{1}{2}N\langle k\rangle+\mathcal{O}(\log N)}\right.$$
$$\times\exp\left[\frac{1}{2}\langle k\rangle N\sum_{\xi\xi'}\int\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{\theta}'\,D(\xi,\boldsymbol{\theta})D(\xi',\boldsymbol{\theta}')\int\mathrm{d}J\,P(J)\,\mathrm{e}^{\beta J\xi\xi'\sum_\alpha\cos(\theta_\alpha-\theta'_\alpha)}\right]$$
$$\left.\times\exp\left[N\sum_{\xi k}p(\xi,k)\log\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\theta}\int_{-\pi}^{\pi}\frac{\mathrm{d}\omega}{2\pi}\,\mathrm{e}^{\mathrm{i}\omega k-\mathrm{i}\hat{D}(\xi,\boldsymbol{\theta})\mathrm{e}^{-\mathrm{i}\omega}}\right]\right\}. \tag{B.2}$$

With $z = \lim_{N\to\infty} N^{-1}\log\mathcal{Z}_N = \langle k\rangle\log\langle k\rangle - \langle k\rangle - \sum_k p(k)\log k!$ (which has already been calculated earlier), we evaluate $\overline{f}$ by the steepest descent:

$$\overline{f}_B = \lim_{n\to 0}\frac{1}{n}\,\mathrm{extr}_{\{D,\hat{D}\}} f_{n,B}[\{D,\hat{D}\}] \tag{B.3}$$

$$f_{n,B}[\ldots] = -\frac{1}{\beta}\left\{\mathrm{i}\sum_\xi\int\mathrm{d}\boldsymbol{\theta}\hat{D}(\xi,\boldsymbol{\theta})D(\xi,\boldsymbol{\theta}) - \frac{1}{2}\langle k\rangle - z\right.$$
$$+\frac{1}{2}\langle k\rangle\sum_{\xi\xi'}\int\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\boldsymbol{\theta}'\,D(\xi,\boldsymbol{\theta})D(\xi',\boldsymbol{\theta}')\int\mathrm{d}J\,P(J)\,\mathrm{e}^{\beta J\xi\xi'\sum_\alpha\cos(\theta_\alpha-\theta'_\alpha)}$$
$$\left.+\sum_{\xi k}p(\xi,k)\log\int_{-\pi}^{\pi}\mathrm{d}\boldsymbol{\theta}\int_{-\pi}^{\pi}\frac{\mathrm{d}\omega}{2\pi}\,\mathrm{e}^{\mathrm{i}\omega k-\mathrm{i}\hat{D}(\xi,\boldsymbol{\theta})\mathrm{e}^{-\mathrm{i}\omega}}\right\}. \tag{B.4}$$

Functional variation of (B.3) with respect to $\{D, \hat{D}\}$, followed by application of (A.10) and transformation via $\hat{D}(\xi,\theta) = \mathrm{i}\langle k\rangle F(\xi,\boldsymbol{\theta})$, gives the saddle-point equations

$$F(\xi,\boldsymbol{\theta}) = \sum_{\xi'} \int \mathrm{d}\boldsymbol{\theta}' D(\xi',\boldsymbol{\theta}') \int \mathrm{d}J\, P(J)\, \mathrm{e}^{\beta J\xi\xi' \sum_\alpha \cos(\theta_\alpha-\theta'_\alpha)} \tag{B.5}$$

$$D(\xi,\boldsymbol{\theta}) = \sum_{k>0} p(\xi,k) \frac{k}{\langle k\rangle} \frac{F^{k-1}(\xi,\boldsymbol{\theta})}{\int \mathrm{d}\boldsymbol{\theta}' F^k(\xi,\boldsymbol{\theta}')}. \tag{B.6}$$

Again $\sum_\xi \int \mathrm{d}\boldsymbol{\theta}\, D(\xi,\boldsymbol{\theta}) F(\xi,\boldsymbol{\theta}) = 1$ at the saddle-point, and we obtain

$$\overline{f}_B = -\lim_{n\to 0} \frac{1}{\beta n} \sum_{\xi k} p(\xi,k) \log\left[\int \mathrm{d}\boldsymbol{\theta} F^k(\xi,\boldsymbol{\theta})\right]. \tag{B.7}$$

B.2. Replica symmetric theory

For real-valued variables the replica-symmetric ansatz is less straightforward. Permutation invariance with respect to $\boldsymbol{\theta}$ components now implies that $D(\xi,\boldsymbol{\theta})$ and $F(\xi,\boldsymbol{\theta})$ are functional integrals over the space of normalized functions $P : [-\pi,\pi] \to \mathbb{R}$ (i.e. $\int_{-\pi}^{\pi} \mathrm{d}\theta\, P(\theta) = 1$), with functional measures $W_D[\xi,\{P\}]$ and $W_F[\xi,\{P\}]$:

$$D(\xi,\boldsymbol{\theta}) = \int \{\mathrm{d}P\} W_D[\xi,\{P\}] \prod_\alpha P(\theta_\alpha) \tag{B.8}$$

$$F(\xi,\boldsymbol{\theta}) = \int \{\mathrm{d}P\} W_F[\xi,\{P\}] \prod_\alpha P(\theta_\alpha) \tag{B.9}$$

(we may use the same symbol P as employed to define the bond probabilities via $P(J)$; the arguments will always prevent ambiguity). Insertion of (B.8), (B.9) into the two equations (B.5), (B.6) then gives, in the limit $n \to 0$ and after some manipulations, the following closed equations for the RS measures $W_D[\xi,\{P\}]$ and $W_F[\xi,\{P\}]$:

$$W_F[\xi,\{P\}] = \sum_{\xi'} \int \{\mathrm{d}P'\} W_D[\xi',\{P'\}] \int \mathrm{d}J\, P(J) \times \prod_\theta \delta\left[P(\theta) - \frac{\int \mathrm{d}\theta' \mathrm{e}^{\beta J\xi\xi'\cos(\theta-\theta')} P'(\theta')}{2\pi I_0(\beta J\xi\xi')}\right] \tag{B.10}$$

$$W_D[\xi,\{P\}] = \sum_{k>0} \frac{p(\xi,k)k/\langle k\rangle}{\left[\int \{\mathrm{d}P'\} W_F[\xi,\{P'\}]\right]^k} \int \prod_{\ell<k} [\{\mathrm{d}P_\ell\} W_F[\xi,\{P_\ell\}]] \times \prod_\theta \delta\left[P(\theta) - \frac{\prod_{\ell<k} P_\ell(\theta)}{\int \mathrm{d}\theta' \prod_{\ell<k} P_\ell(\theta')}\right]. \tag{B.11}$$

Functional integration of both equations over P shows that $\int \{\mathrm{d}P\} W_F[\xi,\{P\}] = \sum_{\xi'} \int \{\mathrm{d}P\} W_D[\xi',\{P\}] = 1$. This allows us to write $W_F[\xi,\{P\}] = W_F[\{P\}|\xi]$ with $\int \{\mathrm{d}P\} W_F[\{P\}|\xi] = 1$, which also allows here for probabilistic interpretations of the order parameters $W_F[\xi,\{P\}]$, which are now functionals acting on the space of probability distributions over the interval $[-\pi,\pi]$. Upon eliminating $W_D[\xi,\{P\}]$, the RS saddle-point equations then take the following form (where $\theta, \theta' \in [-\pi, \pi]$):

$$W_F[\{P\}|\xi] = \sum_{k\xi'} p(\xi', k)\frac{k}{\langle k\rangle}\int \mathrm{d}J\, P(J)\int \prod_{\ell<k}[\{\mathrm{d}P_\ell\}W_F[\{P_\ell\}|\xi']]$$
$$\times \prod_\theta \delta\left[P(\theta) - \frac{\int \mathrm{d}\theta' \mathrm{e}^{\beta J\xi\xi'\cos(\theta-\theta')}\prod_{\ell<k}P_\ell(\theta')}{2\pi I_0(\beta J\xi\xi')\int \mathrm{d}\theta' \prod_{\ell<k}P_\ell(\theta')}\right]. \tag{B.12}$$

We observe that $W_F[\{P\}|0] = \prod_\theta \delta[P(\theta) - (2\pi)^{-1}]$. To identify the physical meaning of our observables we define and calculate the quantity $P(\xi, k, \boldsymbol{\theta}) = \lim_{N\to\infty} N^{-1}\sum_i \overline{\langle \delta_{\xi,\xi_i}\delta_{k,k_i}\delta[\boldsymbol{\theta}, \boldsymbol{\theta}_i]\rangle}$. Within the RS ansatz it is found to be

$$P_{\mathrm{RS}}(\xi, k, \boldsymbol{\theta}) = p(\xi, k)\int \{\mathrm{d}P\}W[\{P\}|\xi, k]\prod_\alpha P(\theta_\alpha) \tag{B.13}$$

$$W[\{P\}|\xi, k] = \int \prod_{\ell\leqslant k}[\{\mathrm{d}P_\ell\}W_F[\{P_\ell\}|\xi]]\prod_\theta \delta\left[P(\theta) - \frac{\prod_{\ell\leqslant k}P_\ell(\theta)}{\int \mathrm{d}\theta' \prod_{\ell\leqslant k}P_\ell(\theta')}\right]. \tag{B.14}$$

The functional measure $W[\{P\}|\xi, k]$ generalizes the concept of an effective field to an 'effective' angle distribution of those oscillators with $(\xi_i, k_i) = (\xi, k)$. Note that $W[\{P\}|0, k] = \prod_\theta \delta[P(\theta) - (2\pi)^{-1}]$. With $W[\{P\}] = \sum_{\xi k} p(\xi, k)W[\{P\}|\xi, k]$ we can write the conventional types of scalar-order parameters in a compact form:

$$\lim_{N\to\infty}\frac{1}{N}\sum_i \overline{\langle f(\theta_i)\rangle} = \int \{\mathrm{d}P\}W[\{P\}]\int_{-\pi}^{\pi}\mathrm{d}\theta\, P(\theta)f(\theta) \tag{B.15}$$

$$\lim_{N\to\infty}\frac{1}{N}\sum_i \overline{\langle f(\theta_i)\rangle\langle g(\theta_i)\rangle} = \int \{\mathrm{d}P\}W[\{P\}]\left[\int_{-\pi}^{\pi}\mathrm{d}\theta\, P(\theta)f(\theta)\right]\left[\int_{-\pi}^{\pi}\mathrm{d}\theta\, P(\theta)g(\theta)\right]. \tag{B.16}$$

For $T \to \infty$ (i.e. $\beta \to 0$) the only solution of our equations is the trivial P state of fully random phases θ_i: $W_F[\{P\}|\xi] = W[\{P\}|\xi, k] = W[\{P\}] = \prod_\theta \delta[P(\theta) - (2\pi)^{-1}]$, which solves out equations at any temperature, but will destabilize at some T in favour of ordered states with (partially) frozen relations between the phases of the oscillators.

B.3. Continuous phase transitions away from the incoherent state

To find continuous bifurcations away from the incoherent (P) state one has to carry out a Guzai expansion [29] of the functional order parameter equations (B.12) around the solution $W_F[\{P\}|\xi] = \prod_\theta \delta[P(\theta) - (2\pi)^{-1}]$. This will involve the modified Bessel functions $I_m(z)$ [32]. One writes $P(\theta) = (2\pi)^{-1} + \Delta(\theta)$ and $W_F[\{P\}|\xi] \to \tilde{W}[\{\Delta\}|\xi]$, with $\tilde{W}[\{\Delta\}|\xi] = 0$ as soon as $\int_{-\pi}^{\pi}\mathrm{d}\theta\,\Delta(\theta) \neq 0$ and one expands (B.12) in $\Delta(\theta)$:

$$\tilde{W}[\{\Delta\}|\xi] = \sum_{k\xi'} p(\xi', k)\frac{k}{\langle k\rangle}\int \mathrm{d}J\, P(J)\int \prod_{\ell<k}\left[\{\mathrm{d}\Delta_\ell\}\tilde{W}[\{\Delta_\ell\}|\xi']\right]$$
$$\times \prod_\theta \delta\left[\Delta(\theta) - \frac{1}{2\pi I_0(\beta J\xi\xi')}\sum_{\ell<k}\int \mathrm{d}\theta' \mathrm{e}^{\beta J\xi\xi'\cos(\theta-\theta')}\Delta_\ell(\theta')\right.$$
$$\left. - \frac{1}{2}\sum_{\ell\neq\ell'}^{k-1}\int \mathrm{d}\theta'\left(\frac{\mathrm{e}^{\beta J\xi\xi'\cos(\theta-\theta')}}{I_0(\beta J\xi\xi')} - 1\right)\Delta_\ell(\theta')\Delta_{\ell'}(\theta') + \mathcal{O}(\Delta^3)\right]. \tag{B.17}$$

We next evaluate functional moments of both sides of this equation. If the first bifurcation away from the P state is of order Δ, we multiply by $\Delta(\theta)$ and integrate (functionally) over all

Δ, leading to an eigenvalue problem for the functions $\Psi_\xi(\theta) = \int\{\mathrm{d}\Delta\}\tilde{W}[\{\Delta\}|\xi]\Delta(\theta)$ subject to the constraint $\int_{-\pi}^{\pi}\mathrm{d}\theta\,\Psi_\xi(\theta) = 0$:

$$\Psi_\xi(\theta) = \sum_{\xi'}\gamma(\xi')\int\frac{\mathrm{d}J\,P(J)}{I_0(\beta J\xi\xi')}\int_{-\pi}^{\pi}\frac{\mathrm{d}\theta'}{2\pi}\,\mathrm{e}^{\beta J\xi\xi'\cos(\theta-\theta')}\Psi_{\xi'}(\theta') \qquad \text{(B.18)}$$

with $\gamma(\xi)$ as defined in (A.23). The solutions are of the form $\Psi_\xi(\theta) = \psi(\xi)\mathrm{e}^{\mathrm{i}m\theta}$, with $m \in \{1, 2, 3, \ldots\}$ and with $\psi(\xi)$ to be solved from the eigenvalue equation

$$\mathcal{O}(\Delta)\text{ bifurcations:}\quad \psi(\xi) = \sum_{\xi'}\left(\int\mathrm{d}J\,P(J)\frac{I_m(\beta J\xi\xi')}{I_0(\beta J\xi\xi')}\right)\gamma(\xi')\psi(\xi'). \qquad \text{(B.19)}$$

For $m = 1$ the bifurcating state (F) is one where the oscillators synchronize (partly) to a preferred overall phase, whereas for $m > 1$ the transition is towards a state with non-uniform phase statistics but without global synchronization [29].

If the first bifurcation away from the P state is of order Δ^2 rather than Δ, so $\int\{\mathrm{d}\Delta\}\tilde{W}[\{\Delta\}|\xi]\Delta(\theta) = 0$, we multiply (B.17) by $\Delta(\theta_1)\Delta(\theta_2)$ and integrate over all functions Δ, leading to an eigenvalue problem for the function $\Psi_\xi(\theta_1,\theta_2) = \int\{\mathrm{d}\Delta\}\tilde{W}[\{\Delta\}|\xi]\Delta(\theta_1)\Delta(\theta_2)$ subject to $\int_{-\pi}^{\pi}\mathrm{d}\theta_1\,\Psi_\xi(\theta_1,\theta_2) = \int_{-\pi}^{\pi}\mathrm{d}\theta_2\,\Psi_\xi(\theta_1,\theta_2) = 0$:

$$\Psi_\xi(\theta_1,\theta_2) = \sum_{\xi'}\gamma(\xi')\int\frac{\mathrm{d}J\;P(J)}{I_0^2(\beta J\xi\xi')}\int_{-\pi}^{\pi}\frac{\mathrm{d}\theta_1'\,\mathrm{d}\theta_2'}{4\pi^2}\,\mathrm{e}^{\beta J\xi\xi'[\cos(\theta_1-\theta_1')+\cos(\theta_2-\theta_2')]}\Psi_{\xi'}(\theta_1',\theta_2'). \qquad \text{(B.20)}$$

The solutions are of the form $\Psi_\xi(\theta_1,\theta_2) = \psi(\xi)\,\mathrm{e}^{\mathrm{i}(m_1\theta_1+m_2\theta_2)}$ with $m_{1,2} \in \{1, 2, 3, \ldots\}$, representing new states with 'frozen' local phase ordering but no global synchronization, i.e. spin-glass-type states (SG), each bifurcating when

$$\mathcal{O}(\Delta^2)\text{ bifurcations:}\quad \psi(\xi) = \sum_{\xi'}\left(\int\mathrm{d}J\;P(J)\frac{I_{m_1}(\beta J\xi\xi')I_{m_2}(\beta J\xi\xi')}{I_0^2(\beta J\xi\xi')}\right)\gamma(\xi')\psi(\xi'). \qquad \text{(B.21)}$$

The right-hand sides of both (B.19) and (B.21) vanish at $\beta = 0$, so the transitions correspond to the smallest β such that solutions of (B.19) and (B.21) exist. Hence, we need the maxima of the right-hand sides over m and (m_1, m_2), respectively. The properties of the modified Bessel functions (see e.g. [29]) ensure that these maxima are found for $m = 1$ and $(m_1, m_2) = (1, 1)$. Finally, if we again choose the bond distribution $P(J) = \frac{1}{2}(1+\eta)\delta(J-J_0) + \frac{1}{2}(1-\eta)\delta(J+J_0)$, the bifurcation conditions can once more be written in the form (A.25), (A.26), but where in the case of coupled oscillators the largest eigenvalues $\lambda_{\max}^{(1)}$ and $\lambda_{\max}^{(2)}$ refer to the following matrices:

$$M_{\xi\xi'}^{(1)}(\beta) = \eta\frac{I_1(\beta J_0\xi\xi')}{I_0(\beta J\xi\xi')}\gamma(\xi'), \qquad M_{\xi\xi'}^{(2)}(\beta) = \frac{I_1^2(\beta J_0\xi\xi')}{I_0^2(\beta J_0\xi\xi')}\gamma(\xi'). \qquad \text{(B.22)}$$

Comparison with (A.27) shows that, inasmuch as the location of the transition lines away from the P state is concerned, the differences between having interacting Ising spins or coupled oscillators on the nodes of the network are accounted for by the simple substitution $\tanh(z) \to I_1(z)/I_0(z)$ in the relevant remaining eigenvalue problem.

References

[1] Albert R, Jeong H and Barabási A L 2000 *Nature* **406** 378
[2] Cohen R, Erez K, Ben-Avraham D and Havlin S 2001 *Phys. Rev. Lett.* **86** 3682
[3] Magnien C, Latapy M and Guillaume J L 2009 arXiv:cond-mat/0908.3154

[4] Albert R and Barabási A L 2002 *Rev. Mod. Phys.* **74** 47–97
[5] Newman M E J, Strogatz S H and Watts D J 2001 *Phys. Rev.* E **64** 026118
[6] Callaway D S, Newman M E J, Strogatz S H and Watts D J 2001 *Phys. Rev. Lett.* **85** 5468
[7] Cohen R, Erez K, Ben-Avraham D and Havlin S 2000 *Phys. Rev. Lett.* **85** 4626
[8] Hase M O and Mendes J F F 2008 *J. Phys. A: Math. Theor.* **41** 145002
[9] Pastor-Satorras R and Vespignani A 2001 *Phys. Rev. Lett.* **87** 3200
[10] Pastor-Satorras R and Vespignani A 2001 *Phys. Rev.* E **63** 066117
[11] Lloyd A L and May R M 2001 *Science* **292** 1316
[12] Simonis N *et al* 2008 *Nature Methods* **6** 47
[13] Parrish J R *et al* 2007 *Genome Biol.* **8** R131
[14] Arifuzzaman M *et al* 2006 *Genome Res.* **16** 686–91
[15] Prasad T S *et al* 2009 *Nucleic Acid Res.* **37** D767
[16] Ito T *et al* 2001 *Proc. Natl Acad. Sci. USA* **98** 4569
[17] Kuehn R, Lin Y-C and Poeppel G 1998 arXiv:cond-mat/9805137
[18] Annibale A, Coolen A C C, Fernandes L P, Kleinjung J and Fraternali F 2009 *J. Phys. A: Math. Theor.* **42** 485001
[19] Bianconi G, Coolen A C C and Pérez-Vicente C J 2008 *Phys. Rev.* E **78** 016114
[20] Newman M E J 2002 *Phys. Rev. Lett.* **89** 208701
[21] Paul G, Tanizawa T, Havlin S and Stanley H E 2004 *Eur. Phys. J.* B **38** 187–91
[22] Dorogovtsev S N and Mendes J F F 2003 *Evolution of Networks: From Biological Networks to the Internet and www* (Oxford: Oxford University Press)
[23] Barabási A L 2002 *Linked: The New Science of Networks* (Cambridge, MA: Perseus)
[24] Barabási A L 2003 *Emergence of Scaling in Complex Networks* (Weinheim: Wiley)
[25] Broder A Z *et al* 2000 *Comput. Netw.* **33** 309–20
[26] Cohen R, Havlin S and Ben-Avraham D 2003 *Structural Properties of Scale Free Networks* (Weinheim: Wiley)
[27] Newman M E J 2003 *Random Graphs as Models of Networks* in Handbook of graphs and networks: from the genome to the internet (Weinheim: Wiley)
[28] Pérez-Vicente C J and Coolen A C C 2008 *J. Phys. A: Math. Theor.* **41** 255003
[29] Coolen A C C, Skantzos N S, Pérez Castillo I, Pérez Vicente C J, Hatchett J P L, Wemmenhove B and Nikoletopoulos T 2005 *J. Phys. A: Math. Gen.* **38** 8289–317
[30] Wemmenhove B and Coolen A C C 2003 *J. Phys. A: Math. Gen.* **36** 9617
[31] Skantzos S, Pérez Castillo I and Hatchett J P L 2005 *Phys. Rev.* E **72** 066127
[32] Menzel D H 1960 *Fundamental Formulas of Physics* vol 1 (New York: Dover)