# ADVANCED TOPICS 5.A DERIVING THE DEGREE DISTRIBUTION

A number of analytical techniques are available to calculate the exact form of the degree exponent provided in Eq. 5.11. Next we derive it using the rate equation approach [12, 13]. The method is rather general, allowing us to explore the properties of a wide range of growing networks. Consequently, the calculations described here will be of direct relevance for many systems, from the WWW to protein interaction networks. Let us denote with $N(k, t)$ the number of nodes with degree $k$ at time $t$. The degree distribution $p_k(t)$ relates to this quantity via $p_k(t) = N(k,t)/N$ .

Since at each time-step we add a new node to the network, we have $N = t$. That is, at any moment the total number of nodes equals the number of timesteps BOX 5.3. We write preferential attachment as

$$\Pi(k) = \frac{k}{\sum_j k_j} = \frac{k}{2mt} \cdot \tag{5.31}$$

where the $2m$ term captures the fact that that in an undirected network each link contributes to the degree of two nodes. Our goal is to calculate the changes in the number of nodes with degree $k$ after a new node is added to the network. For this we inspect the two events that alter $N(k, t)$ (and hence $p_k(t)$) following the arrival of a new node:

**(i)** A new node can link to a degree-$k$ node, turning it into a degree ($k$+1 node, hence decreasing $N(k, t)$.

**(ii)** A new node can link to a degree ($k$-1) node, turning it into a degree $k$ node, hence increasing $N(k, t)$.

The number of links that are expected to connect to degree $k$ nodes after the arrival of a new node is

$$\frac{k}{2mt} \times N p_k(t) \times m = \frac{k}{2} p_k(t), \tag{5.32}$$

where the first term captures the probability that the new node will link to a degree-$k$ node (preferential attachment); the second term provides the

total number of nodes with degree $k$, as the more nodes are in this category, the more likely that a new node will attach to one of them; the third term is simply the degree of the incoming node, as the higher $m$, the higher the chance that the new node will link to a degree-$k$ node. We next apply Eq. 5.32 to cases (i) and (ii) above:

**(i')** The number of degree $k$ nodes that acquire a new link becoming ($k$+1) degree nodes, is

$$\frac{k}{2} p_k(t) \tag{5.33}$$

**(ii')** The number of degree ($k$-1) nodes that acquire a new link, increasing their degree to $k$ is

$$\frac{k-1}{2} p_{k-1}(t). \tag{5.34}$$

Combining Eq. 5.32 and Eq. 5.33 we obtain the expected number of degree-$k$ nodes after the addition of a new node

$$(N+1)p_k(t+1) = Np_k(t) + \frac{k-1}{2} p_{k-1}(t) - \frac{k}{2} p_k(t). \tag{5.35}$$

This equation applies to all nodes with degree $k>m$. As we lack nodes with degree $k$=0,1, ... , $m$-1 in the network (each new node arrives with degree $m$) we need a separate equation for degree $m$ modes. Following the arguments we used to derive Eq. 5.35, we obtain

$$(N+1)p_m(t+1) = Np_m(t) + 1 - \frac{m}{2} p_m(t) . \tag{5.36}$$

Eq. 5.35 and 5.36 are the starting point of the recursive process that provides $p_k$. Let us use the fact that we are looking for a stationary degree distribution, supported by numerical simulations Fig. 5.6. This means that in the $N = t \to \infty$ limit, $p_k(\infty) = p_k$. Using this we can write the l.h.s. of Eq. 5.35 and 5.36 as $(N+1)p_k(t+1) - Np_k(t) \to Np_k(\infty) + p_k(\infty) - Np_k(\infty) = p_k(\infty) = p_k$ $(N+1)\, p_{mk}(t+1) - Np_m(t) \to p_m$. Therefore the rate equations Eq. 5.35 and 5.36 take the form:

$$p_k = \frac{k-1}{k+2} p_{k-1} \qquad k > m \tag{5.37}$$

$$p_m = \frac{2}{m+2} \tag{5.38}$$

Note that Fig. 5.37 can be rewritten as

$$p_{k+1} = \frac{k}{k+3} p_k \tag{5.39}$$

via a $k \to k+1$ variable change. To obtain the degree distribution, we use a recursive approach. That is, we write the degree distribution for the smallest degree, $k=m$, using Eq.. 5.38 and then use Eq. 5.39 to calculate $p_k$ for the higher degrees:

$$\begin{aligned} p_{m+1} &= \frac{m}{m+3} p_m = \frac{2m}{(m+2)(m+3)} \\ p_{m+2} &= \frac{m+1}{m+4} p_{m+1} = \frac{2m(m+1)}{(m+2)(m+3)(m+4)} \\ p_{m+3} &= \frac{m+2}{m+5} p_{m+2} = \frac{2m(m+1)}{(m+3)(m+4)(m+5)} \end{aligned} \tag{5.40}$$

At this point we notice a simple recursive pattern: by replacing $m$+3 with $k$ we obtain the probability to observe a node with degree $k$

$$p_k = \frac{2m(m+1)}{k(k+1)(k+2)}, \tag{5.41}$$

which represents the exact form of the degree distribution for the Barabási-Albert model. Note that:

- For large $k$ this becomes $p_k \sim k^{-3}$, in agreement with the numerical result.
- The prefactor of Eq. 5.11 or Eq. 5.41 is different from the prefactor derived in Eq. 5.9.

This form was derived independently in [12] and [13], and the mathematical proof of its validity is provided in [10]. Note that the rate equation formalism offers an elegant continuum equation satisfied by the degree distribution of the Barabási-Albert model [16]. Starting from the equation

$$p_k = \frac{k-1}{2} p_{k-1} - \frac{k}{2} p_\infty \tag{5.42}$$

we can write

$$2p_k = (k-1)p_{k-1} - kp(k) = -p_{k-1} - k[p_k - p_{k-1}], \tag{5.43}$$

$$2p_k = -p_{k-1} - k\frac{p_k - p_{k-1}}{k-(k-1)} = -p_{k-1} - k\frac{\partial p_k}{\partial k} \tag{5.44}$$

obtaining

$$p_k = \frac{1}{2}\frac{\partial [kp_k]}{\partial k} \tag{5.45}$$

One can check that the solution of Eq. 5.45 is

$$p_k \sim k^{-3}. \tag{5.46}$$

# ADVANCED TOPICS 5.B NONLINEAR PREFERENTIAL ATTACHMENT

The purpose of this section is to derive the degree distribution of an evolving networks governed by a nonlinear preferential attachment. We follow Krapivsky et al. [13]. As the results of Ref. [13] were derived for undirected networks, here we adjusted the calculation to cover undirected networks.

Strictly speaking the degree distribution only exists for $\alpha \leq 1$. For $\alpha > 1$ a few nodes attract a finite fraction of links, as explained in SECT. 5.7, and we do not have a stationary $p_k$. Therefore, we limit ourself to the $\alpha \leq 1$ case. We start with the Barabási-Albert model, in which at each time step a new node is added with m new links. We connect each new link to an existing node with probability

$$\Pi(k_i) = \frac{k_i^{\alpha}}{\mu(\alpha)} . \tag{5.47}$$

where $k_i$ is the degree of node i, $0 < \alpha \leq 1$ and

$$\mu(\alpha,t) = \sum_k k^{\alpha} p_k(t) . \tag{5.48}$$

is the normalization factor. Note that $\mu(0,t) = \sum_k p_k(t) = 1$ and $\mu(1,t) = \sum_k k p_k(t) = \langle k \rangle = 2mt / N$ is the average degree. Since $0 < \alpha \leq 1$,

$$\mu(0,t) \leq \mu(\alpha,t) \leq \mu(1,t) . \tag{5.49}$$

Therefore in the long time limit

$$\mu(\alpha, t \to \infty) = \text{constant} . \tag{5.50}$$

whose precise value will be calculated later. For simplicity, we adopt the notation $\mu \equiv \mu(\alpha, t \to \infty)$.

Following the rate equation approach introduced in Advanced TOPICS 5.A, we write the rate equation for the network's degree distribution as

$$p_k(t+1) = \frac{m}{\mu(\alpha,t)}[(k-1)^\alpha p_{k-1}(t) - k^\alpha p_k(t)] + \delta_{k,m}. \tag{5.51}$$

The first term on the r.h.s. describes the rate at which nodes with degree ($k$-1) gain new links; the second term describes the loss of degree-$k$ nodes when they gain new links, turning into ($k$+1) degree nodes; the last term represents the newly added nodes with degree $m$. Asymptotically, in the $t\to\infty$ limit we can write $p_k=p_k(t+1)=p_k(t)$. Substituting $k=m$ in Eq. 5.51 we obtain:

$$p_m = -\frac{m}{\mu} - m^\alpha p_m + 1 \tag{5.52}$$

$$p_m = -\frac{\mu/m}{\mu/m + m^\alpha}.$$

For $k > m$

$$p_k = \frac{m}{\mu}[(k-1)^\alpha p_{k-1} - k^\alpha p_k] \tag{5.53}$$

$$p_k = \frac{(k-1)^\alpha}{\mu/m + k^\alpha} p_{k-1} \tag{5.54}$$

Solving Eq. 5.53 recursively we obtain

$$p_m = \frac{\mu/m}{\mu/m + m^\alpha} \tag{5.55}$$

$$p_{m+1} = \frac{\mu/m \cdot m^\alpha}{\mu/m + (m+1)} \frac{\mu/m}{\mu/m + m^\alpha} \tag{5.56}$$

$$p_k = \frac{\mu/m}{k^\alpha} \prod_k^{j=m} \left(1 + \frac{\mu/m}{j^\alpha}\right)^{-1} \tag{5.57}$$

To determine the large $k$ behavior of $p_k$ we take the logarithm of (52):

$$\ln p_k = \ln(\mu/m) - \alpha \ln k - \sum_k^{j=m} \left(1 + \frac{\mu/m}{j^\alpha}\right) \tag{5.58}$$

Using the series expansion $\ln(1+x) = \sum_{n=1}^{\infty} (-1)_i^{n+1} / n \cdot x^n$ we obtain

$$\ln p_k = \ln(\mu/m) - \alpha \ln k - \sum_{i=m}^{k} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (\mu/m)^n j^{-\alpha n} \tag{5.59}$$

We approximate the sum over $j$ with the integral

$$\sum_k^{j=m} j_X^{-n\alpha} \approx \int_k^m x^{-n\alpha} dx = \frac{1}{1-n\alpha}(k^{1-n\alpha} - m^{1-n\alpha}) \tag{5.60}$$

which in the special case of $n\alpha$ =1 becomes

$$\sum_{j=m}^{k} j^{-1} \approx \int_m^k x^{-1} dx = \ln k - \ln m. \tag{5.61}$$

Hence we obtain

$$\ln p_k = \ln(\mu/m) - \alpha \ln k - \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \frac{(\mu/m)^n}{1-n\alpha}(k^{1-n.\alpha} - m^{1-n\alpha}) \tag{5.62}$$

Consequently the degree distribution has the form

$$p_k = C_\alpha k^{-\alpha} e^{-\sum_{\infty}^{n=1} \frac{(-1)^{n+1}}{n} \frac{(\mu/m)^n}{1-n\alpha} k^{1-n\alpha}}, \tag{5.63}$$

where

$$C_\alpha = \frac{\mu}{m} e^{\sum_{\infty}^{n=1} \frac{(-1)^{n+1}}{n} \frac{(\mu/m)^n}{1-n\alpha} m^{1-n\alpha}} \tag{5.64}$$

The vanishing terms in the exponential do not influence the $k \to \infty$ asymptotic behavior, being relevant only if $1-n\alpha \geq 1$ . Consequently the precise form of $p_k$ depends on α as:

$$p_k \sim \begin{cases} k^{-\alpha} e^{\frac{-\mu/m}{1-\alpha} k^{1-\alpha}}, & 1/2 < \alpha < 1 \\ k^{-\frac{1}{2}+\frac{1}{2}(\frac{\mu}{m})^2} e^{-\frac{1}{2}\frac{\mu}{m}k^{-2}}, & \alpha = 1/2 \\ k^{-\alpha} e^{\frac{-\mu/m}{1-\alpha}k^{1-\alpha} + \frac{1}{2}\frac{(\mu/m)^2}{1-2\alpha}k^{1-2\alpha}}, & 1/3 < \alpha < 1/2 \\ \vdots \end{cases} \tag{5.65}$$

That is, for $1/2 < \alpha < 1$ the degree distribution follows a stretched exponential. As we lower α, new corrections start contributing each time α becomes smaller than $1/n$, where n is an integer. For $\alpha \to 1$ the degree distribution scales as $k^{-3}$, as expected for the Barabási-Albert model. Indeed for $\alpha = 1$ we have $\mu=2$, and

$$\lim_{a \to 1} \frac{k^{1-a}}{1-a} = \ln k. \tag{5.66}$$

Therefore $p_k \sim k^{-1}\exp(-2\ln k) = k^{-3}$.

Finally we need to calculate $\mu(\alpha) = \sum^{j} j^\alpha p_j$. For this we sum Eq. 5.58:

$$\sum_{k=m}^{\infty} k^\alpha p_k = \sum_{k=m}^{\infty} \frac{\mu(\alpha)}{m} \prod_{k}^{j=m} \left(1 + \frac{\mu(\alpha)/m}{j^\alpha}\right)^{-1} \tag{5.67}$$

$$1 = \frac{1}{m} \sum_{k=m}^{\infty} \prod_{i=m}^{k} \left(1 + \frac{\mu(\alpha)/m}{j^\alpha}\right)^{-1} \tag{5.68}$$

We obtain $\mu(\alpha)$ by solving Eq. 5.52 numerically.

# ADVANCED TOPICS 5.C
# THE CLUSTERING COEFFICIENT

The purpose of this section is to derive the average clustering coefficient, Eq. 5.30, for the Barabási-Albert model. The derivation follows the an argument proposed by Klemm and Eguiluz [36], that was supported by the exact calculation of Bollobás [37]. We aim to calculate the number of triangles expected in the model, as the number of triangles can be linked to the clustering coefficient SECT. 2.10. We denote the probability to have a link between node *i* and *j* with *P(i, j)*. Therefore, the probability that three nodes *i, j, l* form a triangle is *P(i, j) P(i, l) P(j, l)*. The expected number of triangles in which node *l* with degree $k_l$ participates is thus given by the sum of the probabilities that node *l* participates in triangles with an arbitrary chosen node *i* and *j* in the network. This can be written as

$$Nr_l(\triangleleft) = \int_{i=1}^{N} di \int_{j=1}^{N} dj P(i,j)P(i,l)P(j,l) \tag{5.69}$$

To proceed we need to calculate *P(i,j)*, which requires us to consider how the Barabási-Albert model evolves. Let us denote the time when node *j* arrived with $t_j = j$, which we can do as in each time step we added only one new node. Hence the probability that at its arrival node *j* links to node *i* with degree $k_i$ is given by preferential attachment:

$$P(i,j) = m\Pi(k_i(j)) = m\frac{k_i(j)}{\sum_{l=1}^{j} k_l} = m\frac{k_i(j)}{2mj}. \tag{5.70}$$

Using Eq. 5.7, we can write

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\frac{1}{2}} = m\left(\frac{j}{i}\right)^{\frac{1}{2}}, \tag{5.71}$$

where we used the fact that the arrival time of node *j* is $t_j = j$ and the arrival time of node is $t_i = i$. Hence Eq. 5.70 now becomes

$$P(i,j) = \frac{m}{2}(ij)^{-\frac{1}{2}}. \tag{5.72}$$

Using this result we can calculate the number of triangles in Eq. 5.62, writing

$$Nr_l(\triangleleft) = \int_{i=1}^{N} di \int_{j=1}^{N} dj P(i,j)P(i,l)P(j,l) \tag{5.73}$$

$$= \frac{m^3}{8}\int_{i=1}^{N} di \int_{j=1}^{N} dj (ij)^{-\frac{1}{2}}(il)^{-\frac{1}{2}}(jl)^{-\frac{1}{2}}$$

$$= \frac{m^3}{8l}\int_{i=1}^{N} \frac{di}{i} \int_{j=1}^{N} \frac{dj}{j} = \frac{m^3}{8l}(\ln N)^2 \tag{5.74}$$

The clustering coefficient can be written as $C = \frac{2Nr_l(\triangleleft)}{k_l(k_l-1)}$, hence we obtain

$$C = \frac{\frac{m^3}{4l}(\ln N)^2}{k_l(k_l-1)} \tag{5.75}$$

To simplify Eq. 5.74, we note that according to Eq. 5.7 we have

$$k_l(t) = m\left(\frac{N}{l}\right)^{\frac{1}{2}} \tag{5.76}$$

which is the degree of node $l$ at time $t = N$. Hence, for large $k_l$ we have

$$k_l(k_l-1) \approx k_l^{\,2} = m^2\frac{N}{l} \tag{5.77}$$

allowing us to write the clustering coefficient of the Barabási-Albert model as

$$C = \frac{2Nr_l(\triangleleft)}{k_l(k_l-1)} \tag{5.78}$$

Eq. 5.78, apart from a factor 2, is the result Eq. 5.30.

# BIBLIOGRAPHY

[1] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. Science, 286:509-512, 1999.

[2] F. Eggenberger and G. Pólya. Über die Statistik Verketteter Vorgänge. ZAMM - Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik, 3:279-289, 1923.

[3] G. Udny Yule. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, f.r.s. Philosophical Transactions of the Royal Society of London. Series B, 213:21-87, 1925.

[4] Gibrat R. “Les Inégalités économiques”, Paris, France, 1931.

[5] G. K. Zipf. Human behavior and the principle of least reort. Addison-Wesley Press, Oxford, England, 1949.

[6] H. A. Simon. On a class of skew distribution functions. Biometrika, 42:425-440, 1955.

[7] D. De Solla Price. A general theory of bibliometric and othercumulative advantage processes. Journal of the American Society for Information Science, 27:292-306, 1976.

[8] R. K. Merton. The Matthew effect in science. Science, 159(3810):56-63, 1968.

[9] A.-L. Barabási. Linked: The new science of networks. Perseus, New York, 2002.

[10] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process. Random Structures and Algorithms, 18:279-290, 2001.

[11] A.-L. Barabási, H. Jeong, R. Albert. Mean-field theory for scale free

random networks. Physica A, 272:173-187, 1999.

[12] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Structure of growing networks with preferential linking. Phys. Rev. Lett., 85:4633-4636, 2000.

[13] P. L. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of growing random networks. Phys. Rev. Lett., 85:4629-4632, 2000.

[14] H. Jeong, Z. Néda. A.-L. Barabási. Measuring preferential attachment in evolving networks. Europhysics Letters, 61:567-572, 2003.

[15] M. E. J. Newman. Clustering and preferential attachment in growing networks, Phys. Rev. E 64, 025102, 2001.

[16] S.N. Dorogovtsev and J.F.F. Mendes. Evolution of networks. Oxford Clarendon Press, 2002.

[17] J. M. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. The Web as a graph: measurements, models and methods. Proceedings of the International Conference on Combinatorics and Computing, 1999.

[18] R. Kumar, P. Raghavan, S. Rajalopagan, D. Divakumar, A. S. Tomkins, and E. Upfal. The Web as a graph. Proceedings of the 19th Symposium on principles of database systems, 2000.

[19] Pastor-Satorras, R., Smith, E. & Sole, R. Evolving proteinminteraction networks through gene duplication. J. Theor. Biol. 222:m199–210, 2003.

[20] Vazquez, A., Flammini, A., Maritan, A. & Vespignani, A. Modeling of protein interaction networks. ComPlexUs 1:38–44, 2003.

[21] G.S. Becker, The economic approach to Human Behavior. Chicago, 1976.

[22] A. Fabrikant, E. Koutsoupias, and C. Papadimitriou. Heuristically optimized trade-offs: a new paradigm for power laws in the internet. In Proceedings of the 29th International Colloquium on Automata, Languages, and Programming (ICALP), pages 110-122, Malaga, Spain, July 2002.

[23] RM. D'Souza, C. Borgs, J. T. Chayes, N. Berger, and R. D. Kleinberg, Emergence of tempered preferential attachment from optimization, PNAS 104, 6112-6117, 2007.

[24] F. Papadopoulos, M. Kitsak, M. Angeles Serrano, M. Boguna, and D. Krioukov, Popularity versus similarity in growing networks, Nature, 489: 537, 2012.

[25] A.-L. Barabási Network science: luck or reason, Nature 489: 1-2, 2012.

[26] H. A. Simon. On a class of skew distribution functions. Biometrika 42:3, 425-440, 1955.

[27] B. Mandelbrot. An Informational Theory of the Statistical Structure of Languages. In Communication Theory, edited by W. Jackson, pp. 486-502. Woburn, MA: Butterworth, 1953.

[28] B. Mandelbrot. A note on a class of skew distribution function: analysis and critique of a Paper by H.A. Simon. Information and control 2: 90-99, 1959.

[29] H. A. Simon. Some Further Notes on a class of skew distribution functions. Information and Control 3: 80-88, 1960.

[30] B. Mandelbrot. Final Note on a Class of Skew Distribution Functions: Analysis and Critique of a Model due to H.A. Simon. Information and Control 4: 198-216., 1961.

[31] H. A. Simon. Reply to final note. Information and Control 4:, 217-223, 1961.

[32] B. Mandelbrot. Post scriptum to final note. Information and Control 4: 300-304 1961.

[33] H. A. Simon. Reply to Dr. Mandelbrot's Post Scriptum. Information and Control 4: 305-308, 1961.

[34] R. Cohen and S. Havlin. Scale-free networks are ultrasmall. Phys. Rev. Lett., 90:058701, 2003.

[35] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. Combinatorica, 24:5-34, 2004.

[36] K. Klemm and V. M. Eguluz. Growing scale-free networks with small-world behavior. Phys. Rev. E, 65:057102, 2002.

[37] B. Bollobás, O.M. Riordan. Mathematical results on scale-free random graphs, in the Handbook of Graphs and Networks, edited by S. Bormholdt and A. G. Schuster, Wiley, 2003.

# CHAPTER 6

# EVOLVING NETWORKS



**Figure 6.0 (front cover)**
**Network representation by Mauro Martino**

# INTRODUCTION

Founded six years after birth of the World Wide Web, Google was a latecomer to search. By the late 1990s Alta Vista and Inktomi, two search engines with an early start, have been dominating the search market. Yet Google, the third mover, soon not only became the leading search engine, but acquired links at such an incredible rate that by 2000 became the most connected node of the Web as well [1]. But its status didn't last: in 2011 Facebook, with an even later start, took over as the Web's biggest hub.

This competition for the top spot is by no means unique to the online world: the history of business is full of companies whose consumers were hijacked by a more successful latecomer. Take Apple, whose ingenious Newton handheld, introduced in 1987, was wiped off the market by Palm. A decade later Apple engineered a dramatic comeback, creating the iPad, that changed the concept of a handheld computer. If we view the market as a bipartite network whose nodes are products and whose links are purchasing decisions, we can say that Apple's links in the 1990s were rewired to Palm, only to be re-captured by Apple again a decade later. This competitive landscape highlights an important limitation of our current modeling framework: the network models we encountered so far cannot account for it. Indeed, in the Erdős-Rényi model the identity of the biggest node is driven entirely by chance. The Barabási-Albert model offers a more realistic picture, predicting that each node increases its degree following $k(t) \sim t^{1/2}$ Eq. 5.6. This means that the oldest node always has the most links, a phenomena called the "first mover's advantage" in the business literature. It also means that late nodes can never turn into the largest hubs.

In reality a node's growth does not depend on the node's age only. Instead webpages, companies, or actors have intrinsic qualities that influence the rate at which they acquire links. Some show up late and nevertheless grab most links within a short timeframe. Others rise early yet never quite make it. The goal of this chapter is to understand how the differences in the node's ability to acquire links, and other processes not captured by the Barabási-Albert model, like node and link deletion or aging, affect the network topology.

# THE BIANCONI-BARABÁSI MODEL

Some people have a knack for turning each random encounter into a lasting social link; some companies turn each consumer into a loyal partner; some webpages turn visitors into addicts. A common feature of these successful nodes is some intrinsic property that propels them ahead of the other nodes. We will call this property fitness. Fitness is an individual's skill to turn a random encounter into a lasting friendship; it is a company's competence in acquiring consumers relative to its competition; a webpage's ability to bring us back on a daily basis despite the many other pages that compete for our attention. Fitness may have genetic roots in people, it may be related to management quality and innovativeness in companies and may depend on the content offered by a website. In the Barabási-Albert model we assumed that a node's growth rate is determined solely by its degree. To incorporate the role of fitness we assume that preferential attachment is driven by the product of a node's fitness, $\eta$, and its degree $k$.

**Movie 6.1**

**The evolution of the Bianconi-Barabási model**

The movie shows a growing network in which each new node acquires a randomly chosen fitness parameter at birth, represented by the color of the node. Each new node chooses the nodes it links to following generalized preferential attachment, making each node's growth rate proportional to its fitness. The node size is shown proportionally to its degree, illustrating that with time the nodes with the highest fitness turn into the largest hubs.

*Video courtesy of D. Wang.*

The resulting model consists of the following two steps [2, 3]:

- **Growth:** In each timestep a new node $j$ with $m$ links and fitness $\eta_j$ is added to the system, where $\eta_j$ is a random number chosen from a distribution $\rho(\eta)$. Once assigned, a node's fitness does not change.

- **Preferential Attachment:** The probability that a link of a new node connects to a pre-existing node $i$ is proportional to the product of node $i$'s degree $k_i$ and its fitness $\eta_i$

$$\Pi_i = \frac{\eta_i k_i}{\sum_j \eta_j k_j}. \qquad (6.1)$$

In Eq. 6.1 the dependence of $\Pi_i$ on $k_i$ captures the fact that higher-degree nodes are easier to encounter, hence we are more likely to link to them. The dependence of $\Pi_i$ on $\eta_i$ implies that between two nodes with the same degree, the one with higher fitness is selected with a higher probability. Hence, Eq. 6.1 assures that even a relatively young node with initially only a few links can acquire links rapidly if it has larger fitness than the rest of

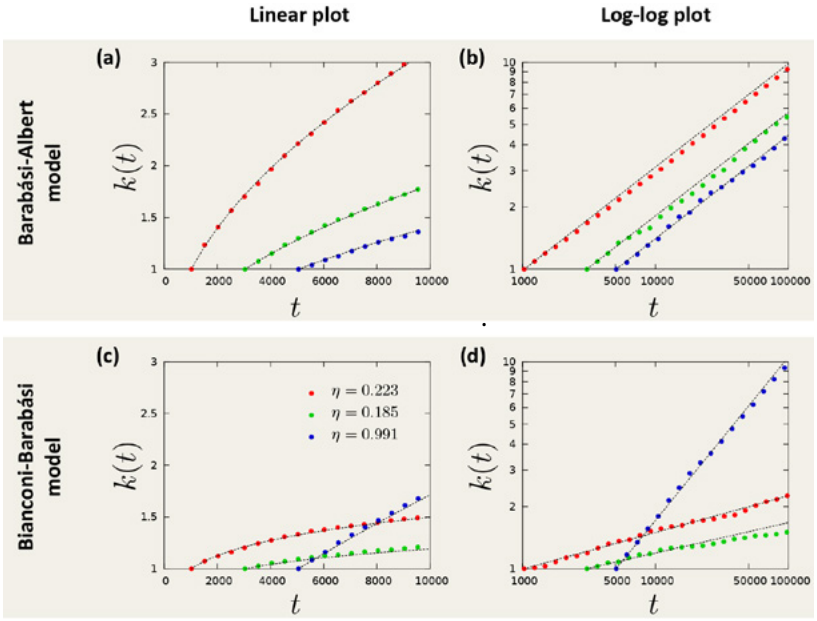



**Figure 6.1**
**Competition in the Bianconi-Barabási model**

**(a)** In the Barabási-Albert model all nodes increase their degree at the same rate, hence the earlier a node joins the network, the larger will be its degree. The figure shows the time dependence of the degree for nodes that arrived at different times, indicating that the later nodes are unable to pass the earlier nodes.

**(b)** Same as in (a) but in a log-log plot, demonstrating that each node follows the same growth law with identical dynamical exponents $\beta = 1/2$.

**(c)** In the Bianconi-Barabási model nodes increase their degree at a rate that is determined by their individual fitness. Hence a latecomer node (blue symbols) can overcome the earlier nodes.

**(d)** Same as in (c) but on a log-log plot, demonstrating that each node follows a growth curve with its own fitness-dependent dynamical exponent $\beta$, as predicted by **Eq. 6.3** and **Eq. 6.4**.

In (a)-(d) each curve corresponds to average over several independent runs using the same fitness sequence.

the nodes. We will call the model introduced above the *Bianconi-Barabási* model after the authors of the paper that introduced it [2, 3]. In the literature one may also account it as the *fitness model.*

### DEGREE DYNAMICS

We can use the continuum theory to predict a node's temporal evolution in the model defined above. According to **Eq. 6.1**, the degree of node $i$ changes at the rate

$$\frac{\partial k_i}{\partial t} = m\frac{\eta_i k_i}{\sum_k \eta_j k_j} \tag{6.2}$$

Let us assume that the time evolution of $k_i$ follows a power law with a fitness-dependent exponent $\beta(\eta_i)$ **Fig. 6.1**,

$$k_{\eta_i}(t,t_i) = m\left(\frac{t}{t_i}\right)^{\beta(\eta_i)}. \tag{6.3}$$

Inserting **Eq. 6.3** into **Eq. 6.2** we find that the dynamic exponent satisfies

**ADVANCED TOPICS 6.A**

$$\beta(\eta) = \frac{\eta}{C} \tag{6.4}$$

with

$$C = \int \rho(\eta)\frac{\eta}{1-\beta(\eta)}d\eta. \tag{6.5}$$

In the Barabási-Albert model we have $\beta = 1/2$, indicating that the degree of each node increases as a square root of time. In contrast, according to **Eq. 6.4**, in the Bianconi-Barabási model the dynamic exponent is proportional to the node's fitness, $\eta$, hence each node has its own dynamic exponent. Consequently, a node with a higher fitness will increase its degree faster. Given sufficient time, the fitter node will leave behind each node that has a smaller fitness **BOX 6.1**. Facebook is a poster child of this phenomenon: a latecomer with an addictive product, it acquired links faster than its competitors, eventually becoming the Web's biggest hub.

### DEGREE DISTRIBUTION

The degree distribution of the network generated by the Bianconi-Barabási model can be calculated using the continuum theory ADVANCED TOPICS 6.A, predicting that

$$p_k \sim C \int d\eta \frac{\rho(\eta)}{\eta} \left(\frac{m}{k}\right)^{\frac{C}{\eta}+1}. \tag{6.6}$$

Eq. 6.6 is a weighted sum of multiple power-laws, indicating that $p_k$ depends on the precise form of the fitness distribution, $\rho(\eta)$. To illustrate the properties of the model we apply Eq. 6.4 and Eq. 6.6 to calculate $\beta(\eta)$ and $p_k$ for two different fitness distributions:

- Equal fitnesses

  When all fitnesses are equal, the Bianconi-Barabási model should reduce to the Barabási-Albert model. Indeed, let us use $\rho(\eta) = \delta(\eta - 1)$, capturing the fact that each node has the same fitness $\eta = 1$. In this case, Eq. 6.5 predicts $C = 2$. Using Eq. 6.4 we obtain $\beta = 1/2$ and Eq. 6.6 predicts $p_k \sim k^{-3}$, the known scaling of the degree distribution in the Barabási-Albert model.

- Uniform fitness distribution

  The model's behavior is more interesting when nodes have different fitnesses. Let us choose $\eta$ to be uniformly distributed in the [0,1] interval. In this case $C$ is the solution of the transcendental equation Eq. 6.5

$$\exp(-2/C) = 1 - 1/C \tag{6.7}$$

  whose numerical solution is $C^* = 1.255$. According to Eq. 6.4, each node $i$ has a different dynamic exponent, $\beta(\eta_i) = \eta_i / C^*$. Using Eq. 6.6 we obtain

$$p_k \sim \int_1^0 d\eta \frac{C^*}{\eta} \frac{1}{k^{1+C^*/\eta}} \sim \frac{k^{-(1+C^*)}}{\ln k}, \tag{6.8}$$

  predicting that the degree distribution follows a power law with degree exponent $\gamma = 2.255$, affected by an inverse logarithmic correction $1/\ln k$.

Numerical support for these predictions is provided in Fig. 6.1 and Fig. 6.2. The simulations confirm that $k_i(t)$ follows a power law for each $\eta$ and that the dynamical exponent $\beta(\eta)$ increases with the fitness $\eta$. As Fig. 6.2 a indicates, the measured dynamical exponents are in excellent agreement with the prediction of Eq. 6.4. Fig. 6.2b also documents an agreement between Eq. 6.8 and the numerically obtained degree distribution.

In summary, the Bianconi-Barabási model can account for the different rate at which nodes with different internal characteristics acquire links. It predicts that a node's growth rate is directly determined by its fitness $\eta$ and allows us to calculate the dependence of the degree distribution on the fitness distribution $\rho(\eta)$.

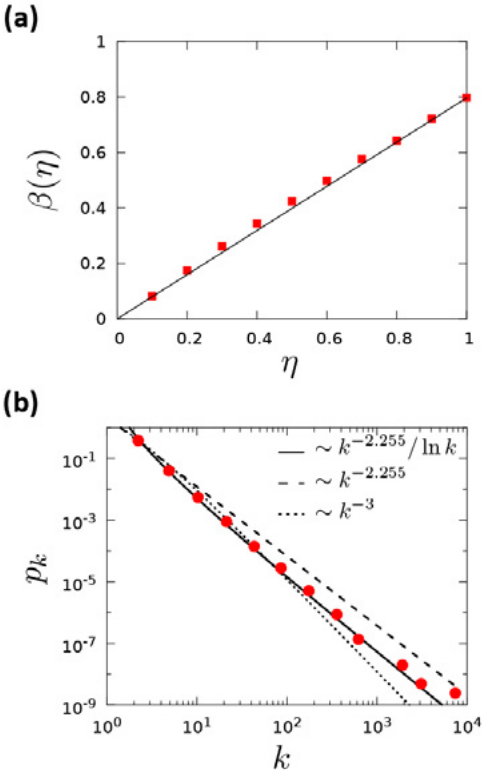



**Figure 6.2**
**Characterizing the Bianconi-Barabási model**

**(a)** The measured dynamic exponent $\beta(\eta)$ shown in function of $\eta$ in the case of a uniform $\rho(\eta)$ distribution. The squares were obtained from numerical simulations while the solid line corresponds to the analytical prediction $\beta(\eta) = \eta/1.255$.

**(b)** Degree distribution of the fitness model obtained numerically for a network with $m=$ and $N = 10^6$ and for fitnesses chosen uniformly from the $\eta \in [0, 1]$ interval. The solid line corresponds to the theoretical prediction Eq. 6.8 with $\gamma = 2.255$. The dashed line corresponds to a simple fit $p_k \sim k^{-2.255}$ without the logarithmic correction, while the long-dashed curve correspond to $p_k \sim k^{-3}$, expected if all fitness are equal. Note that the best fit is provided by Eq. 6.8.

# MEASURING FITNESS

Measuring the fitness of a node could help us identify web sites that are poised to grow in visibility, research papers that will become influential, or actors on their way to stardom. Yet, our ability to determine the utility of a webpage is prone to errors: while a small segment of the population might find a webpage on sumo wrestling fascinating, most individuals are indifferent to it and some might even find it repulsive. Hence, different individuals will inevitably assign different fitnesses to the same node. Yet, according to Eq. 6.1 fitness reflects the network's collective perception of a node's importance relative to the other nodes. Thus, we can determine a node's fitness by comparing its time evolution to the time evolution of other nodes in the network. In this section we show that if we have dynamical information about the evolution of the individual nodes, the conceptual framework of the Bianconi-Barabási model allows us to determine the fitness of each node.

To relate a node's growth rate to its fitness we take the logarithm of Eq. 6.3,

$$\log k_{n_i}(t,t_i) = \beta(n_i)\log t + \beta_i \,. \tag{6.9}$$

where $B_i = \log(m/t_i^{\beta(\eta_i)})$ is a time-independent parameter. Hence, the slope of $\log k_{\eta_i}(t, t_i)$ is a linear function of the dynamical exponent $\beta(\eta_i)$, which depends linearly on $\eta_i$ according to Eq.6.4. Therefore, if we can track the time evolution of the degree for a large number of nodes, the distribution of the obtained growth exponent $\beta(\eta_i)$ will be identical with the fitness distribution $\rho(\eta)$. Such measurement were first carried out in the context of the WWW, relying on a dataset that crawled the links of about 22 million web documents per month for 13 months [9]. While most nodes (documents) did not change their degree during this time frame, 6.5% of nodes showed sufficient changes to allow the determination of their growth exponent via Eq. 6.9. The obtained fitness distribution $\rho(\eta)$ has an exponential form Fig. 6.3, indicating that high fitness nodes are exponentially rare. This is somewhat unexpected, as one would be tempted to assume that on the web fitness varies widely: Google is probably significantly more interesting to Web

## BOX 6.1

### THE GENETIC ORIGINS OF FITNESS

Could fitness, an ability to acquire friends in a social network, have genetic origins? To answer this researchers examined the social network characteristics of 1,110 school-age twins [6, 7], using a technique previously developed to identify the heritability of a variety of traits and behaviors. The measurements indicate that:

- Genetic factors account for 46% of the variation in a student's in-degree (i.e. the number of students that name a given student as a friend).

- Generic factors are not significant for out-degrees (i.e. the number of students a given student names as friends).

This suggests that an individual's ability to acquire links, i.e. its fitness, is heritable. Hence, fitness may have genetic origins. This conclusion is also supported by research that associated a particular genetic variation with variation in popularity [8].

surfers than my personal webpage. Yet the exponential form of $\rho(\eta)$ indicates that most Web documents have comparable fitness. Consequently, the observed large differences in the degree of various web documents is the result of the system's dynamics: growth and preferential attachment amplifies the small fitness differences, turning nodes with slightly higher fitness into much bigger nodes. To illustrate this amplification, consider two nodes that arrived at the same time, but have different fitnesses $\eta_2 > \eta_1$. According to Eq. 6.3 and Eq. 6.4, the relative difference between their degrees grows with time as

$$\frac{k_2 - k_1}{k_1} \sim t^{\frac{\eta_2 - \eta_1}{C}}. \tag{6.10}$$

while the difference between $\eta_2$ and $\eta_1$ may be small, far into the future (large $t$) the relative difference between their degrees can become quite significant.

### CASE STUDY: MEASURING THE FITNESS OF A SCIENTIFIC PUBLICATION

Eq. 6.9 assumes that Eq. 6.3 fully captures a Web document's temporal evolution. In some systems nodes follow a more complex dynamics, that we must account for when we try to measure their fitness. We illustrate this by determining the fitness of a research publication, allowing us to predict its future impact [11]. While most research papers acquire only a few citations, a small number of publications collect thousands and even tens of thousands of citations. These differences capture the considerable impact disparity characterizing the scientific enterprise BOX 6.3. These impact differences mirror differences in the novelty and the content of various publications. In general, we can write the probability that paper i is cited at time t after publication as [11]

$$\Pi_i \sim \eta_i c_i^t P_i(t), \tag{6.11}$$

where $\eta_i$ is the paper's fitness, accounting for the perceived noveltyand importance of the reported discovery; $c_t^i$ is the cumulative number of citations acquired by paper i at time t after publication, accounting for the fact that well-cited papers are more likely to be cited again than less-cited contributions. The last term in Eq. 6.11 captures the fact that new ideas are integrated in subsequent work, hence the novelty of each paper fades with time [11, 12]. The measurements indicate that this decay has the log-normal form

$$P_i(t) = \frac{1}{\sqrt{2\pi\sigma_i t}} e^{-\frac{(\ln t - \mu_i)^2}{2\sigma_i^2}} \tag{6.12}$$

By solving the master equation behind Eq. 6.11, we obtain

$$c_i^t = m\left(e^{\frac{\beta\eta_i}{A}\Phi\left(\frac{\ln t - \mu_i}{\sigma_i}\right)}\right), \tag{6.13}$$

where

$$\Phi(x) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy \tag{6.14}$$

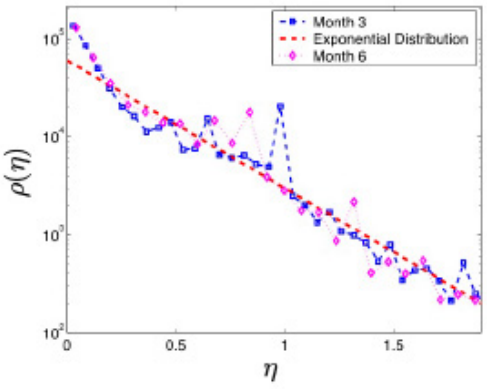



**Figure 6.3**
**The fitness distribution of the WWW**

The fitness distribution obtained from Eq. 6.9 by comparing the degree evolution of a large number of Web documents. The measurements indicate that each node's degree has a power law time dependence, as predicted by Eq. 6.3. The slope of each curve is $\beta(\eta_j)$, which corresponds to the node's fitness $\eta_i$ up to a multiplicative constant according to Eq. 6.4. The plot shows the result of two measurements based on datasets recorded three months apart, demonstrating that the fitness distribution is time independent. The dashed line indicates that the fitness distribution follows an exponential form. After [9].

## BOX 6.2

### ULTIMATE IMPACT

Citation counts offer only the momentary impact of a research paper. Therefore, they represent an inherently weak measure of long-term impact, as we do not know if a paper that acquired a hundred citations in two years has already had its run, or will continue to grow in impact, acquiring thousands more. Ideally we would like to predict how many citations will a paper acquire during its lifetime, or its ultimate impact. The citation model Eq. 6.11 and Eq. 6.14 allows us to determine the ultimate impact by taking the $t \to \infty$ limit in Eq. 6.13, finding [11]

$$c_i^{\infty} = m(e^{\eta_i} - 1). \tag{6.15}$$

is the cumulative normal distribution and $m$, $\beta$, and A are global parameters. Eq. 6.13 and Eq. 6.14 predict that the citation history of paper $i$ is characterized by three fundamental parameters: the relative fitness $\eta_i' \equiv \eta_i \beta / A$, measuring a paper's importance relative to other papers; the immediacy $\mu_i$, governing the time for a paper to reach its citation peak and the longevity $\sigma_i$, capturing the decay rate.

We fit Eq. 6.13 to the citation history of individual papers published by a given journal to obtain the journal's fitness distribution Fig. 6.4. We find that *Cell* has a fitness distribution shifted to the right, indicating that *Cell* papers tend to have high fitness. By comparison the fitness of papers published in *Physical Review* are shifted to the left, indicating that the journal publishes fewer high impact papers.

In summary, the framework offered by the Bianconi-Barabási model allows us to experimentally determine the fitness of individual nodes and the shape of the fitness distribution $\rho(\eta)$. The fitness distribution is typically bounded, meaning that differences in fitness between different nodes are small. With time these differences are magnified however, resulting in an unbounded (power law) degree distribution in incoming links in the case of the WWW or broad citation distribution in citation networks.

Eq. 6.15 predicts that despite the myriad of factors that contribute to the success and the citation history of a research paper, its ultimate impact is determined only by its fitness $\eta'_i$. As fitness can be determined by fitting Eq. 6.13 to a paper's existing citation history, we can use Eq. 6.15 to predict the ultimate impact of a publication.

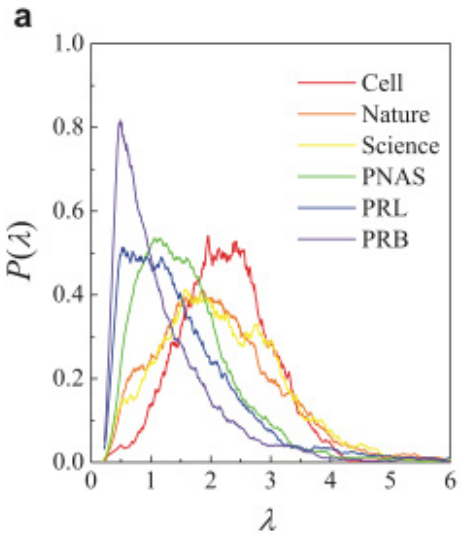



**Figure 6.4**
**Fitness distribution of research papers**

The fitness distribution of papers published in six journals in 1990. Each paper's fitness was obtained by fitting Eq. 6.13 to the paper's citation history for a decade long time interval following 1990. Two journals are from physics (Physical Review B and Physical Review Letters), one from biology (Cell) and three are interdisciplinary, meaning that they publish papers from different areas of science (Nature, Science, and PNAS).

The obtained fitness distributions are shifted relative to each other, indicating that Cell publishes papers with the highest fitness, followed by Nature and Science, PNAS, Physical Reviews Letters and Physical Review B. After [11].

# BOX 6.3

### THE TOP ONE PERCENT

The "one percent" phrase has dominated the discourse during the 2012 US presidential election, reminding everyone that one percent of the population earns a disproportional 17.42% of the total US income. To those familiar with power laws this is hardly surprising: it is a consequence of the fat-tailed nature of the income distribution. Therefore, the "one percent" phenomenon is present in any quantity that follows a power law, from the links of the WWW to scientific impact [10].

The "one percent" debate is not as much about the magnitude of the income disparity, but its trends: income disparity dropped between 1940 and 1970, only to skyrocket again in the past decades (see black line). As models explaining the distribution of income predict a time-invariant distribution, the observed changes offer evidence of endogenous shifts in the share of the top one percent. As the red line indicates, the impact disparity in physical sciences has also been rising steadily over the past century. Indeed, while in 1930 a year after publication the top 1% of papers got only about 5% of the citations, today the magnitude of this impact disparity is comparable to the income disparity.

This shift of the bulk of the citations to a few of publications may reflect the fact that while the number of research papers exploded, the time we devote to reading them has not. Hence, we increasingly rely on crowdsourcing to discover relevant work, a process that favors the highly cited publications.

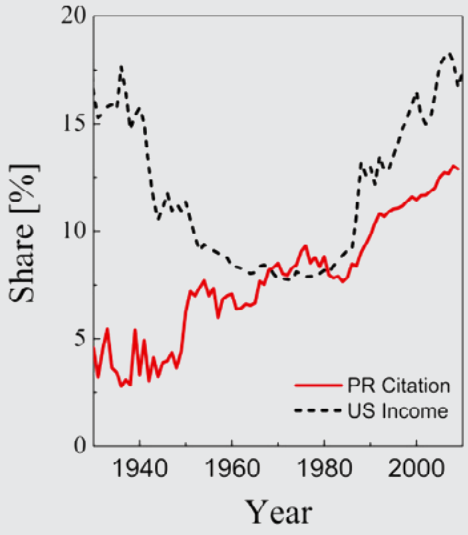



**Figure 6.5**
**The 1% of Science**

The share of citations in a given year received by the top 1% of all papers published during the previous year in Physical Review. The data captures the citation history of 463,348 papers published between 1893 and 2009 (red line). Also shown is the fraction of income earned by the top 1% of the population in US (black dashed line).

# BOSE-EINSTEIN CONDENSATION

In the previous section we found that the Web's fitness distribution follows a simple exponential Fig. 6.3, while the fitness of research papers follows a peaked distribution Fig. 6.4. The diversity of the observed fitness distributions raises an important question: how does the network topology depend on the precise shape of $\rho(\eta)$? Technically, the answer is provided by Eq. 6.6 that links $p_k$ to $\rho(\eta)$. Yet, the true impact of the fitness distribution was realized only after the discovery that some networks can undergo a Bose-Einstein condensation BOX 6.5, with significant consequences on the network topology [13]. We start by establishing a formal link between the Bianconi-Barabási model and a Bose gas, whose properties have been extensively studied in physics Fig. 6.5:

- **Fitness → Energy**: to each node with fitness $\eta_i$ we assign an energy $\varepsilon_i$ using

$$\varepsilon_i = \frac{1}{\beta_T} \log \eta_i \,. \tag{6.16}$$

  In physical systems $\beta_T$ plays the role of the inverse temperature. Hence, we use the subscript $T$ to distinguish $\beta_T$ from the dynamic exponent $\beta$. According to Eq. 6.16, each node in a network corresponds to an energy level in a Bose gas. The larger the node's fitness, the lower is its energy.

- **Links → Particles:** for each link between nodes $i$ and $j$ we add a particle at the energy levels $\varepsilon_i$ and $\varepsilon_j$, respectively.

- **Nodes → Energy levels:** the arrival of a new node with m links corresponds to adding a new energy level $\varepsilon_j$ and $2m$ new particles to the Bose gas. Half of these particles land on level $\varepsilon_j$, corresponding to the links that start from the node $j$; the remaining m particles are distributed between the energy levels that correspond to the nodes to which the new node links to.

If we follow the mathematical consequences of this mapping, we find that in the resulting gas the number of particles on each energy level fol-

**Movie 6.2**
**Bose-Einstein condensation in networks**

The movie shows the time evolution of a growing network in which one node (purple) has a much higher fitness than the rest of the nodes. Consequently this high fitness node attracts most links, forcing the system to undergo a Bose-Einstein condensation.

*Video courtesy of D. Wang.*

lows a Bose statistics, a formula derived by Satyendra Nath Bose in 1924, representing a fundamental result in quantum statistics BOX 6.6. Consequently, the links of the fitness model behave like subatomic particles in a quantum gas. This mapping to a Bose gas is exact and predicts the existence of two distinct phases [13, 14]:

#### SCALE-FREE PHASE

For most fitness distributions the network displays a fit-gets-rich dynamics, meaning that the degree of each node is ultimately determined by its fitness. While the fittest node will inevitably become the largest hub, in the scale-free phase the fittest node is not significantly bigger than the next fittest node.

# BOX 6.4

### BOSE-EINSTEIN CONDENSATION

In classical physics atoms can be distinguished and individually numbered, like the numbered balls used to pick the winning number in lottery. In the subatomic world particles differ in our ability to distinguish them: Fermi particles, like electrons, can be distinguished; in contrast Bose particles, like photons, are indistinguishable. Distinguishability impacts the energy of a particle. In classical physics the kinetic energy of a moving particle, $E=mv^2/2$, can have any value between zero (at rest) and an arbitrarily large $E$, when it moves very fast. In quantum mechanics energy is quantized, which means that it can only take up discrete (quantized) values. This is where distinguishability matters: the distinguishable Fermi particles are forbidden to have the same energy. Hence, only one electron can occupy a given energy level Fig. 6.7a. As Bose particles cannot be distinguished, many can crowd on the same energy level Fig. 6.7b.

At high temperatures, when thermal agitation forces the particles to take up different energies, the difference between a Fermi and a Bose gas is negligible Fig. 6.7a, b. The difference becomes significant at low temperatures when all particles are forced to take up their lowest allowed energy. In a Fermi gas at low temperatures the particles fill the energy levels from bottom up, just like pouring water fills up a vase Fig. 6.7c. However, as any number of Bose particles can share the same energy level, they can all crowd at the lowest energy Fig. 6.7d. Hence, no matter how much "Bose water" we pour into the vase, it will stay at the bottom of the vessel, never filling it up. This phenomenon is called a Bose-Einstein condensation and it was first proposed by Einstein in 1924. Experimental evidence for Bose-Einstein condensation emerged only in 1995 and was recognized with the 2001 Nobel prize in physics.

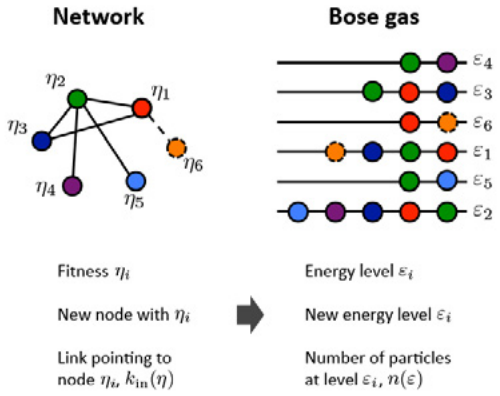



**Figure 6.6**
**Mapping networks to a Bose gas**

**Left:** A network of six nodes, each node characterized by a unique fitness, $\eta_i$, indicated by the color of the node. The individual fitnesses are chosen from the distribution $\rho(\eta)$.

**Right:** The mapping assigns an energy level $\varepsilon$ to each fitness $\eta$, resulting in a Bose gas with random energy levels. A link from node $i$ to node $j$ corresponds to two particles, one at level $\varepsilon_i$ and the other at level $\varepsilon_j$.

**Growth:** The network grows by adding a new node, like the node with fitness $\eta_6$, at each time step. The new node connects to $m$=1 other nodes (dashed link), chosen randomly following Eq. 6.1. In the Bose gas this results in the addition of a new energy level $\varepsilon_6$ (dashed line), populated by two particles, and the deposition of another particle at $\varepsilon_1$, the energy level to which $\eta_6$ connects to.

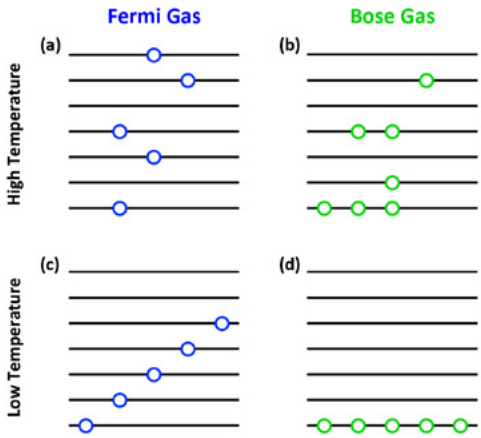



**Figure 6.7**
**Bose and Fermi statistics**

In a Fermi gas (a, c) only one particle is allowed on each energy level, while in a Bose gas (b,d) there is no such a restriction. At high temperatures it is hard, to notice the difference between the two gases. At low temperatures, however, particles want to occupy the lowest possible energy and the difference between the two gasses becomes significant.

Indeed, at any moment the degree distribution follows a power law, indicating that the largest hub is closely followed by a few slightly smaller hubs, with almost as many links as the fittest node Fig. 6.8a. The uniform fitness distribution discussed in the previous section results in a scale-free network.

### BOSE-EINSTEIN CONDENSATION

The unexpected outcome of the mapping to a Bose gas is the possibility of a Bose-Einstein condensation for some fitness distributions $\rho(\eta)$ BOX 6.7. In a Bose-Einstein condensate all particles crowd to the lowest energy level, leaving the rest of the energy levels unpopulated BOX 6.5. In a network this means that the fittest node grabs a finite fraction of the links, turning into a super-hub Fig. 6.8b, and the network develops a hub-and-spoke topology. In these networks the rich-gets-richer process is so dominant that becomes a winner takes-all phenomenon. Consequently, the network will loose its scale-free nature.

In summary, the precise shape of the fitness distribution, $\rho(\eta)$, plays an important role in shaping the topology of a growing network. While most fitness distributions (like the uniform distribution) lead to a power law degree distribution, some $\rho(\eta)$ allow for Bose-Einstein condensation. If a network undergoes a Bose-Einstein condensation, then one or a few nodes grab most of the links. Hence, the rich-gets-richer process that generates the scale-free state, turns into a winner-takes-all phenomenon. The Bose-Einstein condensation has such an obvious impact on a network's structure that, if present, it is hard to miss: it destroys the hierarchy of hubs characterizing a scale-free network, turning it into a star-like topology BOX 6.8.

## BOX 6.5

### FROM FITNESS TO A BOSE GAS

In the context of the Bose gas of Fig. 6.6 the probability that a particle lands on level $i$ is given by

$$\Pi_i = \frac{e^{-\beta_T \varepsilon_i} k_i}{\sum_j e^{-\beta_T \varepsilon_j} k_i} . \qquad (6.17)$$

Hence, the rate at which the energy level $\varepsilon_i$ accumulates particles is [13]

$$\frac{\partial k_i(\varepsilon_i,t,t_i)}{\partial t} = m \frac{e^{-\beta_T \varepsilon_i} k_i(\varepsilon_i,t,t_i)}{Z_t} \qquad (6.18)$$

where $k_i(\varepsilon_i, t, t_i)$ is the occupation number of level $i$ Fig. 6.6 and

$$Z_t \equiv \sum_{j=1}^{t} t e^{-\beta_T \varepsilon_j} k_j(\varepsilon_i,t,t_j) \qquad (6.18a)$$

is the partition function. The solution of Eq. 6.18 is

$$k_i(\varepsilon_i,t,t_i) = m\left(\frac{t}{t_i}\right)^{f(\varepsilon_i)} \qquad (6.19)$$

where $f(\varepsilon) = e^{-\beta_T(\varepsilon-\mu)}$ and $\mu$ is the chemical potential satisfying

$$\int \deg(\varepsilon) \frac{1}{e^{\beta_T(\varepsilon-\mu)}-1} = 1 . \qquad (6.20)$$

Here, $\deg(\varepsilon)$ is the degeneracy of the energy level $\varepsilon$. Eq. 6.20 suggests that in the limit $t \to \infty$ the occupation number, representing the number of particles with energy $\varepsilon$, follows the well-known Bose statistics

$$n(\varepsilon) \frac{1}{e^{\beta_T(\varepsilon-\mu)}-1} . \qquad (6.21)$$

This concludes the mapping of the fitness model to a Bose gas, indicating that the node degrees in the fitness model follow Bose statistics.

# BOX 6.6

### FITNESSES DISTRIBUTION LEADING TO BOSE-EINSTEIN CONDENSATION

In physical systems Bose-Einstein condensation is induced by lowering the temperature of the Bose gas below some critical temperature. In networks, the temperature $\beta_T$ in Eq. 6.16 is a dummy variable, disappearing from all topologically relevant quantities, like the degree distribution $p_k$. Hence, the presence or absence of Bose-Einstein condensation depends only on the form of the fitness distribution $\rho(\eta)$. In order for a network to undergo Bose-Einstein condensation, the fitness distribution needs to satisfy the following conditions:

(a) $\rho(\eta)$ must have a maximum $\eta_{\max}$.This means that $\eta$ needs to have a clear upper bound.

(b) $\rho(\eta_{max})$=0, i.e. the system requires an infinite time to reach $\eta_{\max}$.

The uniform distribution, $\eta \in [0, 1]$ satisfies (a), as it is bounded, having $\eta_{\max}$=1. It fails, however, the criteria (b), as it can reach $\eta_{\max}$= 1 with a finite probability. Consequently, we cannot observe a Bose-Einstein condensation in this case. A fitness distribution that can lead to a Bose-Einstein condensation is

$$\rho(\eta) = (1-\eta)^{\zeta} \tag{6.22}$$

satisfying both (a) and (b). Indeed, $\eta_{\max} = 1$ and $\eta(1) = 0$, which is the reason why, upon varying $\zeta$, we can observe Bose-Einstein condensation Fig. 6.8. Indeed, the existence of the solution of Eq. 6.20 depends on the functional form of the energy distribution, $g(\varepsilon)$, determined by the $\rho(\eta)$ fitness distribution. Specifically, if Eq. 6.22 has no non-negative solution for a given $g(\varepsilon)$, we observe a Bose-Einstein condensation, indicating that a finite fraction of the particles agglomerate at the lowest energy level.

# BOX 6.7

### MICROSOFT AND BOSE-EINSTEIN CONDENSATION

Think of the operating systems (OS) that run on each computer as nodes that compete for links in terms of users or computers. Each time a user installs Windows on his or her computer, a link is added to Microsoft. If a fit-gets-rich behavior of scale-free networks prevails in the marketplace, there should be a hierarchy of operating systems, such that the most popular node is followed closely by several less popular nodes.

In the OS market, however, such hierarchy is absent. True, Windows is not the only available operating system. All Apple products run Mac OS; DOS, the precursor of Windows, is still installed on some PCs; Linux, a free operating system, continues to gain market share and UNIX runs on many computers devoted to number crunching.

But all these operating systems are dwarfed by Windows, as in 2010 its different versions were humming on a whopping 86 percent of all personal computers. The second most popular operating system had only a 5 percent market share. Hence the OS market carries the signatures of a network that has undergone Bose-Einstein condensation [1].

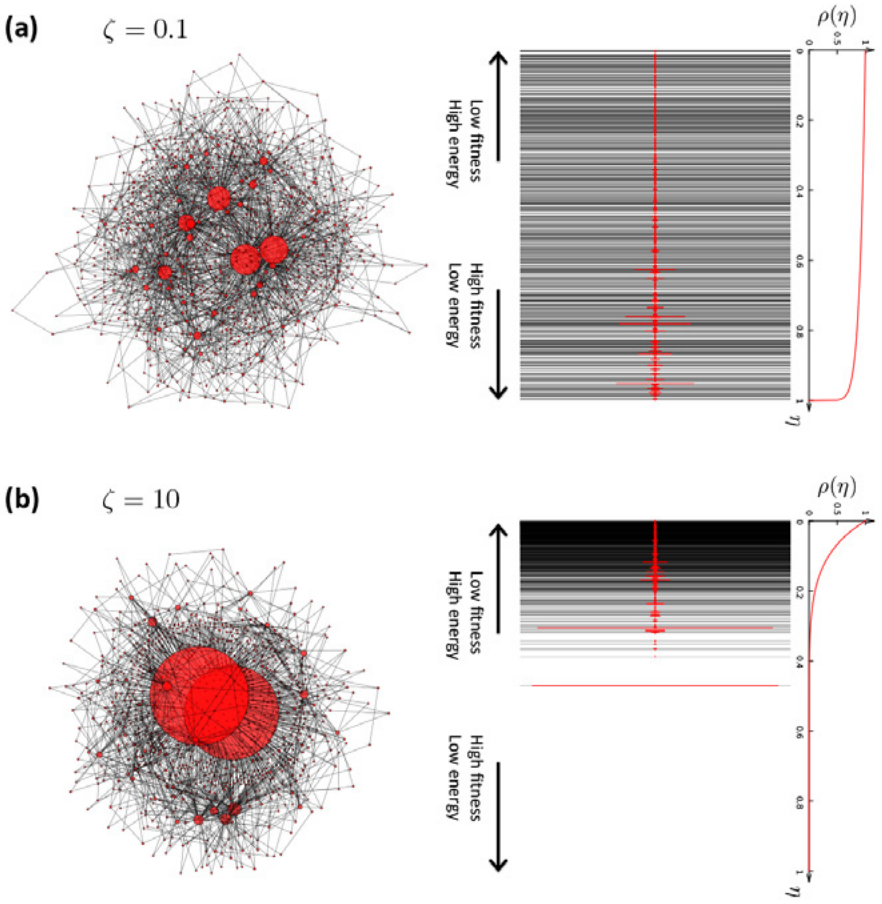



**Figure 6.8**
**Bose Einstein Condensation in Networks**

**Left panels:** (a) A scale-free network and (b) a network that has undergone a Bose-Einstein condensation generated by the fitness model with $\rho(\eta)$ following **Eq. 6.22**.

**Middle panels:** the energy levels (black lines) and the deposited particles (red dots) for a network with $m$=2 and $N$=1,000. Each energy level corresponds to the fitness of a node on the network shown in the left. Each link connected to a node is represented by a particle on the corresponding energy level.

**Right panels:** the fitness distribution $\rho(\eta)$, given by **Eq. 6.22**, illustrating the difference in the shape of the two $\rho(\eta)$ functions. The difference is determined by the parameter $\zeta$.

# EVOLVING NETWORKS

The Barabási-Albert model is a minimal model, its main purpose being to capture the core mechanisms responsible for the emergence of the scale-free property. Consequently, it has several well-known limitations:

(i) It predicts $\gamma = 3$ while the experimentally observed degree exponents vary between 2 and 4 Table 4.1.

(ii) It predicts a pure power-law degree distribution, while real systems are characterized by various deviations from a power law, like small-degree saturation or high-degree cutoff BOX 4.18.

(iii) It ignores a number of elementary processes that are obviously present in many real networks, like the addition of internal links and node or link removal.

These limitations have inspired considerable research in the network science community, clarifying how various elementary processes influence the network topology. The purpose of this section is to systematically extend the Barabási-Albert model to capture the wide range of phenomena shaping the structure of real networks.

### INITIAL ATTRACTIVENESS

In the Barabási-Albert model an isolated node cannot acquire links, as according to preferential attachment Eq. 4.1 the likelihood that a new node attaches to a $k$=0 node is strictly zero. In real networks, however, even isolated nodes acquire links. Indeed, each new research paper has a finite probability of being cited or a person that moves to a new city will quickly acquire acquaintances. In growing networks zero-degree nodes can acquire links if we add a constant to the preferential attachment function Eq. 4.1, obtaining

$$\Pi(k) \sim A + k \,. \tag{6.23}$$

In Eq. 6.23 the parameter $A$ is called initial attractiveness. As $\Pi(0) \sim A$,

initial attractiveness represents the probability that a node will acquire its first link. We can detect the presence of initial attractiveness in real networks by measuring $\Pi(k)$ Fig 6.9. Once present, initial attractiveness has two immediate consequences:

- **Increases the degree exponent:** If in the Barabási-Albert model we place Eq. 4.1 with Eq. 6.23, the degree exponent becomes [15, 16]

$$\gamma = 3 + \frac{A}{m} . \tag{6.24}$$

  By increasing $\gamma$, initial attractiveness makes a network more homogeneous and reduces the size of the hubs. Indeed, initial attractiveness adds a random component to the probability of attaching to a node. This random component favors the numerous small-degree nodes, weakening the role of preferential attachment. For high-degree nodes the $A$ term in Eq. 6.23 is negligible.

- **Generates a small-degree cutoff**: The solution of the continuum equation indicates that the degree distribution of a network governed by Eq. 6.23 does not follow a pure power-law, but has the form

$$p_k = C(k + A)^{-\gamma}. \tag{6.25}$$

Therefore, initial attractiveness induces a small-degree saturation at $k<A$. This saturation is again rooted in the fact that initial attractiveness enhances the probability that new nodes link to the small-degree nodes, which decreases the number of nodes with small $k$. For high degrees ($k \gg A$), where initial attractiveness loses its relevance, the degree distribution continues to follow a power law.

### INTERNAL LINKS

In many networks most new links are added between pre-existing nodes. For example, the vast majority of new links on the WWW are internal links, corresponding to newly added URLs between existing web documents. Similarly, virtually all new social/friendship links form between individuals that already have other acquaintances and friends. Measurements show that in collaboration networks internal links follow double preferential attachment, i.e. the probability for a new internal link to connect two nodes with degree $k$ and $k'$ is [18]

$$\Pi(k,k') \sim (A + Bk)(A' + B'k') . \tag{6.26}$$

We explore several limiting cases to understand the impact of internal links:

- **Double preferential attachment (A=A′=0):** In this case both ends of a new link are chosen proportional to the degree of the nodes they connect. Consider an extension of the Barabási-Albert model, where in each time step we add a new node with m links, followed by n internal links, each selected with probability Eq. 6.26 with $A=A'=0$. The

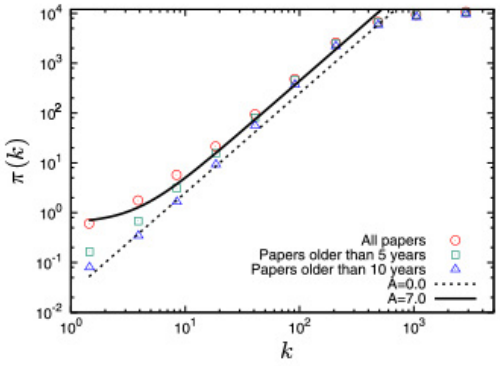



**Figure 6.9**
**Initial Attractiveness**

Cumulative preferential attachment function

$$\pi(k) = \sum_{k' \le k} \pi(k')$$

for the citation network, capturing the citation patterns of research papers published from 2007 to 2008. The $\pi(k)$ curve was measured using the methodology described in SECTION 5.7. The continuous line corresponds to $C(k+A)-\gamma$ with initial attractiveness $A \sim 7.0$. The dashed line corresponds to $A = 0$, i.e. the case without attractiveness. After [17].

degree exponent of the resulting network is [19, 20]

$$\gamma = 2 + \frac{m}{m+2n}, \tag{6.27}$$

indicating that $\gamma$ varies between 2 and 3. This means that double preferential attachment lowers the degree exponent from 3 to 2, hence increasing the network's heterogeneity. Indeed, by preferentially connecting the hubs to each other, it simultaneously helps both hubs to grow faster than they do in the Barabási-Albert model.

- **Random attachment (B=B′=0):** In this case the internal links are blind to the degree of the nodes they connect, implying that they are added between randomly chosen node pairs. Let us again consider the Barabási-Albert model, where after each new node we add n randomly selected links. In this case the degree exponent becomes [20]

$$\gamma = 3 + \frac{2n}{m}. \tag{6.28}$$

Hence, $\gamma \geq 3$ for any $n$, indicating that the resulting network will be more homogenous than the network generated by the Barabási-Albert model. Indeed, randomly added internal links mimic the process observed in random networks, making the node degrees more similar to each other.

### NODE DELETION

In many real systems nodes and links systematically disappear, leading to node or link deletion. For example, nodes are deleted from an organizational network when employees leave the company or from the WWW when web documents are removed. At the same time in some networks node removal is virtually impossible Fig. 6.10.

To explore the impact of node removal, let us start again from the Barabási-Albert model. In each time step we add a new node with $m$ links and with probability $r$ we remove a node. The observed topologies depend on the value of $r$ [23, 24, 25, 26, 27, 28]:

- **Scale-free phase:** For $r < 1$ the number of removed nodes is smaller than the number of new nodes, hence the network continues to grow. In this case the degree exponent has the value

$$\gamma = 3 + \frac{2r}{1-r}. \tag{6.29}$$

Hence, random node removal increases $\gamma$, homogenizing the network.

- **Exponential phase:** For $r=1$ the network has fixed size, as nodes arrive and are removed at the same rate (i.e. $N$=constant). In this case the network will loose its scale-free nature. Indeed, for $r \to 1$ we have $\gamma \to \infty$ in Eq. 6.29.

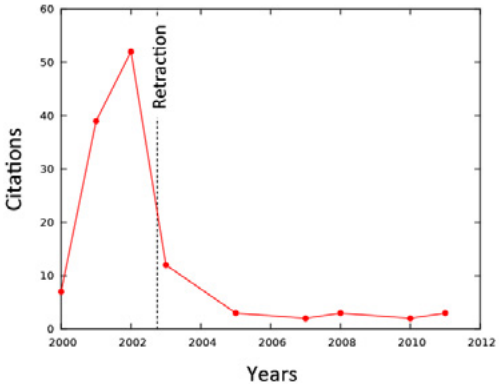



**Figure 6.10**
**The impossibility of node removal**

The citation history of a research paper by Jan Hendrik Schön published in *Science* [21]. Schön rose to prominence after a series of apparent breakthroughs in the area of semico ductors. Schön's findings were published by prominent scientific journals, like *Science* and *Nature*. His productivity was phenomenal: in 2001 he has coauthored one research paper every eight days. However, research groups around the world had difficulty reproducing some of his results.

Soon after Schön published a paper reporting a groundbreaking discovery on single-molecule semiconductors, researchers noticed that two experiments carried out at very different temperatures had identical noise [22].

The ensuing questions prompted Lucent Technologies, which ran the storied Bell Labs where Schön worked, to start a formal investigation. Eventually, Schön admitted falsifying some data to show more convincing evidence for the behavior that he observed. Several dozens of his papers, like the one whose citation pattern is shown in this figure, were retracted and the *University of Konstanz* revoked his PhD degree for "dishonorable conduct."

While the papers' retraction lead to a dramatic drop in citations, his papers continue to be cited even after their official "removal" from the literature, as shown in the figure above. This indicates that in the citation network it is virtually impossible to remove a node.

- **Declining networks:** For $r > 1$ the number of removed nodes exceeds the number of new nodes, hence the network declines BOX 6.10. Declining networks are important in several areas: Alzheimer's research focuses on the progressive loss of neurons with age and ecology focuses on the role of gradual habitat loss [29, 30, 31]. A classical example of a declining network is the telegraph, that dominated long distance communication in the second part of the 19th century and early 20th century. It was once a growing network: in the United States the length of the telegraph lines grew from 40 miles in 1846 to 23,000 in 1852. Yet, following the second World War, the telegraph gradually disappeared.

Note that node removal is not always random but can depend on the removed node's degree BOX 6.9. Furthermore, the behavior of a network can be rather complex if additional elementary processes are considered, inducing phase transitions between scale-free and exponential networks Box 6.10.

In summary, in most networks some nodes can disappear. Yet as long as the network continues to grow, its scale-free nature can persist. The degree exponent depends, however, on the detail governing the node removal process.

### ACCELERATED GROWTH

In the models discussed so far the number of links increases linearly with the number of nodes. In other words, we assumed that $L=\langle k\rangle N$, where $\langle k\rangle$ is independent of time. This is a reasonable assumption for many real networks. Yet, some real networks experience accelerated growth, meaning that the number of links grows faster than $N$, hence $\langle k\rangle$ increases. For example the average degree of the Internet increased from $\langle k\rangle$=3.42 in November 1997 to 3.96 by December 1998 [32]; the WWW increased its average degree from 7.22 to 7.86 during a five month interval [33, 34]; in metabolic networks the average degree of the metabolites grows approximately linearly with the number of metabolites [35]. To explore the consequence of such accelerated growth let us assume that in a growing network the number of links arriving with each new node follows [36, 37, 38, 39]

$$m(t) = m_0 t^{\theta}. \tag{6.30}$$

For $\theta$=0 each node has the same number of links; for $\theta$>0, however, the network follows accelerated growth. The degree exponent of the Barabási-Albert model with accelerated growth Eq. 6.30 is

$$\gamma = 3 + \frac{2\theta}{1-\theta}\,. \tag{6.31}$$

Hence, accelerated growth increases the degree exponent beyond $\gamma$=3,

making the network more homogenous. For $\theta$=1 the degree exponent diverges, leading to hyper-accelerating growth [37]. In this case $\langle k \rangle$ grows linearly with time and the network looses its scale-free nature.

### AGING

In many real systems nodes have a limited lifetime. For example, actors have a finite professional life span, capturing the period when they still act in movies. So do scientists, whose professional lifespan corresponds to the time frame they continue to publish scientific papers. These nodes do not disappear abruptly, but fade away through a slow aging process, gradually reducing the rate at which they acquire new links [40, 41, 42, 43]. Capacity limitations can induce a similar phenomena: if nodes have finite resources to handle links, once they approach their limit, they will stop accepting new links [41].

To understand the impact of aging let us assume that the probability that a new node connects to node $i$ is $\Pi(k, t-t_i)$, where ti is the time node $i$ was added to the network. Hence, $t-t_i$ is the node's age. In analytical calculations slow aging is often modeled by choosing [40]

$$\Pi(k,t-t_i) \sim k(t-t_i)^{-\nu}, \tag{6.32}$$

where $\nu$ is a tunable parameter governing the dependence of the attachment probability on the node's age. Depending on the value of $\nu$ we can distinguish three scaling regimes:

- **For negative $\nu$** the older is node $i$, the more likely that a new node will link to it. Hence, $\nu < 0$ enhances the role of preferential attachment. In the extreme case $\nu \to -\infty$, each new node will only connect to the oldest node, resulting in a hub-and-spoke topology Fig. 6.11a. The calculations show that the scale-free state persists in this regime, but the degree exponent drops under 3. Hence, $\nu < 0$ makes the network more heterogeneous

- **A positive $\nu$** will encourage the new nodes to attach to younger nodes. In the extreme case $\nu \to \infty$ each node will connect to its immediate predecessor Fig. 6.11a. We do not need a very large $\nu$ to experience the impact on aging: the degree exponent diverges as we approach $\nu$=1. Hence gradual aging homogenizes the network by shadowing the older hubs.

- **For $\nu > 1$** the aging effect overcomes the role of preferential attachment, leading to the loss of the scale-free property.

In summary, the results discussed in this section indicate that a wide range of elementary processes can affect the structure of a growing network Table 6.1. These results highlight the true power of the evolving network paradigm: it allows us to address, using a mathematically predictive framework, the impact of a wide range of elementary processes on the network topology and evolution.

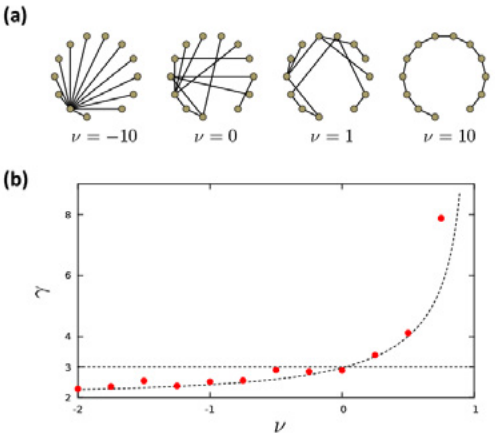



**Figure 6.11**
**The impact of aging**

**(a)** A schematic illustration of the expected network topologies for various aging exponents $\nu$. In the context of a growing network model we assume that the probability to attach to a node is proportional to $kr^{-\nu}$, where $r$ is the age of the node. For negative $\nu$ nodes prefer the oldest nodes, turning the network into a hub-and-spoke topology. For positive $\nu$ the most recent nodes are the most attractive. Hence for large $\nu$ the network turns into a chain, as the last (hence the youngest) node is the most attractive for the new node. The network is shown for $m$=1 for clarity but note that the degree exponent is independent of $m$.

**(b)** The degree exponent $\gamma$ vs the aging exponent $\nu$, as predicted by the analytical solution of the rate equation. The red symbols are the result of simulations, each representing a single network of $N$=10,000 and $m$=1. The degree exponent is estimated using the method described in **CHAPTER 4**. Redrawn after Ref. [40].

# BOX 6.8

## DECLINING NETWORKS AND FASHION

The properties of declining networks is well illustrated by the New York City garment industry, whose nodes are designers and contractors that are connected to each other by the annual coproduction of lines of clothing. As the industry decayed, the network has persistently shrunk. This is illustrated by the fate of the largest connected component, that collapsed from 3,249 nodes in 1985 to 190 nodes in 2003. Interestingly, the network's degree distribution remained relatively unchanged during this period. The analysis of the network's evolution allowed researchers to uncover several interesting properties of declining networks [23]:

- **Preferential Attachment:** While overall the network was shrinking with time, new nodes continued to arrive. The measurements indicate that the attachment probability of these new nodes follows $\Pi(k) \sim k^{\alpha}$ with $\alpha=1.20 \pm 0.06$ **PANEL A**, offering evidence of superlinear preferential attachment.

- **Link deletion:** The measurements also show that the probability that a firm looses a link decreased proportionally with the firms' degree, as $k(t)-\eta$ with $\eta=0.41\pm0.04$. This documents a weak-gets-weaker phenomenon where the less connected firms are more likely to loose their links.

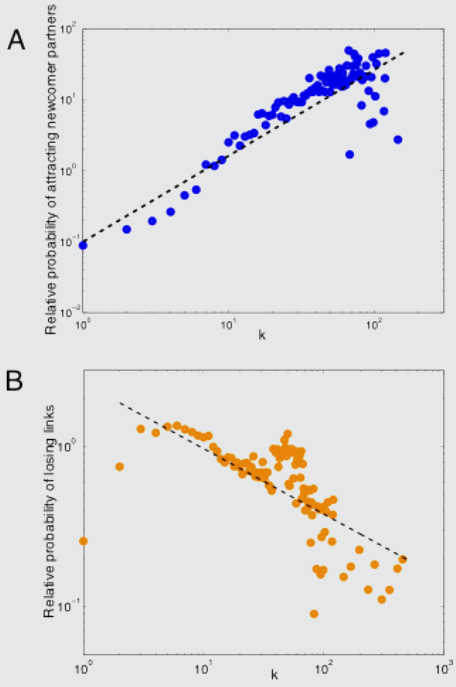




**Figure 6.12**
**The decline of the garment industry**

**(a)** Preferential attachment. The relative probability $\Pi(k)$ that a newcomer firm added at time t connects to an incumbent firm with $k$ links. The dashed line has slope $\alpha=1.2$.

**(b)** Link deletion. The relative probability, $R_{k(t)}$ of deleting a link from a degree node, compared with random link removal. The dashed line has slope $\eta=0.41$.

If link addition and removal were to be random, we would expect $\Pi(k) \sim 1$ and $R_{k(t)} \sim 1$ for all $k$. After [23].

**Figure 6.13**
**Garment district**

The Garment District is a Manhattan neighborhood located between Fifth Avenue and Ninth Avenue, from 34th to 42nd Street. Since the early 20th century it has been the center for fashion manufacturing and design in the United States. The Needle threading a button sculpture and a sculpture of a tailor, located in the heart of the district, pay tribute to the neighborhood's past and present.

| PROCESS | PROCESS | $\gamma$ | OBSERVATIONS |
|---|---|---|---|
| Preferential attachment | $\Pi(k) \sim k$ | 3 | |
| Initial attractiveness | $\Pi(k) = A + k$ | $3 + \frac{A}{m}$ | Small degree cutoff $p_k \sim (k + A)^{-\gamma}$ |
| Internal links | $\Pi(k, k') = (A + Bk)(A' + B'k')$ | $2 + \frac{m}{m + 2n}$ | Double preferential attachment $A = A' = 0, B \quad 0, B' \quad 0$ |
| | | $3 + \frac{2n}{m}$ | Random internal attachment $B = B' = 0$ |
| Node (link) deletionN | ode removal rate $r$ | $3 + \frac{2r}{1 - r}$ | Scale-free for $r < r^*$<br>Stretched exponential for $r = r^*$<br>Exponential for $r > r^*$ |
| Accelerated growth | $m(t) = t^{\Theta}$ | $\frac{3 - \theta}{1 - \theta}$ | For $\theta = 1$ we have hyper-accelerated growth and the scale-free state disappears. |
| Aging | $\Pi(k) \sim (t - t_i)^{-\nu}$ | See Figure 6.11 | For $\nu > 1$ the network looses its scale-free topology. |

**Table 6.1 Elementary processes**
A summary of the various elementary processes discussed in this section and their impact on the degree distribution.

# BOX 6.9

## NODE REMOVAL INDUCED PHASE TRANSITIONS

The coexistence of node removal with other elementary processes can lead to interesting topological phase transitions. This is illustrated by a simple model in which the network's growth is governed by Eq. 6.23, i.e. preferential attachment with initial attractiveness, and we also remove nodes with rate *r*. The network displays three distinct phases, captured by the phase diagram shown below:

**Subcritical node removal ($r < r^*(A)$):** If the rate of node removal is under a critical value $r^*(A)$, the network will be scale-free.

**Critical node removal ($r = r^*(A)$)**: Once r reaches a critical value $r^*(A)$, the degree distribution turns into a stretched exponential SECT. 4.A.

**Exponential networks ($r > r^*(A)$):** In this regime the network looses its scale-free nature, developing an exponential degree distribution.

Therefore, the coexistence of multiple elementary processes in a network can lead to discontinuous changes in the network topology. To be specific, a continuous increase in the node removal rate leads to a transition from a scale-free to an exponential network.

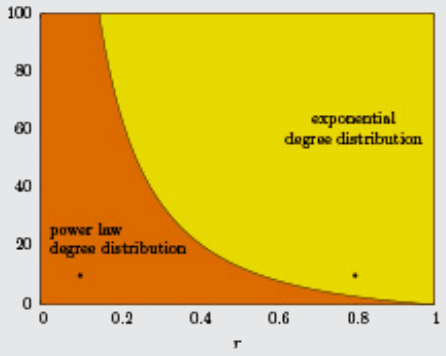



**Figure 6.14**
**Scaling under node deletion**

The degree distribution of a network whose growth is driven by preferential attachment with initial attractiveness A and node removal rate *r*. After [28].

# SUMMARY

As we illustrated in this chapter, rather diverse processes, from node fitness to internal links or aging, can influence the topology of real networks. By exploring these processes, we came to see how to use the evolving network modeling framework to accurately predict the impact of various frequently encountered elementary events on a network's topology and evolution. The most important conclusion of the examples discussed in this chapter is that *if we want to understand the structure of a network we must first get its dynamics right. The topology is the bonus of this approach.*

In CHAPTER 4 we documented the difficulties we encounter when we attempt to fit a pure power law to the degree distribution of real networks. The roots of this problem were revealed in this chapter: if we account for the detailed dynamical processes that contribute to the evolution of real networks, we expect systematic deviations from a pure power law. Indeed, the previous sections predicted several analytical forms for the degree distribution:

- **Power law:** A pure power law emerges if a growing network is governed by preferential attachment only, lacking nonlinearities or initial attractiveness. In its pure form a power law is observed only in a few systems. Yet, it is the starting point for understanding the degree distribution of most real networks.

- **Stretched exponential:** If preferential attachment is sublinear, the degree distribution follows a stretched exponential SECTION 5.7. A stretched exponential degree-distribution can also appear under node removal at the critical point SECTION 6.5.

- **Fitness-induced corrections:** In the presence of fitnesses the precise form of $p_k$ depends on the fitness distribution $\rho(\eta)$, which determines $p_k$ via Eq. 6.6. For example, a uniform fitness distribution induces a logarithmic correction in $p_k$ Eq. 6.8. Other forms of $\rho(\eta)$ can lead to rather exotic $p_k$.

- **Small-degree cutoffs:** Initial attractiveness adds a random component to preferential attachment. Consequently, the degree distribution develops a small-degree saturation.

- **Exponential cutoffs:** Node and link removal, present in many real systems, can also induce exponential cutoffs in the degree distribution. Furthermore, random node-removal can deplete the small-degree nodes, inducing a peak in $p_k$.

In most networks several of the elementary processes discussed in this chapter appear together. For example, in scientific collaboration networks we have sublinear preferential attachment with initial attractiveness and links can be both external and internal. As researchers have different creativity, fitness also plays a role, requiring us to know the appropriate fitness distribution. Therefore, the degree distribution is expected to display small degree saturation (thanks to initial attractiveness), stretched exponential cutoff at high degrees (thanks to sublinear preferential attachment), and some unknown corrections due to the particular form of the fitness distribution $\rho(\eta)$. These findings indicate that if our goal is to obtain an accurate fit to the degree distribution, we first need to build a generative model that analytically predicts the expected functional form of $p_k$. Yet, in many systems developing an accurate theory for $p_k$ may be an overkill. Hence, it is often sufficient to establish if we are dealing with a broad or a bounded degree distribution SECTION 4.9, as the system's properties will be primarily driven by this distinction.

The results of this chapter also allow us to reflect on the role of the various network models. We can categorize these models into three main classes Table 6.2:

**Static Models**: The random network model of Erdős and Rényi CHAPTER 3 and the small world network model of Watts and Strogatz Fig. 3.15 have a fixed number of nodes, prompting us to call them static. They both assume that the role of the network modeler is to cleverly place the links between the nodes. Both models predict a bounded degree distribution.

**Generative Models:** The configuration and the hidden parameter models discussed in SECTION 4.8 generate networks with some predefined degree distribution. Hence, these models are not mechanistic, in the sense that they do not tell us why a network develops a particular degree distribution. Rather, they help us understand how various network properties, from clustering to path lengths, depend on the degree distribution.

**Evolving Network Models:** These models aim to capture the mechanisms that govern the time evolution of a network. The most studied example in the Barabási-Albert model, but equally important are the extenions discussed in this chapter, from the Bianconi-Barabási mod-

el to models involving internal links, aging, node and link deletion, or accelerated growth. These models are motivated by the hypothesis that if we correctly capture all microscopic processes that contribute to a network's evolution, then the network's large-scale characteristics follow from that. There is an important role in network theory for each three modeling frameworks. If our interest is limited to the role of the network environment on some phenomena, like spreading processes or network robustness, the generative models offer an excellent starting point. If, however, we want to understand the origin of a certain network property, we must resort to evolving network models, that capture the processes that have built the network in the first place.

Finally, the results of this chapter allow us to formulate our next network law:

**The fifth law, the role of diversity.**

*With time the fittest nodes turn into the largest hubs.*

**A. Quantitative formulation**

Eq. 6.4 offers the quantitative formulation of the fifth law, predicting that the dynamical exponent, capturing the rate at which a node acquires links, is proportional to the node's fitness. Hence the higher a node's fitness, the higher the rate it acquires links. Consequently, with time the nodes with the highest fitness will turn into the largest hubs.

**B. Universality**

In most networks nodes with different qualities and capabilities compete for links. Hence node fitness, capturing a node's ability to attract links, is present in most real networks.

**C. Non-random origin**

The dynamics of fitness-driven networks is quite different from the dynamics of the random network model, in which nodes acquire links at comparable rate. Hence, the properties of these networks cannot be explained within the random network framework.

| Name | Static Models | Generative Models | Mechanistic Models |
|---|---|---|---|
| Example | Erdős-Rényi<br>Watts-Strogatz | Configuration Model<br>Hidden Parameter Model | Barabási-Albert<br>Fitness Model |
| Characteristics | $N$ fixed<br>$L$ variable<br>$p_k$ bounded | Pre-defined, arbitrary $p_k$. | $p_k$ depends on the nature of the processes that contribute to the networks evolution. |

**Table 6.2**
**Models of network science**

The table shows the three main modeling frameworks we encountered so far, together with their main distinguishing features.

# HOMEWORK

**1.** Calculate the degree exponent and the dynamical exponent for a growing network with two distinct fitnesses. To be specific, let us assume that the fitnesses follow the double delta distribution

$$\rho(\eta) = \delta(\eta - a) + \delta(\eta - 1) \quad \text{with } 0 \le a \le 1. \tag{6.33}$$

Discuss how the degree exponent depends on the parameter $a$.

**2.** Calculate the degree exponent of the directed Barabási-Albert model with accelerated growth, i.e. when $m(t)=t^{\odot}$.

**3.** Assume that a network is driven by a preferential attachment with additive fitness, $\pi(k_i) \sim A_i+k_i$, where $A_i$ is chosen from a $\rho(A_i)$ distribution [44]. Calculate and discuss the degree distribution of the resulting network.

# ADVANCED TOPICS 6.A
# SOLVING THE FITNESS MODEL

The purpose of this section is to derive the degree distribution of the fitness model [2, 13, 14]. We start by calculating the mean of the sum $\sum_j \eta_j k_j$ over all possible realizations of the quenched fitnesses $\eta$. Since each node is born at a different time $t_0$, we can write the sum over $j$ as an integral over $t_0$

$$\left\langle \sum_j \eta_j k_j \right\rangle = \int d\eta \rho(\eta) \eta \int_1^t dt_0 k_\eta(t, t_0) \tag{6.34}$$

By replacing $k_\eta(t, t_0)$ with Eq. 6.3 and performing the integral over $t_0$, we obtain

$$\left\langle \sum_j \eta_j k_j \right\rangle = \int d\eta \rho(\eta) \eta m \frac{t - t^{\beta(\eta)}}{1 - \beta(\eta)} \,. \tag{6.35}$$

The dynamic exponent $\beta(\eta)$ is bounded, i.e. $0<\beta(\eta)<1$ because a node can only increase its degree with time $(\beta(\eta)>0$ and $k_i(t)$ cannot increase faster than $t(\beta(\eta)<1)$. Therefore in the limit $t\to\infty$ in Eq. 6.35 the term $t^{\beta(\eta)}$ can be neglected compared to $t$, obtaining

$$\left\langle \sum_j \eta_j k_j \right\rangle \overset{t\to\infty}{=} Cmt(1 - O(t^{-\varepsilon})) \,, \tag{6.36}$$

where $\varepsilon = (1 - \max_\eta \beta(\eta)) > 0$ and

$$C = \int d\eta \rho(\eta) \frac{\eta}{1 - \beta(\eta)} \,. \tag{6.37}$$

Using Eq. 6.36, and the notation $k_\eta = k_{\eta i}(t, t_0)$, the dynamic equation Eq. 6.2 can be written as

$$\frac{\partial k_\eta}{\partial t} = \frac{\eta k_\eta}{Ct} \,, \tag{6.38}$$

which has a solution of the form Eq. 6.3, given that

$$\beta(\eta) = \frac{\eta}{C} \,, \tag{6.39}$$

confirming the self-consistent nature of the assumption Eq. 6.3. To complete the calculation we need to determine $C$ from Eq. 6.37. After substituting $\beta(n)$ with $\eta/C$ , we obtain

$$1 = \int_0^{\eta_{\max}} d\eta \rho(\eta) \frac{1}{\frac{C}{\eta} - 1}, \tag{6.40}$$

where $\eta_{max}$ is the maximum possible fitness in the system. The integral Eq. 6.40 is singular. However, since $\beta(\eta)=\eta/c < 1$ for any $\eta$, we have $C > \eta_{max}$, thus the integration limit never reaches the singularity. Note also that, since

$$\sum_j \eta_j k_j \le \eta_{\max} \sum_j k_j = 2mt\eta_{\max} \tag{6.41}$$

we have $C \le \eta_{max}$.

If there is a single dynamic exponent $\beta$, the degree distribution should follow the power law $p_k \sim k^{-\gamma}$ , where the degree exponent is given by $\gamma=1/\beta+1$. However, in the fitness model we have a spectrum of dynamic exponents $\beta(\eta)$, thus $p_k$ is given by a weighted sum over different power-laws. To find $p_k$ we need to calculate the cumulative probability that a randomly chosen node's degree satisfies $k_\eta(t)>k$. This cumulative probability is given by

$$P(k_\eta(t) > k) = P\left(t_0 < t\left(\frac{m}{k}\right)^{C/\eta}\right) = t\left(\frac{m}{k}\right)^{C/\eta}. \tag{6.42}$$

Thus, the degree distribution is given by the integral

$$P_k = \int_{\eta_{max}}^{0} d\eta \frac{\partial P(k_\eta(t) > k)}{\partial t} \propto \int d\eta \rho(\eta) \frac{C}{\eta} \left(\frac{m}{k}\right)^{\frac{C}{\eta}+1}. \tag{6.43}$$

# BIBLIOGRAPHY

[1] A.L. Barabási, Linked: The New Science of Networks. (Perseus, Boston) 2001.

[2] G. Bianconi and A.-L. Barabási, Competition and multiscaling in evolving networks, Europhysics Letters 54: 436-442, 2001.

[3] A.-L. Barabási, R. Albert, H. Jeong, and G. Bianconi. Power-law distribution of the world wide web, Science 287: 2115, 2000.

[4] P. L. Krapivsky and S. Redner. Statistics of changes in lead node in connectivity-driven networks, Phys. Rev. Lett. 89:258703, 2002.

[5] C. Godreche and J. M. Luck. On leaders and condensates in a growing network, J. Stat. Mech. P07031, 2010.

[6] J. H. Fowler, C. T. Dawes, and N. A. Christakis. Model of Genetic Variation in Human Social Networks, PNAS 106: 1720-1724, 2009.

[7] M. O. Jackson. Genetic influences on social network characteristics, PNAS 106:1687–1688, 2009.

[8] S. A. Burt. Genes and popularity: Evidence of an evocative gene environment correlation, Psychol. Sci. 19:112–113, 2008.

[9] J. S. Kong, N. Sarshar, and V. P. Roychowdhury. Experience versus talent shapes the structure of the Web, PNAS 105:13724-9, 2008.

[10] A.-L. Barabási, C. Song, and D. Wang. Handful of papers dominates citation, Nature 491:40, 2012.

[11] D. Wang, C. Song, and A.-L. Barabási. Quantifying Long term scientific impact, preprint, 2013.

[12] M. Medo, G. Cimini, and S. Gualdi. Temporal effects in the growth of

networks, Phys. Rev. Lett., 107:238701, 2011.

[13] G. Bianconi and A.-L. Barabási. Bose-Einstein condensation in complex networks, Phys. Rev. Lett. 86: 5632–5635, 2001.

[14] C. Borgs, J. Chayes, C. Daskalakis, and S. Roch. First to market is not everything: analysis of preferential attachment with fitness, STOC'07, San Diego, California, 2007.

[15] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Structure of growing networks with preferential linking, Phys. Rev. Lett. 85: 4633, 2000.

[16] C. Godreche, H. Grandclaude, and J. M. Luck. Finite-time fluctuations in the degree statistics of growing networks, J. of Stat. Phys. 137:1117-1146, 2009.

[17] Y.-H. Eom and S. Fortunato. Characterizing and Modeling Citation Dynamics, PLoS ONE 6(9): e24926, 2011.

[18] A.-L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, and T. Vicsek, Evolution of the social network of scientific collaborations, Physica A 311: 590-614, 2002.

[19] R. Albert, and A.-L. Barabási. Topology of evolving networks: local events and universality, Phys. Rev. Lett. 85:5234-5237, 2000.

[20] G. Goshal, L. Ping, and A.-L Barabási, preprint, 2013.

[21] J. H. Schön, Ch. Kloc, R. C. Haddon, and B. Batlogg. A superconducting field-effect switch, Science 288: 656–8. 2000.

[22] D. Agin. Junk Science: An Overdue Indictment of Government, Industry, and Faith Groups That Twist Science for Their Own Gain (Macmillan; New York), 2007.

[23] S. Saavedra, F. Reed-Tsochas, and B. Uzzi. Asymmetric disassembly and robustness in declining networks, PNAS105:16466–16471, 2008.

[24] F. Chung and L. Lu. Coupling on-line and off-line analyses for random power-law graphs, Int. Math. 1: 409-461, 2004.

[25] C. Cooper, A. Frieze, and J. Vera. Random deletion in a scalefree random graph process, Int. Math. 1, 463-483, 2004.

[26] S. N. Dorogovtsev and J. Mendes. Scaling behavior of developing and decaying networks, Europhys. Lett. 52: 33-39, 2000.

[27] C. Moore, G. Ghoshal, and M. E. J. Newman. Exact solutions for models of evolving networks with addition and deletion of nodes, Phys. Rev. E 74: 036121, 2006.

[28] H. Bauke, C. Moore, J. Rouquier, and D. Sherrington, Topological phase transition in a network model with preferential attachment and node removal, The European Physical Journal B: 83: 519-524, 2011.

[29] M. Pascual and J. Dunne, (eds) Ecological Networks: Linking Structure to Dynamics in Food Webs (Oxford Univ Press, Oxford), 2005.

[30] R. Sole and J. Bascompte. Self-Organization in Complex Ecosystems (Princeton Univ Press, Princeton), 2006.

[31] U. T. Srinivasan, J. A. Dunne, J. Harte, and N. D. Martinez. Response of complex food webs to realistic extinction sequencesm, Ecology 88:671–682, 2007.

[32] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology, ACM SIGCOMM Computer Communication Review 29: 251-262, 1999.

[33] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, and A. Tomkins. Graph structure in the web, Computer networks 33: 309-320, 2000.

[34] J. Leskovec, J. Kleinberg, and C. Faloutsos, Graph evolution: Densification and shrinking diameters, ACM TKDD07, ACM Transactions on Knowledge Discovery from Data (2007), 1(1).

[35] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, The large-scale organization of metabolic networks, Nature 407: 651–655, 2000.

[36] S. Dorogovtsev and J. Mendes. Effect of the accelerating growth of communications networks on their structure, Phys. Rev. E 63: 025101(R), 2001.

[37] M. J. Gagen and J. S. Mattick. Accelerating, hyperaccelerating, and decelerating networks, Phys. Rev. E 72: 016123, 2005.

[38] C. Cooper and P. Prałat. Scale-free graphs of increasing degree, Random Structures & Algorithms 38: 396–421, 2011.

[39] N. Deo and A. Cami. Preferential deletion in dynamic models of web-like networks, Inf. Proc. Lett. 102: 156-162, 2007.

[40] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks with aging of sites, Phys. Rev. E, 62:1842, 2000.

[41] A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley, Classes of small-world networks. Proc. National Academy of Sciences USA 97: 11149, 2000.

[42] K. Klemm and V. M. Eguiluz. Highly clustered scale free networks,

Phys. Rev. E 65: 036123, 2002.

[43] X. Zhu, R. Wang, and J.-Y. Zhu. The effect of aging on network structure, Phys. Rev. E 68: 056121, 2003.

[44] G. Ergun and G. J. Rodgers, Growing random networks with fitness, Physica A 303: 261-272, 2002.

# CHAPTER 7

# DEGREE CORRELATIONS



**Figure 6.0 (front cover)**
**Network representation by Mauro Martino**

# INTRODUCTION

Angelina Jolie and Brad Pitt, Ben Affleck and Jennifer Garner, Harrison Ford and Calista Flockhart, Michael Douglas and Catherine Zeta-Jones, Tom Cruise and Katie Holmes, Richard Gere and Cindy Crawford. An odd list, yet instantly recognizable to those immersed in the headline-driven world of celebrity couples. They are Hollywood stars that are or were married in the past. Their weddings (and breakups) sold millions of magazines and drawn countless hours of media coverage. Thanks to them we take for granted that celebrities marry each other. We rarely pause to ask: is this normal? What is the expected chance that a celebrity marries another celebrity?

Assuming that a celebrity could date anyone from a pool of about a billion ($10^9$) eligible individuals, the chances that their mate would be another celebrity from a generous list of 1,000 other celebrities is only $10^{-6}$. Therefore, if dating is driven by random encounters, celebrities would never marry each other. Yet, they do, with some puzzling implications.

Even if you do not care about the dating habits of celebrities, we must pause and explore what this phenomenon tells us about the structure of the social network. Hollywood celebrities, political leaders, and CEOs of major corporations tend to know an exceptionally large number of individuals and are known by even more. They are hubs. Hence celebrity dating is a manifestation of an interesting property of social network: hubs tend to have ties to other hubs.

As obvious this may sound, this property is not present in all networks. Consider for example the protein-interaction network of yeast, shown in Fig. 7.2. Each node corresponds to a protein and a link between two proteins indicates a binding interaction. A quick inspection of the network reveals its scale-free nature: numerous one- and two-degree proteins coexist with a few highly connected hubs. These hubs, however, tend avoid linking to *each other.* They link instead to many small-degree nodes, generating a hub-and-spoke pattern. This is particularly obvious for the two hubs highlighted in Fig. 7.2: they almost exclusively interact with small-degree proteins while avoiding linking to each other.

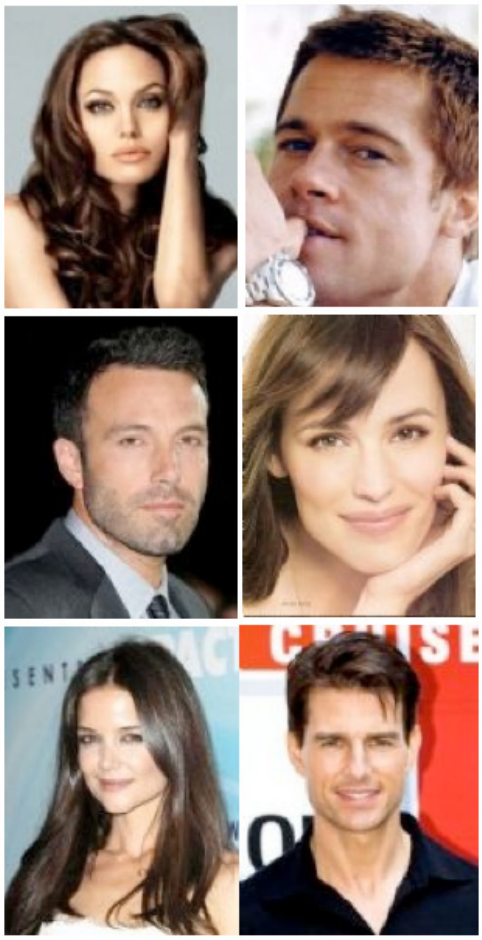


**Figure 7.1**
**Hubs Dating Hubs**

Celebrity couples, offering a vivid demonstration that in social networks hubs tend to know, date and marry each other (Images from http://www.whosdatedwho.com).

A brief calculation illustrates how unusual this pattern is. Let us assume that each node chooses randomly the nodes it connects to. Therefore the probability that two nodes with degree $k$ and $k'$ link to each other is

$$p_{k,k'} = \frac{kk'}{2L}. \qquad (7.1)$$

Eq. 7.1 tells us that hubs, by the virtue of the many links they have, are much more likely to connect to each other than to small degree nodes. Yet, the hubs highlighted in Fig. 7.2 almost exclusively connect to degree one nodes. By itself this is not unexpected: Fig. 7.1 also predicts that a hub with $k$ = 56 connections should link to $N_1 P_{1,56} \approx 12$ nodes with degree 1. The problem is that this hub connects to 46 degree one neighbors, i.e. four times the expected number.

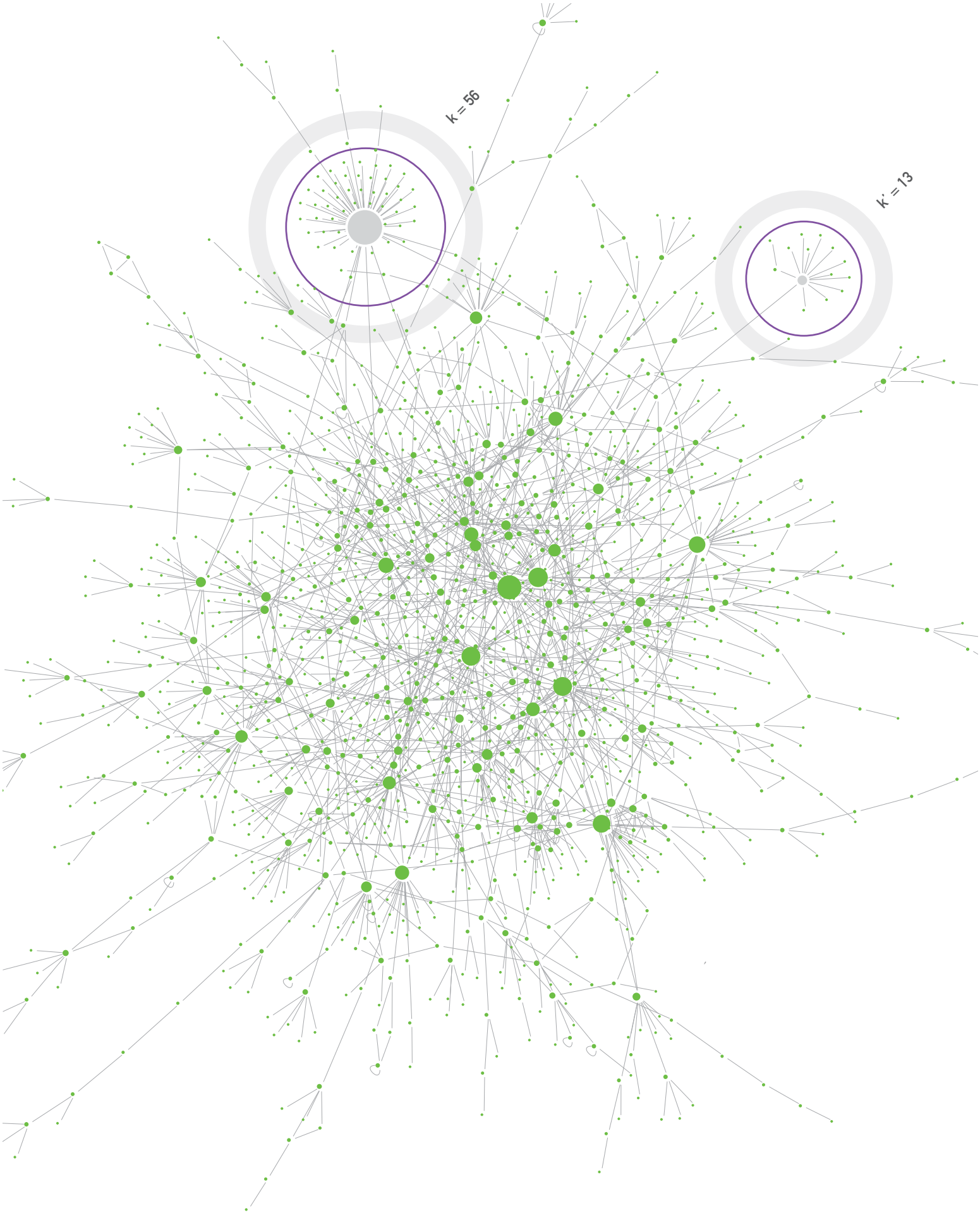
Furthermore, the likelihood that two largest hubs with degrees $k$=56 and $k'$ = 13 have a direct link between them Fig. 7.2, is $p_{k_1,k'} = 0.15$, which is 400 times larger than $p_{1,2} = 0.0004$, the likelihood that a degree-two node links to a degree-one node. Yet, there are no direct links between the hubs in Fig. 7.2, but we observe numerous direct links between small degree nodes.

In summary, while in social networks hubs tend to "date" each other, in the protein interaction network the opposite is true: the hubs avoid linking to other hubs. While it is dangerous to extrapolate generic principles from two examples, the purpose of this chapter is to show that these patterns are manifestations of a general property of real networks: they exhibit a phenomena called *degree correlations*. We discuss how to measure such degree correlations and explore their impact on the network topology.

**Figure 7.2 (folowing page)**
**Hubs Avoiding Hubs**

The protein interaction map of yeast. Each node corresponds to a protein and two proteins are linked if there is experimental evidence that they can bind to each other in the cell. The two largest hubs, with degrees $k$ = 56 (left) and $k'$ = 13 (right) are highlighted in the figure. They both connect to many small degree nodes and avoid linking to each other.

The network has $N$ = 1,870 proteins connected by $L$ = 2,277 links, representing one of the earliest protein interaction maps [1, 2]. Only the largest component is shown. Note that the protein interaction network of yeast, discussed in **TABLE 4.1**, represents a later, more detailed map. Hence, it contains more nodes and links than the network shown in this figure. Redrawn after [3].


k = 56
k' = 13

# ASSORTATIVITY AND DISASSORTATIVITY

Just by the virtue of the many links they have, hubs are expected to link to each other. Yet, as we have seen in the previous section, in some networks they do, in others they don't. This is illustrated in Fig. 7.3, that shows three networks with identical degree sequence but different topology:

- **Neutral Network**
  Fig. 7.3b shows a network whose wiring is truly a random. The network of Fig. 7.3b is *neutral*, meaning that the number of links between the hubs coincides with what we expect by chance, as predicted by Eq. 7.1. For clarity we highlighted in red the five largest nodes and the direct links between them, observing a few red links as the likelihood that two nodes link to each other increases with their degree.

- **Assortative Network**
  The network of Fig. 7.3a has precisely the same degree sequence as the one in Fig. 7.3b. Yet, there is a noticeable difference between the two networks: the hubs in Fig. 7.3a tend to link to each other, while avoiding linking to small-degree nodes. At the same time the small-degree nodes tend to connect to other small-degree nodes. Networks displaying such trends are *assortative*. An extreme manifestation of this pattern is a perfectly assortative network, in which degree-$k$ nodes connect only to other degree-$k$ nodes Fig. 7.4.

- **Disassortative Network**
  We observe the opposite trend in Fig. 7.3c, where the hubs completely avoid each other, linking mainly to small-degree nodes. Consequently the network displays a hub and-spoke character, making it *disassortative*.

In general a network displays degree correlations if the number of links between the high and low-degree nodes is systematically different from what is expected by chance. In other words, in correlated networks the number of links between nodes of degrees $k$ and $k'$ deviates from Eq. 7.1.

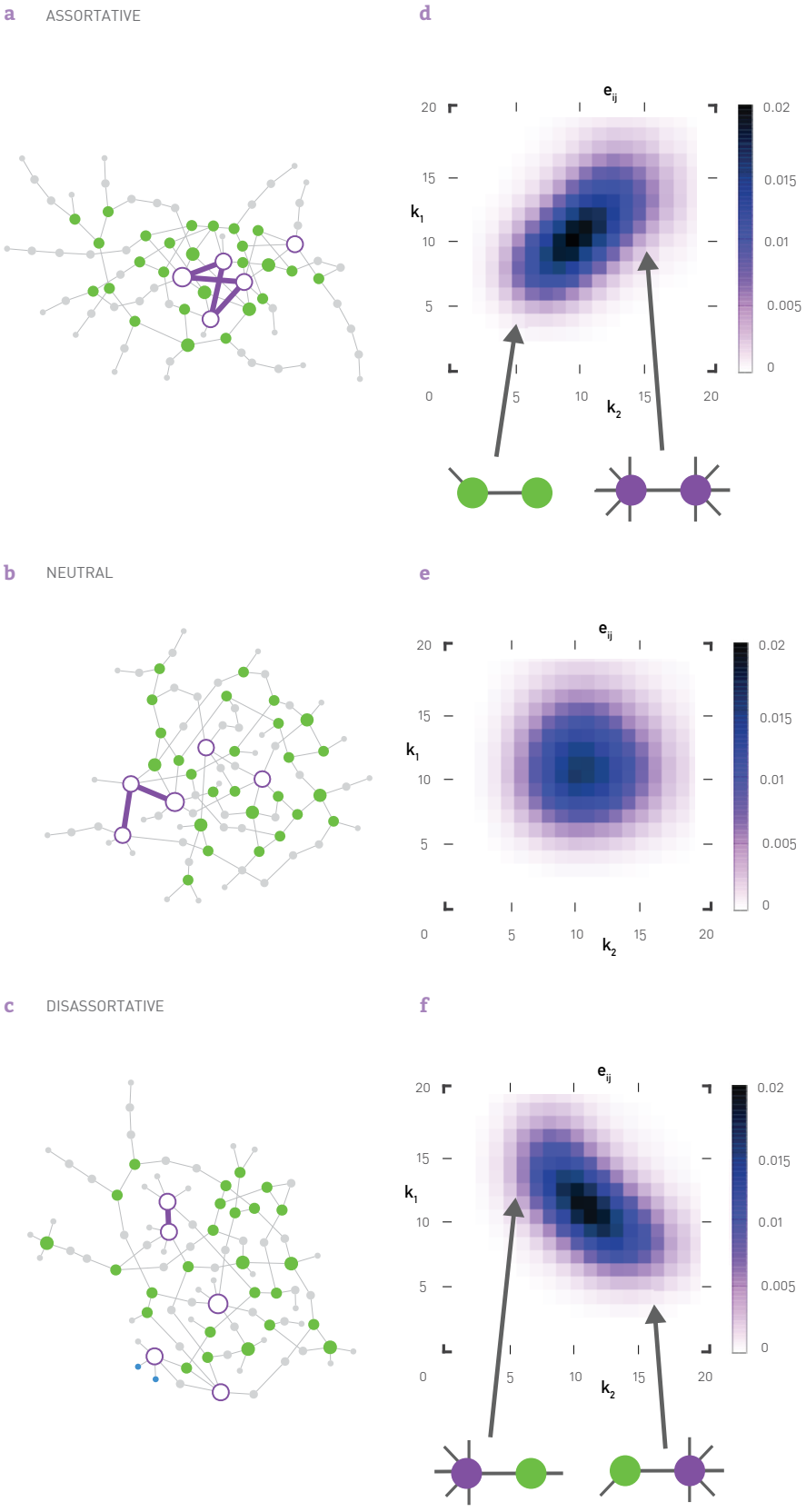



**Figure 7.3**

**Degree correlation matrix**

**a, b, c** Three networks that have precisely the same degree distribution (Poisson $p_k$), but display different degree correlations. We show only the largest component and we highlight the five nodes with the highest degree in red, together with the direct links between them.

**d, e, f** The degree correlation matrix $e_{ij}$ for (d) an assortative, (e) a neutral and (f) a disassortative network with Poisson degree distribution and $N$=1,000, and $\langle k \rangle$=10. The colors correspond to the probability that there is a link between nodes with degrees $k_1$ and $k_2$.

**a, d** For assortative networks $e_{ij}$ takes higher values along the main diagonal. This indicates that nodes of similar degree tend to link to each other: small-degree nodes to small-degree nodes and hubs to hubs. The network in (a) illustrates this by having numerous links between its hubs.

**b, e** In neutral networks nodes link to other nodes randomly. Hence, the density of links is symmetric around the average degree, indicating the lack of correlations in the linking pattern.

**c, f** In disassortative networks $e_{ij}$ is higher along the secondary diagonal, indicating that hubs tend to connect to small-degree nodes, and small-degree nodes to hubs. This is illustrated by the hub and spoke character of the network in (c).

The complete information about potential degree correlations is contained in the degree correlation matrix, $e_{ij}$, which is the probability of finding *a* node with degrees *i* and *j* at the two ends of a randomly selected link. As $e_{ij}$ is a probability, it obeys the normalization condition

$$\sum_{i,j} e_{ij} = 1. \tag{7.2}$$

In SECT. 5.8 we derived the probability $q_k$ that there is a degree-*k* node at the end of the randomly selected link Eq. 5.29.

$$q_k = \frac{k p_k}{\langle k \rangle} \tag{7.3}$$

We can connect $q_k$ to $e_{i,j}$ via

$$\sum_j e_{ij} = q_i. \tag{7.4}$$

In neutral networks, we expect

$$e_{ij} = q_i q_j. \tag{7.5}$$

A network displays degree correlations if eij deviates from the random expectation captured by Eq. 7.5, Eqs. 7.2 - 7.5 are valid for networks with an arbitrary degree distribution, hence they apply to both random and scale-free networks. Given that $e_{ij}$ contains the complete information about potential degree correlations, we start with its visual inspection. Figs. 7.3 d, e, f shows $e_{ij}$ for an assortative, a neutral and a disassortative network. In a neutral network small and high-degree nodes connect to each other randomly, hence $e_{ij}$ lacks any trend Fig. 7.3e. In contrast, assortative networks show high correlations along the main diagonal, indicating that nodes predominantly connect to other nodes with comparable degree. Therefore low-degree nodes tend to link to other low-degree nodes and hubs to hubs Fig. 7.3d. In disassortative networks $e_{ij}$ displays the opposite trend: it has high correlations along the secondary diagonal, indicating that high-degree nodes tend to connect to low-degree nodes Fig. 7.3f.

In summary information about degree correlations is carried by the degree correlation matrix $e_{ij}$. Yet, the study of degree correlations through the inspection of $e_{ij}$ has numerous disadvantages:

- It is difficult to extract information from the visual inspection of a matrix.
- Unable to infer the magnitude of the correlations, it is difficult to compare networks with different correlations.
- $e_{jk}$ contains approximately $k^2_{max}$ independent variables, representing a huge amount of information that is difficult to model in analytical calculations and simulations.

We therefore need to develop a more compact way to detect the presence and the magnitude of degree correlations.

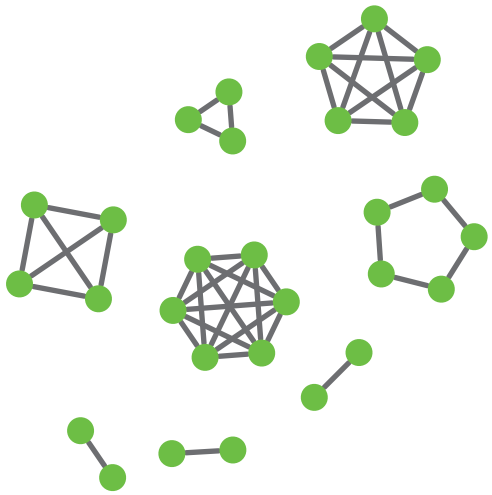


**Figure 7.4**
**A perfectly associative network**

Maximal assortativity is obtained when each degree-*k* node links only to other degree-*k* nodes. For such a perfectly assortative network $e_{jk} = \delta_{jk} q_k$, where $\delta_{jk}$ is the Kronecker delta. In this case the non-diagonal elements of the $e_{jk}$ matrix are zero. The figure shows such a perfectly assortative network, consisting of complete *k*-clusters.

# MEASURING DEGREE CORRELATIONS

While $e_{ij}$ contains the complete information about the potential degree correlations characterizing a network, it is difficult to interpret its content. The purpose of this section is to introduce the degree correlation function, which offers a simpler way to measure degree correlations.

Degree correlations capture the relationship between the degrees of nodes that link to each other. One way to quantify their magnitude is to measure for each node $i$ the average degree of its neighbors Fig. 7.5.

$$k_{nn}(k_i) = \frac{1}{k_i}\sum_{j=1}^{N} A_{ij}k_j \,. \tag{7.6}$$

If we wish to calculate Eq. 7.6 for all nodes with the same degree $k$, we define the degree correlation function as [4, 5]

$$k_{nn}(k) \equiv \sum_{k'} k' P(k' \mid k) \tag{7.7}$$

where $P(k' \mid k)$ is the conditional probability that following a link of a $k$-degree node we reach a degree-$k$ node. To quantify degree correlations we inspect the dependence of $k_{nn}(k)$ on $k$. For *neutral networks*, using Eqs. 7.3-7.5, we have

$$P(k' \mid k) = \frac{e_{kk'}}{\sum_{k'} e_{kk'}} = \frac{e_{kk'}}{q_k} = \frac{q_{k'}q_k}{q_k} = q_{k'} \,. \tag{7.8}$$

Hence $k_{nn}(k)$ can be expressed as

$$k_{nn}(k) = \sum_{k'} k' q_{k'} = \sum_{k'} k' \frac{k' p(k')}{\langle k \rangle} = \frac{\langle k^2 \rangle}{\langle k \rangle} \,. \tag{7.9}$$

Therefore, in a neutral network the average degree of a node's neighbors is independent of the node's degree $k$ and depends only on $\langle k \rangle$ and $\langle k^2 \rangle$. So plotting $k_{nn}(k)$ in function of $k$ is expected to result in a horizontal line at $\langle k^2 \rangle / \langle k \rangle$, as observed in the case of the power grid in Fig. 7.6b. Eq. 7.9 also reflects an intriguing property of real networks: that our friends are more popular than we are, a phenomenon called the *friendship paradox* BOX 7.1.

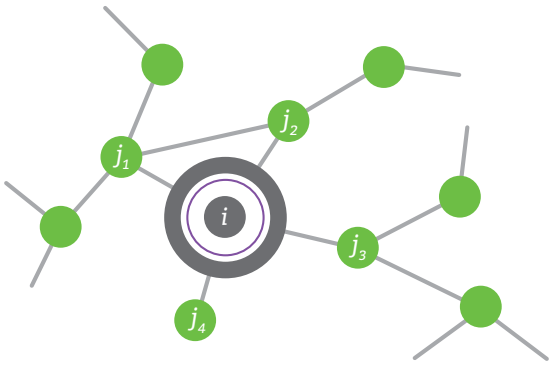



**Figure 7.5**
**Nearest neighbor degree: $K_{nn}$**

To determine $k_{nn}(k)$, we calculate the average degree of a node's neighbors. The figure illustrates the calculation of $k_{nn}(k)$ for node $i$ shown in red. As the degree of the node $i$ is $k_i = 4$, by averaging the degree of its neighbors $j_1$, $j_2$, $j_3$ and $j_4$, we obtain $k_{nn}(4) = (4 + 3 + 3 + 1)/4 = 2.75$.

- **Assortative Network**
  In this case hubs tend to connect to other hubs, hence the higher is the degree $k$ of a node, the higher should be the average degree of its nearest neighbors. Consequently for assortative networks $k_{nn}(k)$ increases with $k$, as observed in collaboration networks in Fig. 7.6a.

- **Disassortative Network**
  In this case hubs prefer to link to low-degree nodes. Consequently $k_{nn}(k)$ decreases with $k$, as observed for the protein-protein interaction network Fig. 7.6c.

The scaling observed in Fig. 7.6 prompts us to approximate the degree correlation function with [4]

$$k_{nn}(k) = ak^{\mu}. \tag{7.10}$$

If Eq. 7.10 holds, then the nature of degree correlations characterizing a network is determined by the sign of the *correlation exponent* $\mu$:

- **For assortative Networks $\mu > 0$**
  Indeed, a fit to $k_{nn}(k)$ for the science collaboration network provides $\mu = 0.37 \pm 0.11$ Fig. 7.6a.

- **For neutral networks we have $\mu = 0$**
  As according to Eq. 7.9 $k_{nn}(k)$ is independent of $k$. For the power grid we obtain $\mu = 0.04 \pm 0.05$, which is indistinguishable from zero Fig. 7.6b.

- **For disassortative networks we expect $\mu < 0$**
  Indeed, for the metabolic network we obtain $\mu = -0.76 \pm 0.04$ Fig. 7.6c.

In summary, the degree correlation function helps us capture the presence or absence of correlations in real networks. The $k_{nn}(k)$ function also plays an important role in analytical calculations, allowing us to calculate the impact of degree correlations on various network characteristics SECT. 7.6. Note that it is often convenient to extract a single number to capture the magnitude of correlations present in a network. This can be achieved either through the correlation exponent $\mu$ defined in Eq. 7.10, or using the degree correlation coefficient discussed in BOX 7.2.

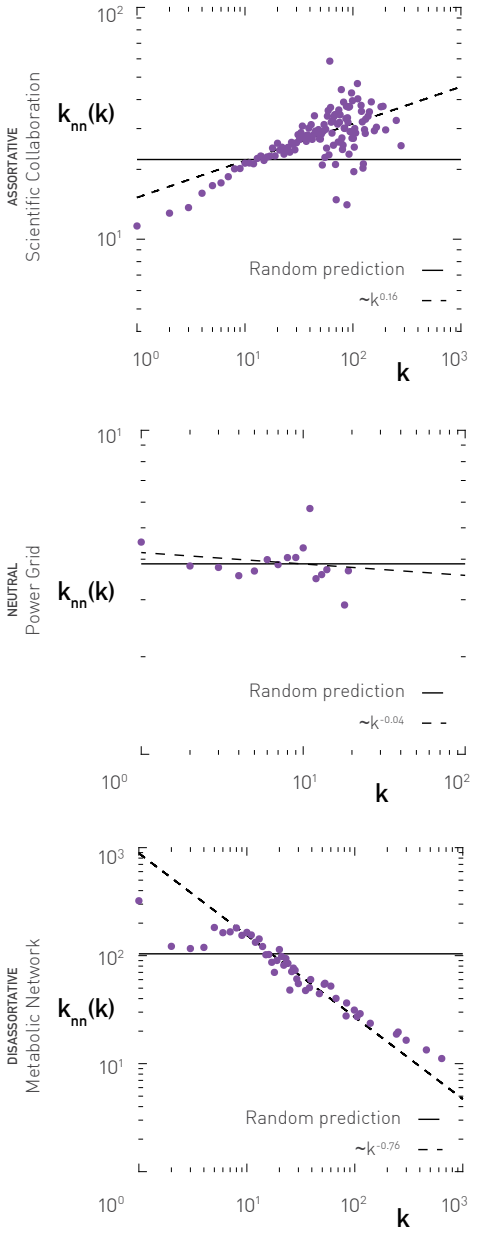



**Figure 7.6**
**Degree correlation function**

The degree correlation function $k_{nn}(k)$ for three real networks. The panels show $k_{nn}(k)$ on a log-log plot to test the validity of **Eq. 7.10**.

**(a)** Collaboration network of astrophysicists. The increasing $k_{nn}(k)$ with $k$ indicates that the network is assortative.

**(b)** Power grid. The horizontal $k_{nn}(k)$ indicates the lack of degree correlations, as predicted by **Eq. 7.9** for neutral networks.

**(c)** Metabolic network. The decreasing $k_{nn}(k)$ documents the network's disassortative nature.

On each panel the horizontal dotted line corresponds to the prediction **Eq. 7.9** and the oblique dashed line is a fit to **Eq. 7.10**. The slope in (a) is $\mu = 0.37$, in (b) is $\mu = 0.04$ while the slope in (c) is $\mu = -0.76$.

# BOX 7.1

**FRIENDSHIP PARADOX**

While most people believe that they have more friends than their friends [7], the friendship paradox, discovered by sociologist Scott L. Feld, states the opposite: on average your friends are more popular than you are [6].

The roots of the friendship paradox is Eq. 7.9, telling us that the average degree of a node's neighbors is not simply $\langle k \rangle$, but depends on $\langle k^2 \rangle$ as well. Consider for example a random (Erdős-Rényi) network, for which $\langle k^2 \rangle = \langle k \rangle (1 + \langle k \rangle)$ . According to Eq. 7.9

$$k_{nn}(k) = 1 + \langle k \rangle. \tag{7.11}$$

Therefore the average degree of a node's neighbors is always higher than the average degree of the network $\langle k \rangle$. The gap between $\langle k \rangle$ and our friends' degree can be particularly large in scale-free networks, for which $\langle k^2 \rangle / \langle k \rangle$ is significantly larger than $\langle k \rangle$ Fig. 4.7. Consider for example the email network, for which $\langle k^2 \rangle / \langle k \rangle = 390.45$, or the actor network, for which $\langle k^2 \rangle / \langle k \rangle = 565.70$. Hence in these networks the average degree of the friends of a randomly selected node can be hundreds of times higher than the expected degree of the node itself, which is $\langle k \rangle$.

To understand the origin of the friendship paradox, we must realize that for a randomly chosen node, the degree distribution of the nodes at the other end of each link do not follow $p_k$, but are biased towards higher-degree nodes, as indicated by Eq. 7.3. In other words, we are more likely to be friends with hubs than with small-degree nodes, simply because hubs have more friends than the small-nodes. Hence our friends do not reflect the whole population - they are biased towards the hubs.

# STRUCTURAL CUTOFFS

Throughout this book we assumed that the networks we explore are simple, meaning that there is at most one link between any two nodes CHAPTER 2. For example, in the email network we place a single link between two individuals that are in email contact, despite the fact that they may have exchanged multiple messages; in the actor network we connect two actors with a single link if they acted together, independent of the number of movies they jointly made. All datasets discussed in TABLE 4.1 are simple networks. In simple networks there is a puzzling conflict between the scale-free property and degree correlations [10, 11]. Consider for example the scale-free network of Fig. 7.7a, whose two largest hubs have degrees $k = 55$ and $k' = 46$, connected by a link. In a network with degree correlations $e_{kk'}$ the expected number of links between $k$ and $k'$ is

$$E_{kk'} = e_{kk'}\langle k\rangle N \tag{7.14}$$

For a neutral network $e_{kk'}$ is given by Eq. 7.5, which, using Eq. 7.3, predicts

$$E_{kk'} = \frac{k_k p_k k' p_{k'}}{\langle k\rangle} N = \frac{\frac{55}{300}\frac{46}{300}}{3} 300 = 2.8\,. \tag{7.15}$$

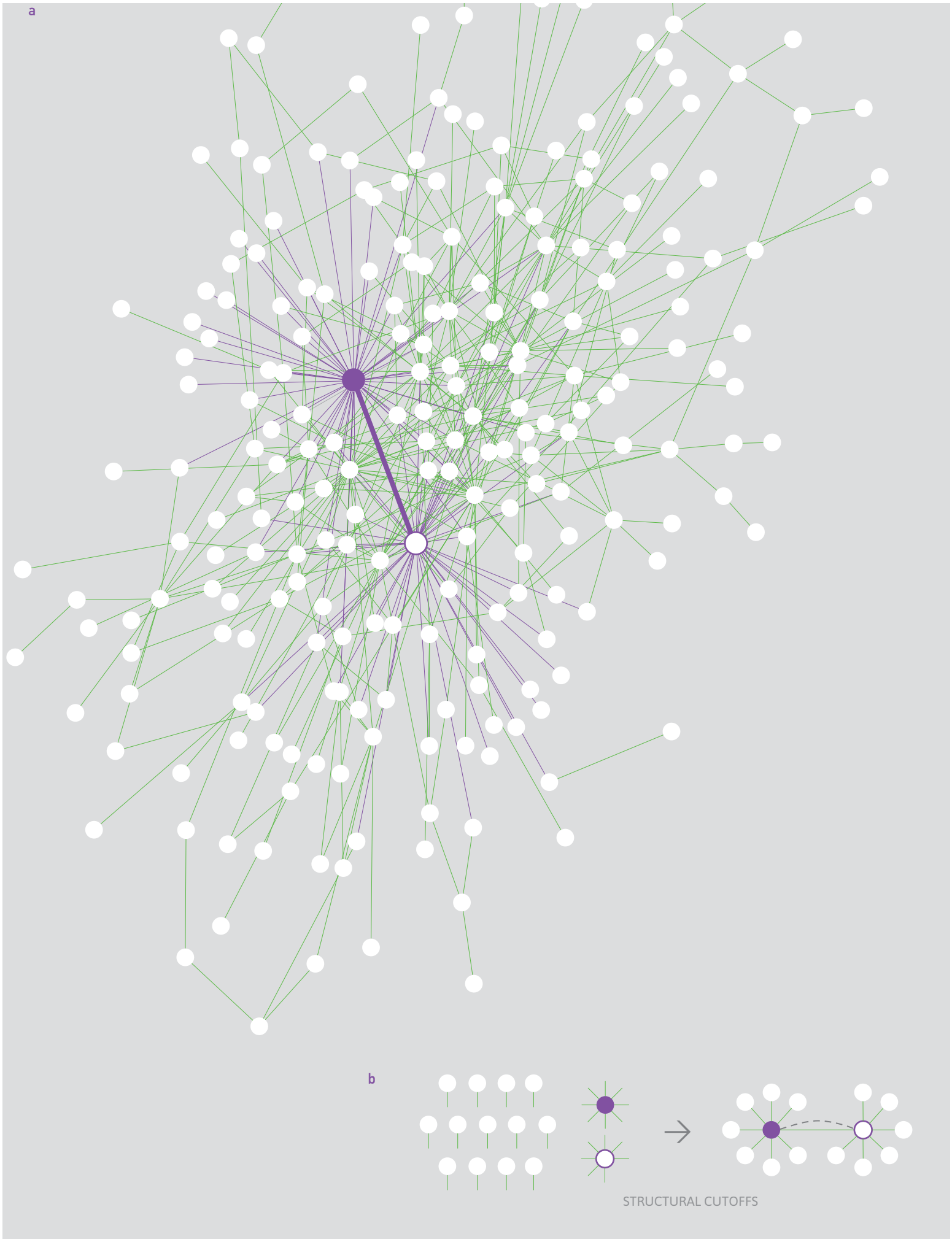
Therefore, given the size of these two hubs, they should be connected to each other by two to three links to comply with the network's neutral nature. Yet, in a simple network we are allowed only one link between them, raising a conflict between degree correlations and the scale-free property. Such conflict emerges in a simple network each time the degrees violate the $E_{kk'} \leq 1$ condition. The goal of this section is to understand the origin and the consequences of this conflict.

For small $k$ and $k$ Eq. 7.15 predicts that $E_{kk'}$ is also small, i.e. we expect less than one link between the two nodes. Only for nodes whose degree exceeds some threshold $k_s$ will Eq. 7.15 predict multiple links. As we show in ADVANCED TOPICS 7.B, this $k_s$, that we

**Figure 7.7 (following page)**
**Structural disassortativity**

**(a)** A scale-free network with $N$=300, $L$=450, according **Eq. 7.15**, and $\gamma$=2.2, generated by the configuration model, while forbidding self-loops and multiple links between two nodes, making the network simple. The blue and the red nodes are the two largest nodes in the network and are connected by the red link. As **Eq. 7.15** predicts,to maintain the network's neutral nature, we would need two to three links between these two nodes. The fact that we do not allow multiple links (simple network representation) makes the network disassortative, a phenomena we call structural disassortativity.

**(b)** To illustrate the origins of structural correlations, we start from a fixed degree sequence, shown as stubs on the left, and we randomly connect the stubs (configuration model). In this case, the expected number of links between the nodes with degree 8 and 7 is 8 x 7/28 ≈ 2. Yet, if we do not allow multilinks, there can only be one link, making the network structurally dissasortative.


a
b
STRUCTURAL CUTOFFS

$$k_s(N) \sim (\langle k \rangle N)^{1/2}. \quad (7.16)$$

In other words, nodes whose degree exceeds Eq. 7.16 are expected to have $E_{kk'} > 1$, a conflict that as we show below gives rise to degree correlations.

To fully understand the consequences of the described conflict, we must first ask if a network has nodes whose degrees exceeds Eq. 7.16. For this we compare the structural cutoff, $k_s$, with the natural cutoff, $k_{max}$, which is the expected largest degree in a network with degree distribution $p_k$. According to Eq. 7.14, for a scale-free network $k_{max} \sim N^{\frac{1}{\gamma-1}}$. The relative magnitude of $k_{max}$, vs. $k_s$, gives raise to two regimes:

- For scale-free networks with $\gamma \geq 3$ and random networks, ks is always larger than $k_{max}$, hence we lack nodes for which $E_{kk'} > 1$.

- For scale-fee networks with $\gamma < 3$, $k_s$ is smaller than $k_{max}$, hence all nodes between $k_s$ and $k_{max}$ violate $E_{kk'} > 1$. Consequently, the network has fewer links between its hubs than expected based on Eq. 7.15. As a result, these networks will be disassortative, a phenomenon we call *structural disassortativity*. This is illustrated in Figs. 7.8a, b that show a simple scale-free network generated by the configuration model. The network shows disassortative tendencies, despite the fact that we did not impose degree correlations.

We have two avenues to generate networks that are free of structural disassortativity:

(i) We relax the simple network requirement, allowing multiple links between the nodes. The conflict disappears and the network will be neutral Figs. 7.8c, d.

(ii) If we insist of having a simple scale-free network that is neutral or associative, we must remove all hubs with degrees larger than $k_s$. This is illustrated in Fig 7.8 e, f: the obtained network, missing nodes with $k \geq 100$, is neutral.

How can we convince ourselves that the correlations observed in a particular network are a consequence of structural dissasortativity, or are generated by some unknown process? Degree-preserving randomization Fig. 4.14 helps us distinguish these two possibilities:

(i) Degree preserving randomization with simple links (*R-S*): We apply degree-preserving randomization to the original network, while making sure that we do not allow for more than one link between any pair of nodes. On the algorithmic side this means that each rewiring that results in multiple links between two nodes is discarded. If the real $k_{nn}(k)$ and the randomized $k_{nn}^{R-S}(k)$ are indistinguishable, then the correlations observed in a real system are all structural,

## BOX 7.2

### DEGREE CORRELATION COEFFICIENT

If we wish to characterize degree correlations using a single number, we can also use the degree correlation coefficient, introduced by Mark Newman and defined as [8,9]

$$r = \sum_{jk} \frac{jk(e_{jk} - q_j q_k)}{\sigma_r^2} \quad (7.12)$$

with

$$\sigma_r^2 = \sum_k k^2 q_k - \left[\sum_k^{jk} k q_c\right]^2 \quad (7.13)$$

Hence $r$ is the Pearson correlation coefficient between the degrees at the two end of the same link. It varies between $-1 \leq r \leq 1$: for $v < 0$ the network is assortative, for $r = 0$ the network is neutral and for $v > 0$ the network is disassortative. For example, for the collaboration network we obtain $r = 0.13$, in line with its assortative nature; for the protein interaction network $r = -0.04$, supporting its disassortative nature and for the power grid $r = 0$. Note that the degree correlation coefficient $r$ assumes that $k_{nn}(k)$ is a linear function of $k$ with slope $r$. In contrast the correlation exponent $\mu$ assumes that $k_{nn}(k)$ follows the power law Eq. 7.10. Naturally, both scaling laws cannot be valid simultaneously. The analytical models of SECT. 7.7 offer some guidance, supporting the validity of Eq. 7.10. As we show in ADVANCED TOPICS 7.B, while $r$ correlates with $\mu$, we need to be cautious when we use it to measure degree correlations.

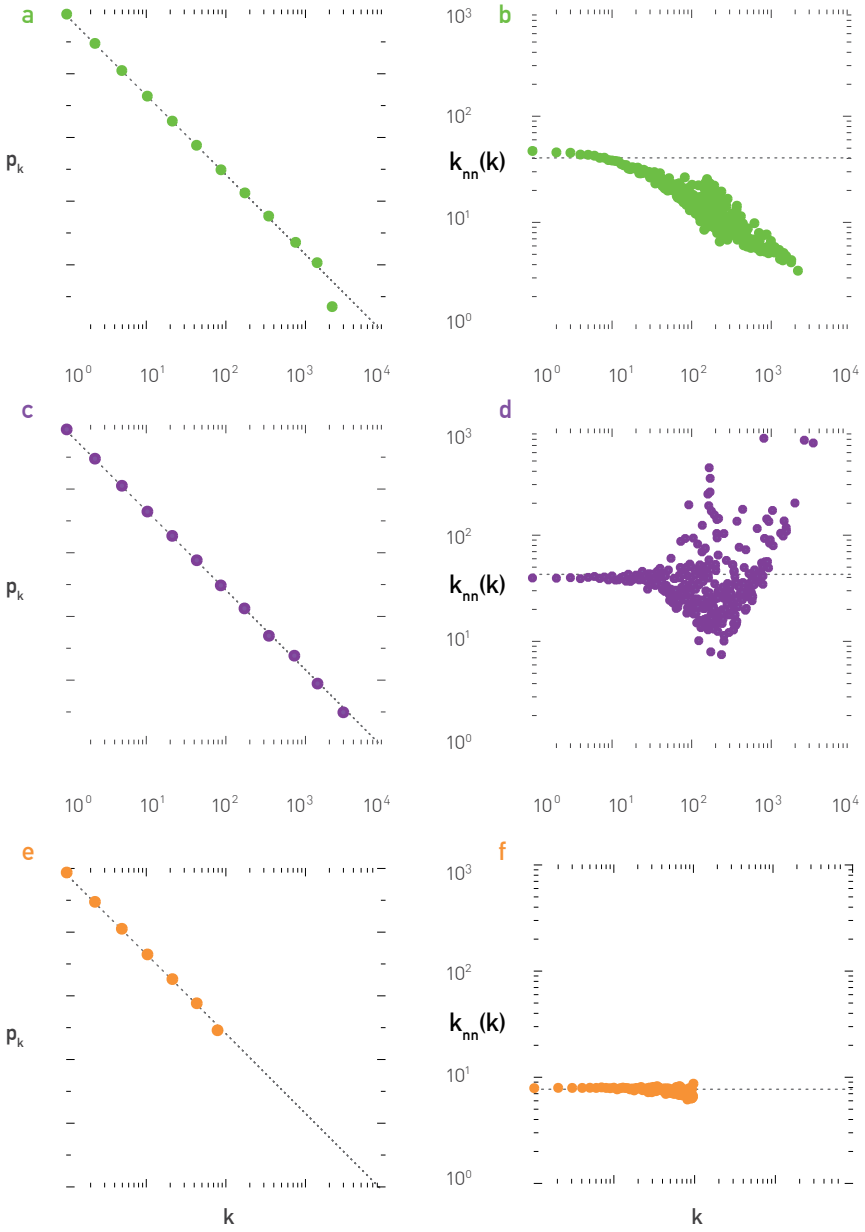



**Figure 7.8**
**Natural and structural cutoffs**

The figure illustrates the tension between the scale-free property and degree correlations. It shows the degree distribution (left panels) and the degree correlation function $k_{nn}(k)$ (right panels) of a scale-free network with $N = 10{,}000$ and $\gamma = 2.5$, generated by the configuration model.

**(a, b)** If we generate a scale-free network with the power-law degree distribution shown in (a), and we forbid self-loops and multi-links, the network displays structural disassortativity, as indicated by $k_{nn}(k)$ in (b). In this case, we lack a sufficient number of links between the high-degree nodes to maintain the neutral nature of the network, hence for high $k$ the $k_{nn}(k)$ function decays.

**(c, d)** We can eliminate structural disassortativity by allowing multiple links, i.e. relaxing the simple network requirement. As shown in (c,d), in this case we obtain a neutral scale-free network.

**(e, f)** If we artificially impose an upper cutoff by removing all nodes with $k \geq k_s$ predicted by **Eq. 7.16**, the network becomes neutral, as seen in (f).

fully explained by the degree distribution. If the randomized $k_{nn}^{R\text{-}S}(k)$ does not show degree correlations while $k_{nn}(k)$ does, there is some unknown process that generates the observed degree correlations.

**(ii)** Degree preserving randomization with multiple links (*R-M*): For a self-consistency check it is useful to also perform degree-preserving randomization that allows for multiple links between the nodes. On the algorithmic side this means that we allow each random rewiring, even if they lead to multiple links. This process eliminates all degree correlations.

We have taken the three networks of in Fig. 7.6 and performed the randomizations discussed above. As Fig. 7.9a shows, the assortative nature of the scientific collaboration network disappears under both randomizations. This indicates that the observed assortative correlations are not linked to the scale-free nature of the underlying network. In contrast, for the metabolic network the observed disassortativity remains unchanged under *R-S* Fig. 7.9c. This indicates that the disassortativity of the metabolic network is structural, induced by its degree distribution.

In summary, the scale-free property can induce disassortativity in simple networks. To be specific, in neutral or assortative networks we expect multiple links between the hubs. If such multiple links are forbidden (simple graph), the network will display disassortative tendencies. This conflict vanishes for scale-free networks with $\gamma \geq 3$ and for random networks. It also vanishes if we allow for multiple links between the nodes.

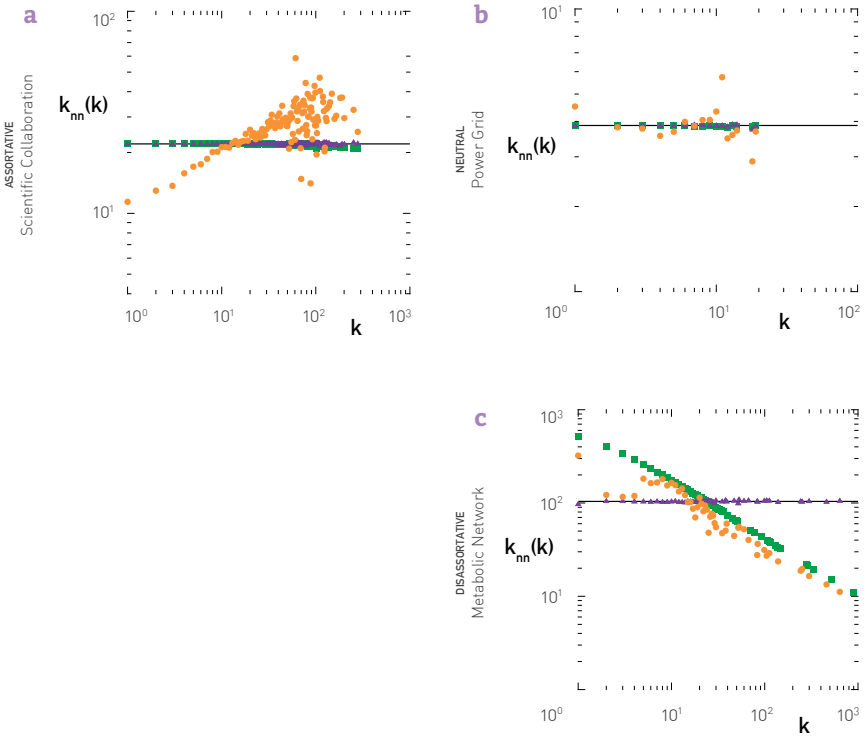



**Figure 7.9**
**Randomization and degree correlations**

To uncover the origin of the observed degree correlations, it is useful to compare $k_{nn}(k)$ with $k_{nn}^{R\text{-}S}(k)$ and $k_{nn}^{R\text{-}M}(k)$ obtained after degree-preserving randomization. We perform two different randomizations for this purpose.

**Green symbols**
Degree-preserving randomization with simple links (*r-s*), in which case at each step of the randomization process we check that we do not have more than one link between any node pairs.

**Blue symbols**
Degree-preserving randomization with multiple links (*R-M*), in which case we allow multilinks during the randomization processes.

We performed these two randomizations for the networks of Fig. 7.6. The *R-M* procedure always generates a neutral network, consequently $k_{nn}^{R\text{-}M}(k)$ is always horizontal.The true insight is provided when we compare $k_{nn}(k)$ with $k_{nn}^{R\text{-}S}(k)$, allowing us to decide if the observed correlations are structural:

**(a) Scientific collaboration network**
The increasing $k_{nn}(k)$ differs from the horizontal $k_{nn}^{R\text{-}S}(k)$, indicating that the network's assortativity is not structural (i.e. it is not a consequence of the degree distribution), but it is generated by some process that governs the network's evolution. This is not unexpected: structural effects can generate only disassortativity, not assortativity.

**(b) Power grid**
The horizontal $k_{nn}(k)$, $k_{nn}^{R\text{-}S}(k)$ and $k_{nn}^{R\text{-}M}(k)$ all support the lack of degree correlations (neutral network).

**(c) Metabolic network**
As both $k_{nn}(k)$ and $k_{nn}^{R\text{-}S}(k)$ decrease, we conclude that the network's disassortative nature is induced by its scale-free property. Hence the observed degree correlations are structural.

# DEGREE CORRELATIONS IN REAL NETWORKS

To truly understand the prevalence of degree correlations, we need to inspect the correlations characterizing various real networks. Therefore, in Fig. 7.10 we show the $k_{nn}(k)$ function for the ten reference networks of TABLE 4.1. Let us discuss the observed behavior:

- **Power grid**
  For the power grid $k_{nn}(k)$ is flat and indistinguishable from its randomized version, indicating a lack of degree correlations Fig. 7.10a. Hence the power grid is neutral.

- **Internet**
  For small degrees ($k \leq 30$) $k_{nn}(k)$ shows a clear assortative trend, an effect that levels off for high degrees Fig. 7.10b. The degree correlations vanish in the randomized networks. Hence the Internet is assortative, but structural cutoffs eliminate the effect for high $k$.

- **Social Networks**
  The three networks capturing social phenomena, like the mobile phone network, science collaboration networks and actor network, all have an increasing $k_{nn}(k)$, indicating that they are assortative Figs. 7.10c-e. Hence in these networks hubs tend to link to other hubs and low-degree nodes tend to link to low-degree nodes. For each of these networks the observed $k_{nn}(k)$, differs from the $k_{nn}^{R\text{-}S}(k)$, indicating that their assortative nature is not rooted in the degree distribution.

- **Email Network**
  While the email network is often used as an example of a social network, its $k_{nn}(k)$ decreases with $k$, documenting a clear disassortative behavior Fig. 7.10f. The randomized $k_{nn}^{R\text{-}S}(k)$ also decays, indicating that we are observing structural disassortativity, a consequence of the network's scale-free nature.

- **Biological Networks**
  The protein interaction and the metabolic network both have a nega-

tive μ, suggesting that these networks are disassortative Eq. 7.10. Yet, the scaling of $k_{min}^{R\text{-}S}(k)$ is indistinguishable from $k_{nn}(k)$, indicating that we are observing structural disassortativity, rooted in the scale-free nature of these networks Fig. 7.10g, h.

- **WWW**
  The decaying $k_{nn}(k)$ implies disassortative correlations Fig. 7.10i. The randomized $k_{min}^{R\text{-}S}(k)$ also decays, but not as rapidly as $k_{nn}(k)$. Hence the disassortative nature of the WWW is not fully explained by its degree distribution.

- **Citation network**
  This network displays a puzzling behavior: for $k \leq 20$, $k_{nn}(k)$ shows a clear assortative trend; for $k > 20$, however, we observe equally clear disassortative scaling Fig. 7.10j. Such mixed behavior can emerge in networks that display extreme assortativity SECT. 7.6. This suggests that the citation network is strongly assortative up to $k_s$, but its scale-free nature reverses the trend for $k \gg k_s$.

In summary, Fig. 7.10 indicates that to understand degree correlations, we must always compare $k_{nn}(k)$ to the degree randomized $k_{nn}^{R\text{-}S}(k)$. It also allows us to draw some interesting conclusions:

**(i)** Of the ten reference networks the power grid appears to be the only that is truly neutral. Hence most real networks display degree correlations.

**(ii)** All networks that display disassortative tendencies (email, protein, metabolic), do so thanks to their scale-free property. Hence, these are all structurally disassortative. Only the WWW shows disassortative correlations that are only partially explained by its degree distribution.

**(iii)** The degree correlations characterizing associative networks are not explained by their degree distribution. Most social networks (mobile phone calls, scientific collaboration, actor network) are in this class and so is the Internet and the citation network.

A number of proposals exist to explain the origin of the observed assortativity. For example, the tendency of individuals to form communities CHAPTER 9 has been shown to induce assortative scaling [12]. Similarly, the society has endless mechanisms, from professional committees to TV shows, to bring hubs together, enhancing the assortative nature of social and professional networks. Finally, homophily, a well documented social phenomena, [13], captures the fact that individuals have a tendency to associate with other individuals of similar background and characteristics. This tendency may also be responsible for the celebrity marriages discussed in SECT. 7.0.

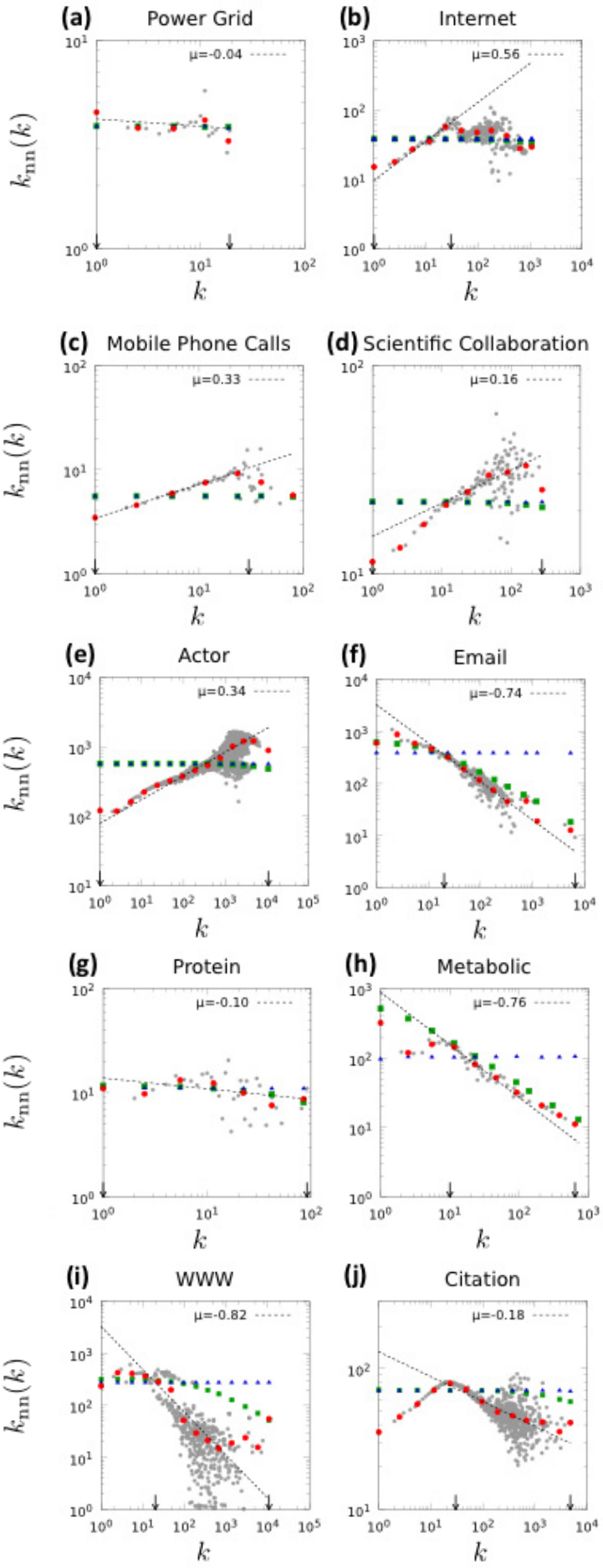



**Figure 7.10**
**Randomization and degree correlations**

The degree correlation function $k_{nn}(k)$ for the ten reference networks of **Table 4.1**. The grey symbols show the $k_{nn}(k)$ function under linear binning: red symbols represent the same data using log-binning **SECT. 4.10**. The dotted line corresponds to the best fit of the form **Eq. 7.10** and the small arrows at the bottom mark the fitting interval. Green squares represent $k_{nn}(k)$ obtained for 100 independent degree-preserving randomizations, making sure that we preserve the simple character of these networks; blue triangles correspond to $k_{nn}(k)$, i.e. randomization that does allow self-loops and multiple links between two nodes. Note that we made directed networks undirected when we measured $k_{nn}(k)$. To fully characterize the correlations emerging in directed networks we must use the directed correlation function **BOX 7.3**.

# BOX 7.3

## CORRELATIONS IN DIRECTED NETWORKS

The degree correlation function $k_{nn}(k)$ in Eq. 7.7 is defined for undirected networks. To measure correlations in directed networks we must take into account that each node $i$ is characterized by an incoming $k_i^{in}$ and an outgoing $k_i^{out}$ degree. Hence, we define four degree correlation functions, $k_{nn}^{\alpha,\beta}(k)$, where $\alpha$ and $\beta$ refer to the in and out indices Figs. 7.11 a-d. In Fig. 7.11e we show $k_{nn}^{\alpha,\beta}(k)$ for citation networks, indicating a lack of in-out correlations, while a detectable assortative scaling for small $k$ for the other three correlations.

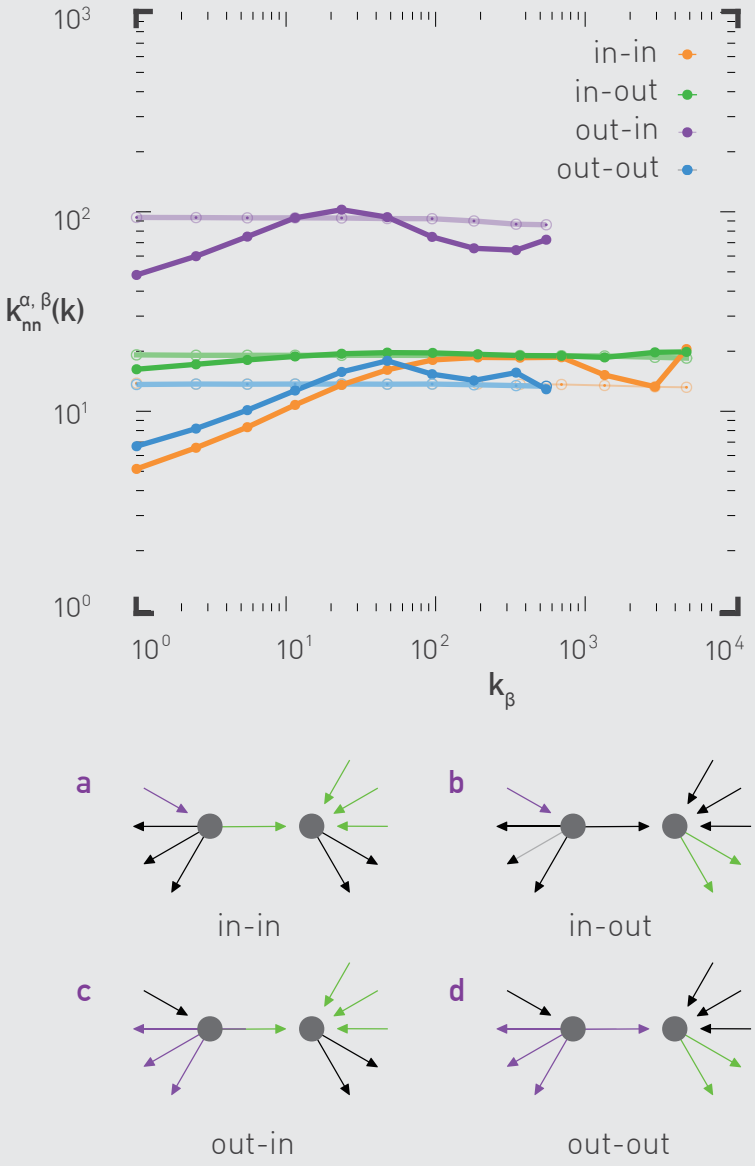



**Figure 7.11**
**Correlation and directed network**

Panels (a)-(d) illustrate the four possible correlations in directed networks. We show in red and green the ($\alpha$, $\beta$) indices that define the appropriate correlation function [14]. For example, (a) describes the $k_{nn}^{in,in}(k)$ correlations between the in-degrees of two nodes connected by a link. (e) The $k_{nn}^{\alpha,\beta}(k)$ correlation function for citation networks, a directed network. For example $k_{nn}^{in,in}(k)$ is the average indegree of the in-neighbors of nodes with in-degree $k_{in}$. These functions show a clear assortative tendency for three of the four function up to degree $k \approx 100$. The lighter symbols capture the degree randomized $k_{nn}^{\alpha,\beta}(k)$ for each correlation function.

# GENERATING CORRELATED NETWORKS

To study degree correlations and to explore their impact on various network characteristics, we need to build networks with tunable correlations. Given the conflicts between the scale-free property and degree correlations, this is not a trivial task. In this section we discuss the degree correlations characterizing some well-known network models, together with an algorithm capable of generating networks with tunable correlations.

## DEGREE CORRELATIONS IN STATIC MODELS

### Erdős-Rényi Model

The random network model is neutral by definition. As it lacks hubs, it does not develop structural correlations either. Hence for the Erdős-Rényi network $k_{nn}(k)$ is given by Eq. 7.9, predicting $\mu = 0$ for any $\langle k \rangle$ and $N$.

*Configuration Model*: The configuration model SECT. 4.7 is also neutral, independent of our choice of the degree distribution $p_k$. This is because the model allows for both multi links and self-loops. Consequently, any conflicts caused by the hubs are relieved by multiple links between them. If, however, we force the network to be simple, then the generated network will develop structural disassortativity Fig. 7.8.

### Hidden Parameter Model

In the model $e_{jk}$ is the product of the hidden variables $\eta_j$ and $\eta_k$, which are chosen randomly, hence the network is technically uncorrelated SECT. 4.8. However, if we do not allow multiple links, for scale-free networks we again observe structural disassortativity. Analytical calculations indicate that in this case $k_{nn}(k) \sim k^{-1}$, i.e. we have $\mu = -1$ [10].

## DEGREE CORRELATIONS IN EVOLVING NETWORKS

To understand the emergence (and absence) of degree correlations in growing networks, let us start with the initial attractiveness model discussed in SECT. 6.4. In the model preferential attachment follows $\Pi(k) \sim A + k$, where A is the initial attractiveness Eq. 6.23. The degree correlation function depends on $A$, the calculations predicting three scaling regimes [15]:

**(i)** If $\gamma < 3$ (i.e. $-m < A < 0$ according to Eq. 6.24, we have

$$k_{nn}(k) \simeq m\frac{(m+A)^{1-\frac{A}{m}}}{2m+A}\varsigma\left(\frac{2m}{2m+A}\right)N^{-\frac{A}{2m+A}}k^{\frac{A}{m}} \tag{7.17}$$

Hence the resulting network is disassortative, $k_{nn}(k)$ being characterized by the power-law decay [15, 16]

$$k_{nn}(k) \simeq k^{-\frac{|A|}{m}} \tag{7.18}$$

**(ii)** If $\gamma = 3$ ($A = 0$), the initial attractiveness model reduces to the Barabási-Albert model CHAPTER 5. In this case

$$k_{nn}(k) \simeq \frac{m}{2}\ln N, \tag{7.19}$$

that is, $k_{nn}(k)$ is independent of $k$, hence the network is neutral.

**(iii)** If $\gamma > 3$ ($A > 0$), the calculations predict

$$K_{nn}(k) \simeq (m+a)\ln\left(\frac{k}{m+a}\right). \tag{7.20}$$

As $k_{nn}(k)$ increases logarithmically with $k$, the resulting network displays a weak assortative tendency, but does not follow the scaling Eq. 7.10.

### Bianconi-Barabási Model

With a uniform fitness distribution the Bianconi-Barabási model generates a disassortative network [5] Fig. 7.12. As the randomized version of the network is also disassortative, this is a structural disassortativity. Note, however, that the real $k_{nn}(k)$ and the randomized $k_{nn}^{R\text{-}S}(k)$ do not overlap, indicating that the Bianconi-Barabási model displays some disassortativity that is not fully explained by its scale-free nature.

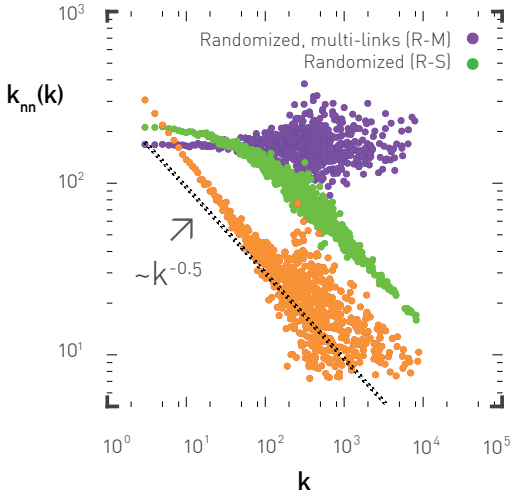



**Figure 7.12**
**Correlations in the Bianconi-Barabási model**

The degree correlation function of the Bianconi-Barabási model for $N = 10,000$, $m = 3$ and uniform fitness distribution SECT. 6.2. As the dotted line indicates, the network is disassortative, with $\mu = 0.5$. The green symbols show $k_{nn}^{R\text{-}S}(k)$, and the blue are for $k_{nn}^{R\text{-}M}(k)$. As $k_{nn}^{R\text{-}S}(k)$ also decreases, the bulk of the observed disassortativity is structural. But the difference between $k_{nn}^{R\text{-}S}(k)$ and correlations in the Bianconi-Barabási model suggests that structural effects cannot fully account for the observed degree correlation.

## TUNING DEGREE CORRELATIONS

Several algorithms exist to generate networks with desired degree correlations [8, 17, 18]. Here we discuss a simplified version of the algorithm proposed by Xalvi-Brunet and Sokolov that generates maximally correlated networks with a predefined degree sequence [19, 20, 21]. It consists of the following steps Fig. 7.13a:

- **Step 1: Link selection**
  Choose at random two links. Label the four nodes at the end of these two links with $a$, $b$, $c$, and $d$ such that their degrees $k_a$, $k_b$, $k_c$, and $k_d$ are ordered as

$$k_a \geq k_b \geq k_c \geq k_d.$$

- **Step 2: Rewiring**
  Break the selected links and rewire them to form new pairs. Depending on the desired degree correlations the rewiring is done in two different ways:

- **Step 2A: Assortative**
  By pairing the two highest degrees ($a$ with $b$) and the two lowest degrees ($c$ with $d$), we are connecting nodes with comparable degrees, enhancing the network's assortative nature.

- **Step 2B, Disassortative**
  By pairing the highest and the lowest degree nodes ($a$ with $d$ and $b$ with $c$), we tend to connect nodes with rather different degrees, enhancing the network's disassortative nature.

By iterating these steps we gradually enhance the network's assortative (2A) or disassortative (2B) features. If we aim to generate a simple network (free of multi-links), after *Step 2* we check whether the particular rewiring leads to multi-links. If it does, we reject it, returning to *Step 1*.

The correlations characterizing the networks generated by this algorithm converge to the maximal or minimal value one can reach for the given degree sequence Fig. 7.13b. We refer to these networks as maximally assortative or maximally disassortative. The model has no difficulty creating disassortative correlations Figs. 7.13e, f. In the assortative limit simple networks displays a mixed $k_{nn}(k)$: assortative for small $k$ and disassortative for high $k$ Figs. 7.13b. This is a consequence of structural cutoff: for scale-free networks the system is unable to sustain assortativity for high $k$. This behavior is reminiscent of the $k_{nn}(k)$ function observed for citation networks Fig. 7.10j.

The version of the Xalvi-Brunet & Sokolov algorithm discussed in Fig. 7.13 generates maximally assortative or disassortative networks. We can tune the magnitude of the generated degree correlations if we use the original version of the proposed algorithm, discussed in Fig. 7.14.

In summary, static models, like the configuration or hidden parameter models, are neutral if we allow multi-links, and develop structural disassortativity if we force them to generate simple networks. To generate networks with tunable correlations, we can use for example the Xalve-Brunet & Sokolov algorithm. An important result of this section is Eq. 7.17, predicts the functional form of the degree correlation function for a growing network, offering analytical backing for the scaling hypothesis Eq. 7.10.

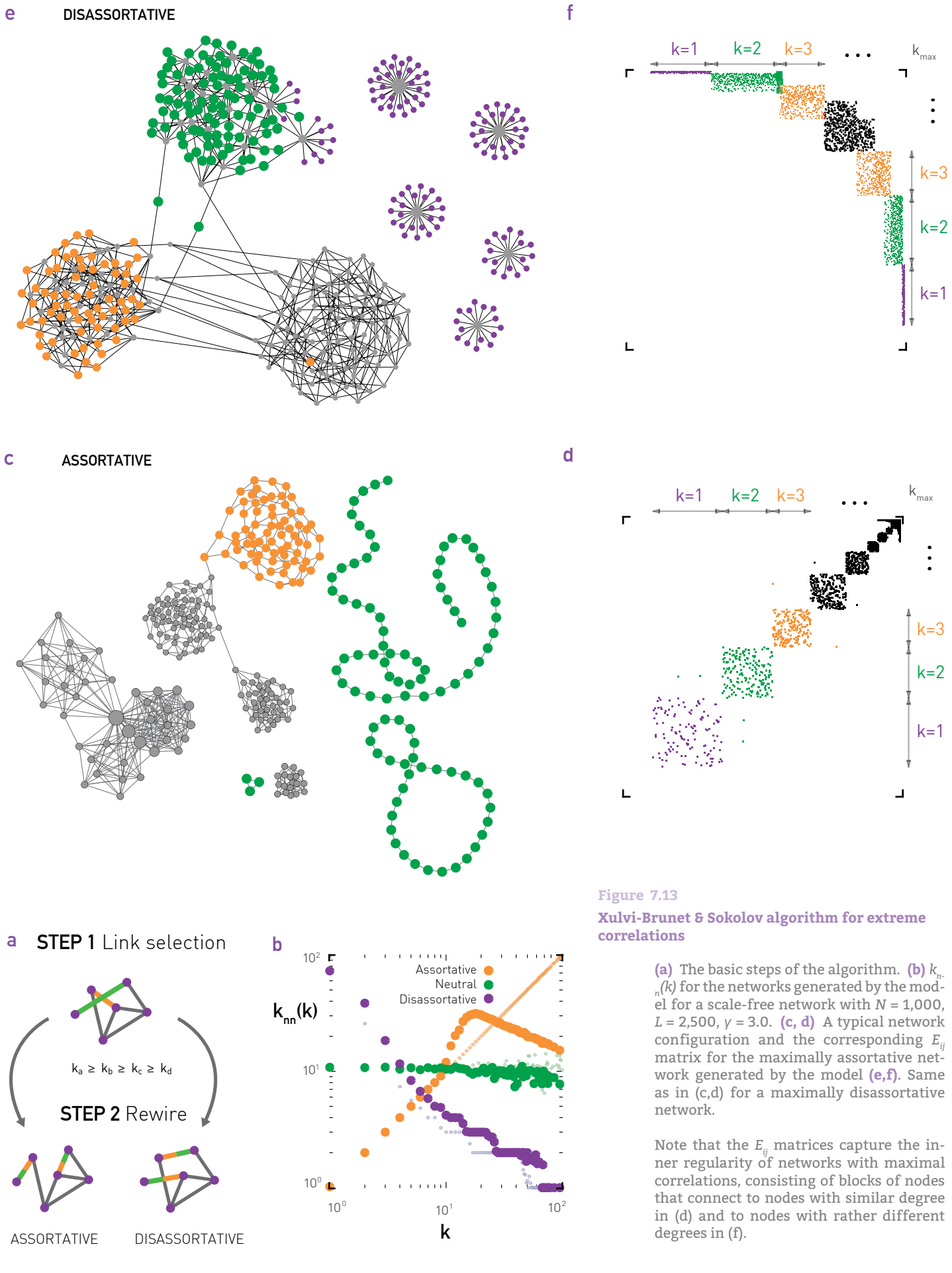



**Figure 7.13**

**Xulvi-Brunet & Sokolov algorithm for extreme correlations**

**(a)** The basic steps of the algorithm. **(b)** $k_{nn}(k)$ for the networks generated by the model for a scale-free network with $N = 1{,}000$, $L = 2{,}500$, $\gamma = 3.0$. **(c, d)** A typical network configuration and the corresponding $E_{ij}$ matrix for the maximally assortative network generated by the model **(e,f)**. Same as in (c,d) for a maximally disassortative network.

Note that the $E_{ij}$ matrices capture the inner regularity of networks with maximal correlations, consisting of blocks of nodes that connect to nodes with similar degree in (d) and to nodes with rather different degrees in (f).

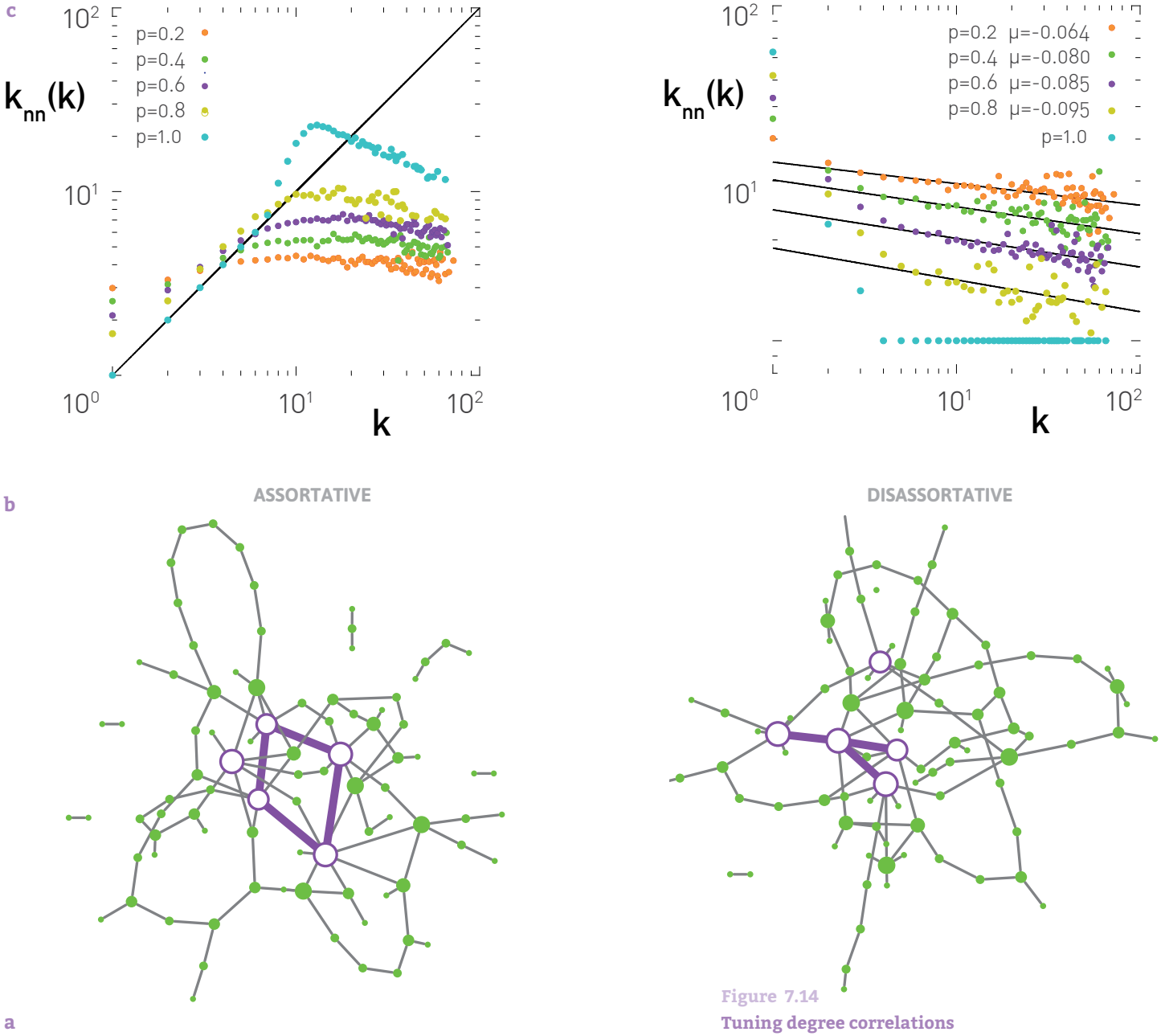



**Figure 7.14**
**Tuning degree correlations**

**(a)** The original Xalvi-Brunet & Sokolov algorithm allows us to tune the magnitude of the observed degree correlations. For this we execute the deterministic rewiring step with probability $p$, and with probability $1 - p$ we randomly pair the $a$, $b$, $c$, $d$ nodes with each other. For $p = 1$ we are back to the model of **Fig. 7.13**, generating maximal degree correlations; for $p < 1$ the induced noise tunes the magnitude of the effect.

**(b)** Typical network configurations generated for $p = 0.5$.

**(c)** The $k_{nn}(k)$ functions for various $p$ values. The simulations are shown for a network with $N = 10{,}000$, $\langle k \rangle = 1$, and $\gamma = 3.0$.

Note that the fit of **Eq. 7.10** is nonconclusive, as the exponents depend on the fitting region, especially in the assortative case.

# THE IMPACT OF DEGREE CORRELATIONS

As we have seen in SECT. 7.5, most real networks are characterized by some degree correlations. Social networks are assortative; biological networks display structural disassortativity. The presence of these correlations raise an important question: why do we care? In other words, do degree correlations alter the properties of a network? And which network properties do they influence? The purpose of this section is to briefly address these questions.

As we have seen in SECT. 3.6, an important property of a random network is the emergence of a phase transition at $\langle k \rangle = 1$, marking the appearance of the giant component. Fig. 7.15 shows the relative size of the giant component for networks with different degree correlations, indicating that [8, 19, 20]:

- **For assortative networks**
  the phase transition point moves to a lower $\langle k \rangle$, hence a giant component emerges for $\langle k \rangle < 1$. The reason is that it is easier to create a giant component if the high-degree nodes tend to link to other high-degree ones.

- **For disassortative networks**
  the phase transition is delayed, as in these networks the hubs tend to connect to small degree nodes. Consequently, these networks have difficulty forming a giant component.

- For large $\langle k \rangle$ the giant component is smaller in assortative networks than in neutral or disassortative networks. Indeed, the high-degree nodes form a core group of high mean degree. As assortativity forces these hubs to mostly link to each other, they fail to attract to the giant component the numerous small degree nodes.

These changes in the size and the structure of the giant component have implications on the spread of diseases [22, 23, 24], a topic discussed in CHAP. 10. Indeed, as we have seen in SECT. 7.4, social networks tend to be

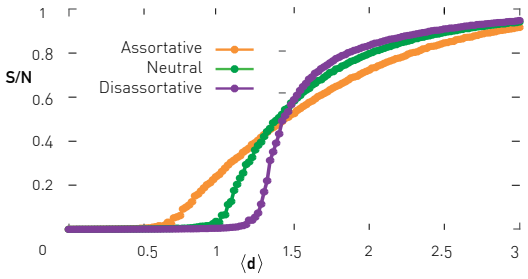



**Figure 7.15**
**Degree correlations and the phase transition point**

Relative size of the giant component for an Erdös-Rényi network of size $N$=10,000 (green curve), which is rewired using the Xalvi-Brunet & Sokolov algorithm with $p = 0.5$, to induce degree correlations (red and blue curve). Each point represents an average of 10 independent runs. The figure indicates that as we move from assortative to disassortative networks, the phase transition point is delayed and the size of the giant component increases for large $\langle k \rangle$.

assortative. The high degree nodes therefore form a giant component that acts as the "reservoir" for the disease, sustaining an epidemic even when on average the network is not sufficintly dense for the virus to persist.

The altered giant component has implications for network robustness as well [25]. As we discuss in CHAPTER 8, a network can be fragmented by the removal of its hubs. In assortative networks hub removal makes less damage because the hubs cluster together, forming a core group, hence many of them are redundant. The removal of the hubs is more damaging in disassortative networks, as in these the hubs connect to many small-degree nodes, which fall off the network once a hub is deleted.

Let us mention a few additional consequences of degree correlations:

- Fig. 7.16 shows the path-length distribution for a random network rewired to display different degree correlations. It indicates that in assortative networks the average path length is shorter than in neutral networks. Yet the most dramatic difference is in the network diameter $d_{max}$, which is significantly higher for assortative networks. Indeed, assortativity favors links between nodes with similar degree, hence it results in long chains of $k = 2$ nodes, enhancing $d_{max}$ Fig. 7.13c.

- Degree correlations influence a system's stability against stimuli and perturbations [26] as well as the synchronization of oscillators placed on a network [27, 28].

- Degree correlations have a fundamental impact on vertex cover problems [29], requiring us to find the minimal set of nodes such that each link is connected to at least one node in the vertex cover BOX 7.4.

- Finally, degree correlations have an impact on our ability to control a network, altering the number of input signals one needs to achieve full control [30].

In summary, degree correlations are not only of academic interest, but they alter numerous network characteristics and have a strong impact on various processes that take place on a network.

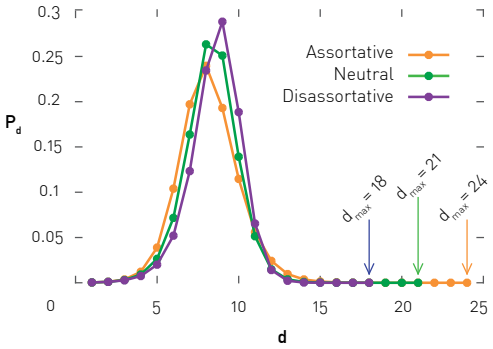



**Figure 7.16**
**Degree correlations and path lengths**

Shortest path distribution for a network with Poisson degree distribution of size $N = 10,000$ and $\langle k \rangle = 3$. Correlations are added using the Xalvi-Brunet & Sokolov algorithm with $p = 0.5$. Each curve presents is an average of 10 independent networks. The plots indicate that as we move from disassortative to assortative networks, the average path length decreases, but the diameter grows.

# BOX 7.4

## VERTEX COVER AND MUSEUM GUARDS

Imagine you are director of an open-air museum situated in a large park with numerous paths. You wish to place guards on crossroads to observe each path, but to save cost you want to use as few guards as possible. Let $N$ be the number of crossroads and $m < N$ is the number of guards you can afford to hire. There are $(Nm)$ ways of placing the $m$ guards in the $N$ positions, but most configurations will leave some paths unobserved [31].

The number of trials one needs to find a perfect solution grows exponentially with $N$. Indeed, this is one of the six basic NP-complete problems, called the vertex cover problem. By definition, the vertex cover of a network is a set of nodes such that each link is connected to at least one node of the set. The NP-completeness means that there is no known algorithm which can identify a vertex cover substantially faster than using as exhaustive search, i.e. checking each possible configuration individually. Obviously, the number of nodes needed to obtain a vertex cover depends on the network topology, being affected by the degree distribution and potential degree correlations [29].

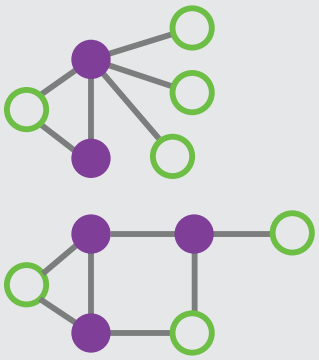


**Figure 7.17**
**The minimum cover**

Formally, a vertex cover of a network $G$ is a set $C$ of nodes such that each link of $G$ connects to at least one node in $C$. A minimum vertex cover is a vertex cover of smallest possible size. The figure above shows examples of minimal vertex covers in two graphs, where the set $C$ is shown in red. One can check that if we turn any of the red nodes into white nodes, we will have at least one link that does not connect to a red node.

# SUMMARY

There are at least three important reasons why we care about degree correlations:

- Degree correlations are present in most real networks SECT. 7.4.
- In the previous chapters we showed how much we can learn about a network by inspecting its degree distribution. Degree correlations force us to go beyond the degree distribution, demonstrating that there are quantifiable patters that govern the way nodes link to each other that are not captured by $p_k$ alone.
- Once present, degree correlations change a network's behavior SECT. 7.6.

Despite the considerable effort devoted to characterizing degree correlations, our understanding of the phenomena is not yet complete. For example, while in SECT. 7.6 we showed how to tune degree correlations, the problem is far from being fully resolved. Indeed, the full degree correlations characterizing a network is contained in the $e_{ij}$ matrix. Generating networks with an arbitrary $e_{ij}$ remains a difficult task.

The results of this chapter allow us to formulate the next network law:

**Structural Correlations**

*Simple scale-free networks are disassortative.*

Let us inspect the validity of this law in the light of the three criteria established in CHAPTER 3:

**A. Quantitative Formulation**

The quantitative basis of this law is provided in SECT. 5.3 and ADVANCED TOPICS 7.B, where we derived the magnitude of the structural cutoff and the emergence of disassortative correlations beyond $k_s$.

### B. Universality

In SECT. 7.4 we showed that many real networks, from biological to email networks, display structural disassortativity.

### C. Non-random Character

As we showed in SECT. 5.3, structural disassortativity cannot appear in the random network model, as the degree of the largest node in a random network is smaller than the structural cutoff $k_s$.

## BOX 7.5

### DEGREE CORRELATIONS: BRIEF HISTORY

Degree correlations were first reported in 2001 in the context of the Internet in a classic paper by Romualdo Pastor-Satorras, Alexei Vazquez, and Alessandro Vespignani [4, 5]. This work introduced the degree correlation function $k_{nn}(k)$ and the scaling Eq. 7.10. A year later Kim Sneppen and Sergey Maslov used the full $p(k_i, k_j)$, rooted in the $e_{ij}$ matrix, to discover the presence of degree correlations in protein-interaction networks [32]. In 2003 Mark Newman introduced the degree correlation coefficient [8, 9], allowing him to realize that two kinds of correlations can emerge in real systems. He also introduced the terminology "assortativity" and "disassortativity" to characterize this diversity. These terms have their roots in social sciences where they are used to capture mating preferences [33].

**Assortative mating**
reflects the tendency of individuals to date or marry individuals that are similar to them. For example, low-income individuals tend to marry low-income individuals, and college graduates marry college graduates. Network theory uses assortativity in the same spirit, capturing the degree-based similarities between nodes: in assortative networks hubs tend to connect to other hubs and small-degree nodes to small-degree nodes. In a network environment we can also encounter the traditional assortativity, when nodes of similar properties link to each other Fig. 7.18.

**Disassortative mixing**
when individuals link to individuals who are unlike them, is also observed in social systems. Sexual networks are perhaps the best example of this phenomena, as most sexual relationships are between individuals of different gender. Disassortative mixing is also common in economic settings. For example, trade typically takes place between individuals of different skills: the baker does not sell bread to other bakers, and the shoemaker rarely fixes other shoemaker's shoes.

**Figure 7.18**
**Politics is rarely neutral**

The network behind the political blogosphere in the US illustrates the presence of assortative mixing, as used in sociology, meaning that nodes of similar characteristics tend to link to each other. In the map each node corresponds to a blog, colored blue if the blog is considered liberal and red if conservative. Blue links connect liberal blogs, red links connect conservative blogs, yellow links go from liberal to conservative, and purple from conservative to liberal. As the image indicates, the linkage patterns is not random: liberal blogs predominantly cite other liberal blogs and conservative blogs connect mainly to conservative-learning blogs. Very few blogs link across the political divide. After [34].

# BOX 7.6

### TWO-POINT, THREE-POINT CORRELATIONS

In their most general form, the degree correlations present in a network are determined by the conditional probability $P(k^{(1)}, k^{(2)}, ..., k^{(k)})$ that a node of degree $k$ connects to nodes with degrees $k^{(1)}, k^{(2)}, ..., k^{(k)}$.

### Two-point correlations

The simplest of these is the two-point degree correlation discussed in this chapter, being the conditional probability $P(k')$ that a node with degree $k$ is connected to a node with degree $k'$. For uncorrelated networks this conditional probability is independent of $k$, hence $P(k') = k'P(k')/\langle k \rangle$ [18]. As the empirical evaluation of $P(k')$ in real networks is a cumbersome task, it is more practical to analyze the degree correlation function $k_{nn}(k)$ defined in Eq. 7.7.

### Three-point correlations

In principle there is no reason to stop at two-point correlations. Correlations involving three nodes are determined by the probability $P(k^{(1)}, k^{(2)} \mid k)$ that a node with degree $k$ is connected to nodes with degrees $k^{(1)}$ and $k^{(2)}$. This conditional probability determines the clustering coefficient Eq. 7.20. Indeed, the average clustering coefficient $C(k)$ of nodes with degree $k$ [22, 23] can be formally written as the probability that a node of degree $k$ is connected to nodes with degrees $k^{(1)}$ and $k^{(2)}$, and that those two are joined by a link, averaged over all the possible values of $k^{(1)}$ and $k^{(2)}$,

$$C(k) = \sum_{k^{(1)},k^{(2)}} P(k^{(1)},k^{(2)} \mid k) p^{k}_{k^{(1)},k^{(2)}},$$

where $p^{k}_{k^{(1)},k^{(2)}}$ is the probability that nodes $k^{(1)}$ and $k^{(2)}$ are connected, provided that they have a common neighbor with degree $k$ [18]. For neutral networks the clustering coefficient is independent of $k$, following

$$C(k) = \frac{\left(\langle k^2 \rangle - \langle k \rangle\right)^2}{\langle k \rangle^3 N}.$$

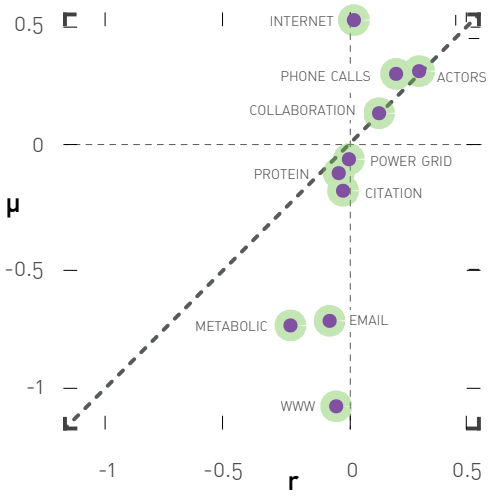



**Figure 7.19**
**Correlation between r and N**

To illustrate the relationship between r and μ, we estimated $\mu$ by fitting the knn function to Eq. 7.10, whether or not the power law scaling was statistically significant Fig. 7.8.

# ADVANCED TOPICS 7.A DEGREE CORRELATION COEFFICIENT

In BOX 7.2 we defined the degree correlation coefficient $r$ as an alternative measure of degree correlations characterizing a network [8, 9]. The use of a single number to characterize degree correlations is extremely attractive, as it also offers an easy way to compare the correlations observed in networks of different nature and size. Yet, before we use $r$ we must be aware of some of its limitations.

The hypothesis behind the correlation coefficient $r$ is that the $k_{nn}(k)$ function can be approximated by the linear function

$$k_{nn}(k) \sim rk\,. \tag{7.21}$$

This is different from the scaling Eq. 7.10, which assumes a power law dependence on $k$. Eq. 7.21 raises several important issues:

- The linear dependence Eq. 7.21 is not supported by empirical data, numerical simulations, or analytical calculations. Indeed, analytical calculation of the initial attractiveness model predict a power law Eq. 7.18 or a logarithmic $k$-dependence Eq. 7.20 for the degree correlation function. Therefore, $r$ forces a linear fit to an inherently nonlinear function. This discrepancy is illustrated in Fig. 7.20, which shows that for assortative and disassortative networks Eq. 7.21 offers a poor to the data.

- As we have seen in Fig. 7.10, the dependence of $k_{nn}(k)$ on $k$ is rather complex, often changing trends for large $k$ thanks to the structural cutoff. A linear fit ignores this inherent complexity. To illustrate the consequences of this phenomena, we calculated $r$ and $\mu$ for the ten reference networks TABLE 7.1. The results are plotted in Fig. 7.19, indicating that while $\mu$ and $r$ correlate for positive $r$, this correlation breaks down for negative $r$.

- As we discuss in BOX 7.8, the maximally correlated model has a vanishing $r$ for large $N$, despite the fact that the network maintains its

| NETWORK | N | r | μ |
|---|---|---|---|
| Internet | 192,244 | 0.03 | 0.56 |
| WWW | 325,729 | -0.05 | -1.11 |
| Power Grid | 4,941 | 0.003 | 0.0 |
| Mobile Phone Calls | 36,595 | 0.21 | 0.33 |
| Email | 57,194 | -0.08 | -0.74 |
| Science Collaboration | 23,133 | 0.13 | 0.16 |
| Actor Network | 702,388 | 0.31 | 0.34 |
| Citation Network | 449,673 | -0.02 | -0.18 |
| E Coli metabolism | 1,039 | -0.25 | -0.76 |
| Protein Interactions | 2,018 | -0.04 | -0.1 |

**Table 7.1**
**Degree correlations in reference networks**

The table shows r and μ for the ten reference networks of **TABLE 4.1**. Directed networks were made undirected to measure $r$ and $\mu$. Alternatively, we can use the directed correlation coefficient to characterize such directed networks **BOX 7.8**.

degree correlations. This suggests that the degree correlation coefficient $r$ has difficulty detecting correlations characterizing large networks.

### RELATIONSHIP BETWEEN $\mu$ AND $r$

If $k_{nn}(k)$ follows the scaling Eq. 7.10, then the sign of the degree coefficient $r$ should agree with the sign of $\mu$. This is supported by Fig. 7.20 as well. To show the origin of this behavior, next we derive a direct relationship between $\mu$ and $r$. To be specific we assume the validity of Eq. 7.10 and determine the value of $r$ for a network with a given correlation exponent $\mu$.

We start by determining a from from Eq. 7.10. We can write the second moment of the degree distribution as

$$\langle k^2 \rangle = \langle k_{nn}(k) k \rangle = \sum_{k'} a k^{\mu+1} p_k = a \langle k^{\mu+1} \rangle ,$$

which leads to

$$a = \frac{\langle k^2 \rangle}{\langle k^{\mu+1} \rangle} .$$

We now calculate $r$ for a network with a given $\mu$:

$$r = \frac{\sum_{k'} k a k^{\mu} q_k - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2}}{\sigma_r^2} = \frac{\sum_{k'} a k^{\mu+2} \frac{p_k}{\langle k \rangle} - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2}}{\sigma_r^2} = \frac{\frac{\langle k^2 \rangle}{\langle k^{\mu+1} \rangle} \frac{\langle k^{\mu+2} \rangle}{\langle k \rangle} - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2}}{\sigma_r^2} =$$

$$= \frac{1}{\sigma_r^2} \frac{\langle k^2 \rangle}{\langle k \rangle} \left( \frac{\langle k^{\mu+2} \rangle}{\langle k^{\mu+1} \rangle} - \frac{\langle k^2 \rangle}{\langle k \rangle} \right). \qquad (7.22)$$

For $\mu = 0$ the term in the last parenthesis vanishes, obtaining $r = 0$. Hence if $\mu = 0$ (neutral network), the network will be neutral based on $r$ as well. For $k > 1$ Eq. 7.22 suggests that for $\mu > 0$ the parenthesis is positive, hence $r > 0$, and for $\mu < 0$ is negative, hence $r < 0$. Therefore, $r$ and $\mu$ predict degree correlations of similar kind.

Therefore, if the degree correlation function follows Eq. 7.10, then the sign of the degree correlation exponent μ will determine the sign of the assortativity coefficient $r$:

$$\mu < 0 \rightarrow r < 0$$
$$\mu = 0 \rightarrow r = 0$$
$$\mu > 0 \rightarrow r > 0 .$$

In summary, the degree correlation coefficient assumes that $k_{nn}(k)$ scales linearly with $k$, a hypothesis that lacks numerical and analytical support. Hence $r$ forces a linear fit to $k_{nn}(k)$, giving occasionally rise to inconsistent results. While typically the sign of $r$ and $\mu$ agree, overall $r$ does not offer a natural characterization of the underlying degree correlations. An accurate characterization starts with $e_{ij}$, whose behavior is reasonably captured by $k_{nn}(k)$.

## BOX 7.7

### AT A GLANCE: DEGREE CORRELATIONS

**Degree Correlation Matrix $e_{ij}$**

probability of finding a node with degrees $i$ and $j$ at the two ends of a link.

Neutral networks:

$$e_{ij} = q_i q_i = \frac{k_i p_{k_i} k_j p_{k_j}}{\langle k \rangle^2}$$

**Degree Correlation Function**

$$k_{nn}(k) = \sum_{k'} k' p(k' | k)$$

Neutral networks:

$$k_{nn}(k) = \frac{\langle k^2 \rangle}{\langle k \rangle}$$

**Scaling Hypothesis**

$$k_{nn}(k) \sim k^{\mu}$$

$\mu > 0$: *Assortative*
$\mu = 0$: *Neutral*
$\mu < 0$: *Dissasortative*

**Degree Correlation Coefficient**

$$r = \sum \frac{jk(e_{jk} - q_j q_k)}{\sigma_r^2}$$

$r > 0$: *Assortative*
$r = 0$: *Neutral*
$r < 0$: *Dissasortative*

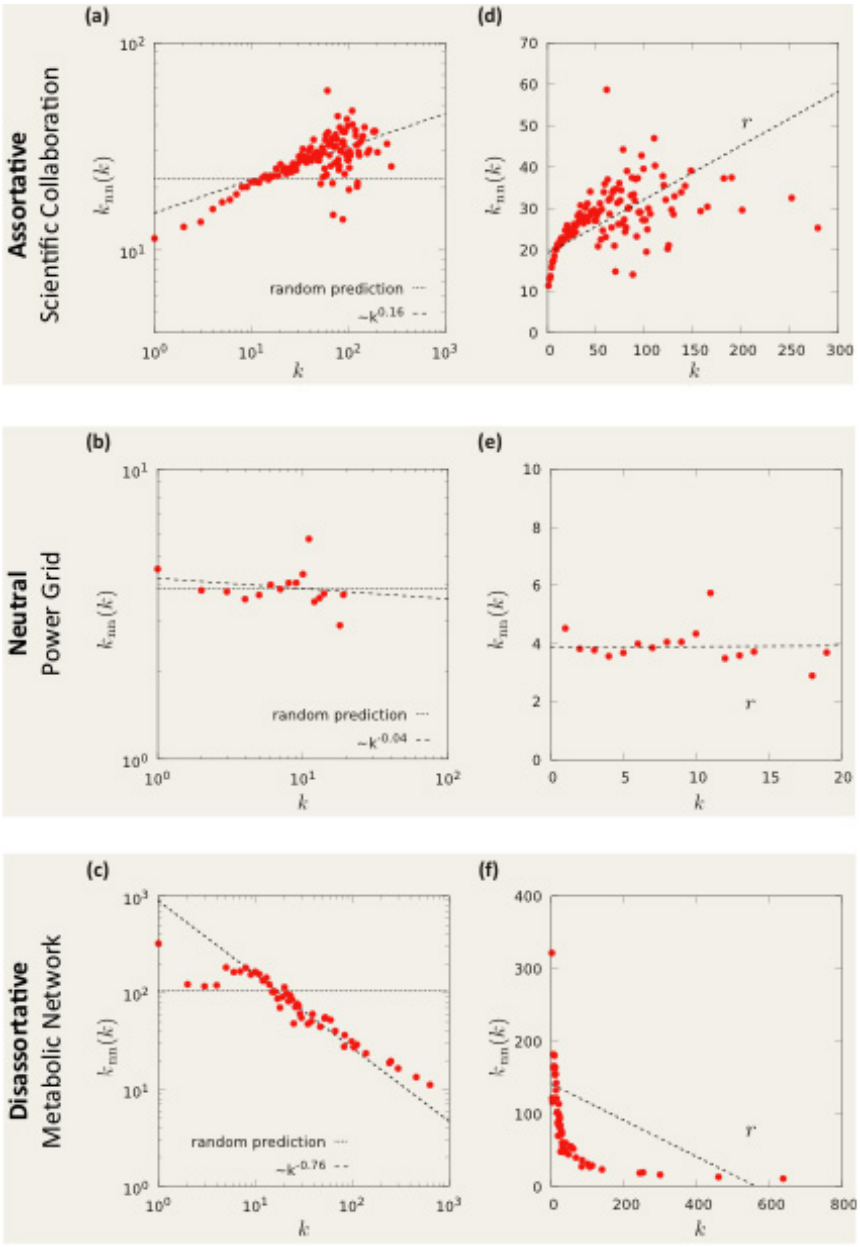



**Figure 7.20**
**Degree correlation function**

The degree correlation function knn for three real networks. The left panels show the cumulative function $k_{nn}(k)$ on a log-log plot to test the validity of **Eq. 7.10**. The right panels show $k_{nn}(k)$ on a *lin–lin* plot to test the validity of **Eq. 7.21**, i.e. the assumption that $k_{nn}(k)$ depends linearly on $k$, the hypothesis behind the correlation coefficient $r$. The slope of the dotted line corresponds to the correlation coefficient $r$. As the lin-lin plots illustrate, **Eq. 7.21** offers a poor fit for assortative (d) and disassortative (f) networks.

# BOX 7.9

## CORRELATION COEFFICIENT FOR DIRECTED NETWORKS

To measure correlations in directed networks we must take into account that each node $i$ is characterized by an incoming $k_i^{in}$ and an outgoing $k_i^{out}$ degree. Hence, we can define four degree correlation coefficients, $r_{in,\,in}$, $r_{in,out}$, $r_{out,in}$, $r_{out,\,out}$ capturing all possible combinations between the incoming and outgoing degrees of two nodes linked to each other Figs. 7.12 a-d. Formally we have [14].

$$r_{\alpha,\beta} = \frac{\sum_{jk} jk(e_{jk}^{\alpha,\beta} - q_j^{\alpha} q_k^{\beta})}{\sigma^{\alpha}\sigma^{\beta}}, \qquad (7.23)$$

where $\alpha$ and $\beta$ refer to the in and out indices. To illustrate the use of Eq. 7.23, we show in Fig. 7.21e the four correlation coefficients for the five directed reference networks TABLE 7.1. For a complete characterization of degree correlations, it is desirable to measure the four $k_{nn}(k)$ functions as well BOX 7.2.

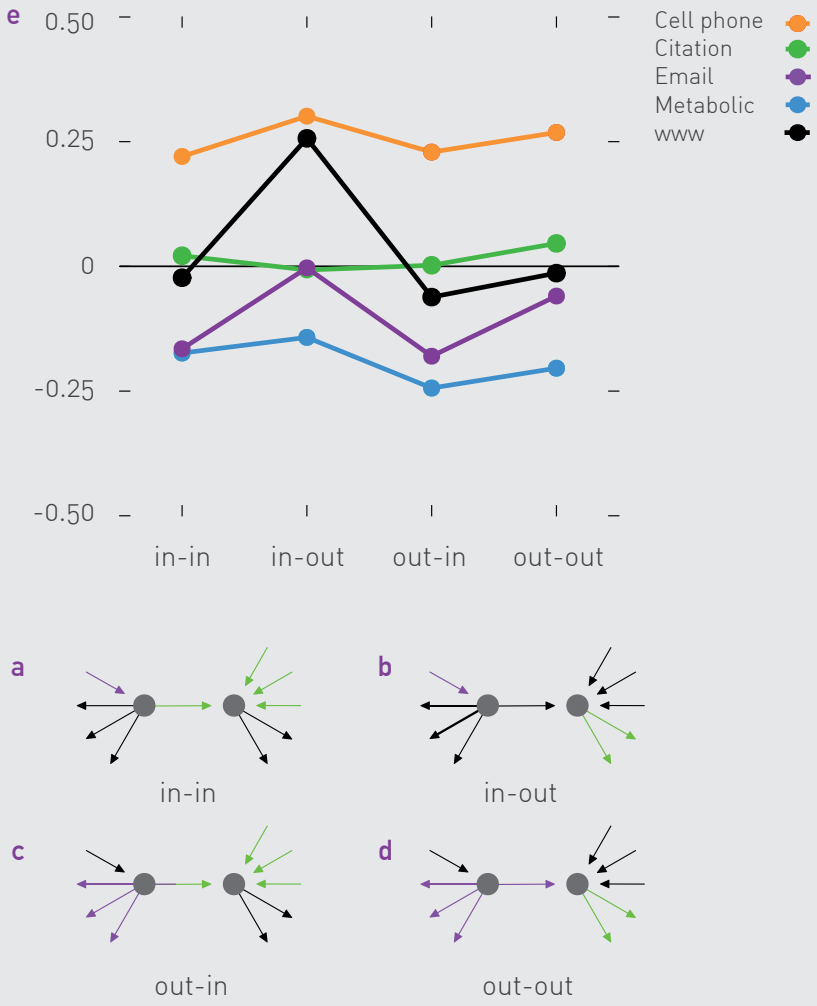



**Figure 7.21**

Panels (a)-(d) illustrate in red and green the ($\alpha$, $\beta$) indices that define the appropriate correlation coefficient for directed ne tworks. (e) The correlation profile of the five directed reference networks, indicating, for example, that while citation networks have negligible correlations, all four correlation coefficients document strong assortative behavior for cell phone calls and strong disassortative behavior for metabolic networks. The case of the WWW is particularly interesting: while three of its correlation coefficients are close to zero, there is a strong assortative tendency for the (in, out) combinations.

# ADVANCED TOPICS 7.B
# STRUCTURAL CUTOFFS

As discussed in SECT. 7.3, there is a fundamental conflict between the scale-free property and degree correlations, which leads to a structural cutoff in simple networks. In this section we derive Eq. 7.16, providing the system size dependence of the structural cutoff [11]. We start by defining

$$r_{kk'} = \frac{E_{kk'}}{m_{kk'}}, \tag{7.24}$$

where $E_{kk'}$ is the number of links between nodes of degrees $k$ and $k'$, and

$$m_{kk'} = \min\left\{kN_k, k'N_k, N_kN_{k'}\right\} \tag{7.25}$$

is the largest possible value of $E_{kk'}$. If multiple links are allowed, $m_{kk'}$ is simply $m_{kk'} = min\{k\,N_k, k'N_k, N_kN_{k'}\}$. The origin of Eq. 7.25 is explained in Fig. 7.22. Consequently, we can write the $r_{kk'}$ ratio as

$$r_{kk'} = \frac{E_{kk'}}{m_{kk'}} = \frac{\langle k\rangle P(k,k')}{\min\left\{kP(k), k'P(k'), NP(k)P(k')\right\}}. \tag{7.26}$$

As $m_{kk'}$ is the maximum of $E_{kk'}$, $r_{kk'}$ must be smaller than or equal to one for any $k$ and $k'$. Yet, for some networks and for some $k$, $k'$ pairs $r_{kk'}$ becomes larger than one. This is clearly non-physical and signals some conflict in the network configuration. Strictly speaking, in simple networks degree pairs for which $r_{kk'} > 1$ cannot exist. Hence, we define the structural cut off $k_s$ as the solution of the equation

$$r_{k_s k_s} = 1. \tag{7.27}$$

Note that as soon as $k > NP(k')$ and $k' > NP(k)$, the effects of the restriction on the multiple links are already felt, turning the expression for $r_{kk'}$ into

$$r_{kk'} = \frac{\langle k\rangle P(k,k')}{Np_k p_{k'}}. \tag{7.28}$$

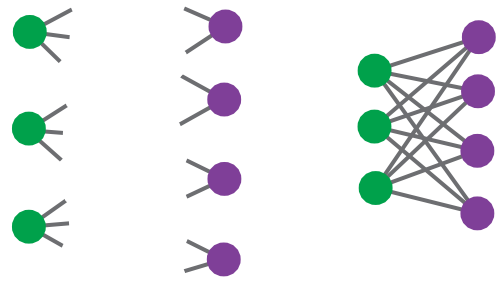



**Figure 7.22**
**Correlation between $r$ and $N$**

Illustrating the maximum number of links one can have between two groups of nodes. The figure shows two groups of nodes, with degree $k$=3 and $k'$=3. The total number of links between these two groups must not exceed

**(a)** The total number of links available in $k$=3 group, which is $kN_k$=9;

**(b)** The total number of links available in $k'$=2 group, which is $k'N_{k'}$=8;

**(c)** The total number of links one can potentially have between the two groups, which is $N_kN_{k'}$.

In the example shown above the smallest of the three is $k'N_k$ = 8 of (b). The resulting configuration is shown on the top right. One can see that in this configuration, one link in the $k$=3 class remains unpaired.

For scale-free networks these conditions are fulfilled in the region $k, k' > (aN)^{1/(\gamma+1)}$, where a is a constant that depends on the function $p_k$. Note that this value is below the natural cut off. As a consequence, this scaling provides a lower bound for the structural cut off, in the sense that whenever the cut off of the degree distribution falls below this limit, the condition $r_{kk'} < 1$ is always satisfied.

For neutral networks the joint distribution factorizes as

$$P(k,k') = \frac{kk' p_k p_{k'}}{\langle k \rangle^2} . \tag{7.29}$$

Hence, the ratio $r_{kk'}$ of Eq. 7.28 takes the form

$$r_{kk'} = \frac{kk'}{\langle k \rangle N} . \tag{7.30}$$

Therefore, the structural cutoff needed to preserve the condition $r_{kk'} \leq 1$ has the form [35, 36, 37]

$$k_s(N) \sim (\langle k \rangle N)^{1/2} , \tag{7.31}$$

which is Eq. 7.16. Note that Eq. 7.31 is independent of the degree distribution of the underlying network. Consequently, for a scale-free network $k_s(N)$ is independent of the degree exponent $\gamma$.

# BIBLIOGRAPHY

[1] P. Uetz, L. Giot, G. Cagney, T. A. Mansfield, RS Judson, JR Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, J. M. Rothberg. A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature 403: 623–627, 2000.

[2] I. Xenarios, D. W. Rice, L. Salwinski, M. K. Baron, E. M. Marcotte, D. Eisenberg DIP: the database of interacting proteins. Nucleic Acids Res. 28: 289–29, 2000.

[3] H. Jeong, S.P. Mason, A.-L. Barabási, Z.N. Oltvai, Nature 411: 41-42, 2001.

[4] R. Pastor-Satorras, A. Vázquez, and A. Vespignani. Dynamical and correlation properties of the Internet. Phys. Rev. Lett. 87: 258701, 2001.

[5] A. Vazquez, R. Pastor-Satorras, and A. Vespignani. Large-scale topological and dynamical properties of Internet. Phys. Rev. E 65: 066130, 2002.

[6] S.L. Feld. Why your friends have more friends than you do. American Journal of Sociology 96: 1464–1477, 1991.

[7] E.W. Zuckerman and J.T. Jost. What makes you think you're so popular? Self evaluation maintenance and the subjective side of the "friendship paradox". Social Psychology Quarterly 64: 207–223, 2001.

[8] M. E. J. Newman. Assortative mixing in networks, Phys. Rev. Lett. 89: 208701, 2002.

[9] M. E. J. Newman. Mixing patterns in networks. Phys. Rev. E 67: 026126, 2003.

[10] S. Maslov, K. Sneppen, and A. Zaliznyak, Pattern detection in complex networks: Correlation profile of the Internet, e-print cond-mat/0205379, 2002.

[11] M. Boguna, R. Pastor-Satorras, A. Vespignani. Cut-offs and finite size effects in scale-free networks. Eur. Phys. J. B 38: 205, 2004.

[12] M. E. J. Newman and Juyong Park. Why social networks are different from other types of networks, arXiv:cond-mat/0305612v1.

[13] M. McPherson, L. Smith-Lovin, J. M. Cook. Birds of a feather: homophily in social networks. Annual Review of Sociology 27:415-444, 2001.

[14] J. G. Foster, D. V. Foster, P. Grassberger, M. Paczuski. Edge direction and the structure of networks. PNAS 107: 10815, 2010.

[15] A. Barrat and R. Pastor-Satorras. Rate equation approach for correlations in growing network models. Phys. Rev. E 71, 036127, 2005.

[16] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. Adv. Phys. 51: 1079, 2002.

[17] J. Berg and M. Lässig. Correlated random networks. Phys. Rev. Lett. 89: 228701, 2002

[18] M. Boguñá and R. Pastor-Satorras. Class of correlated random networks with hidden variables. Phys. Rev. E 68: 036112, 2003.

[19] R. Xulvi-Brunet and I. M. Sokolov. Reshuffling scale-free networks: From random to assortative. Phys. Rev. E 70: 066102, 2004.

[20] R. Xulvi-Brunet and I. M. Sokolov. Changing correlations in networks: assortativity and dissortativity. Acta Phys. Pol. B 36: 1431, 2005.

[21] J. Menche, A. Valleriani, and R. Lipowsky. Asymptotic properties of degree-correlated scale-free networks. Phys. Rev. E 81: 046103, 2010.

[22] V. M. Eguíluz and K. Klemm. Epidemic threshold in structured scale-free networks. Phys. Rev. Lett. 89:108701, 2002.

[23] M. Boguñá and R. Pastor-Satorras. Phys. Epidemic spreading in correlated complex networks. Rev. E 66: 047104, 2002.

[24] M. Boguñá, R. Pastor-Satorras, and A. Vespignani. Absence of epidemic threshold in scale-free networks with degree correlations. Phys. Rev. Lett. 90: 028701, 2003.

[25] A. Vázquez and Y. Moreno. Resilience to damage of graphs with degree correlations. Phys. Rev. E 67: 015101R, 2003.

[26] S.J.Wang,A.C.Wu,Z.X.Wu,X.J.Xu,andY.H.Wang. Response of degree-correlated scale-free networks to stimuli. Phys. Rev. E 75: 046113, 2007.

[27] F. Sorrentino, M. Di Bernardo, G. Cuellar, and S. Boccaletti. Synchronization in weighted scale-free networks with degree–degree correlation. Physica D 224: 123, 2006.

[28] M. Di Bernardo, F. Garofalo, and F. Sorrentino. Effects of degree correlation on the synchronization of networks of oscillators. Int. J. Bifurcation Chaos Appl. Sci. Eng. 17: 3499, 2007.

[29] A. Vazquez and M. Weigt. Computational complexity arising from degree correlations in networks, arXiv:cond-mat/0207035, 2002.

[30] M. Posfai, Y Y. Liu, J-J Slotine. A.-L. Barabási. Effect of correlations on network controllability, Scientific Reports 3: 1067,2013.

[31] M. Weigt and A. K. Hartmann. The number of guards needed by a museum: A phase transition in vertex covering of random graphs. Phys. Rev. Lett. 84: 6118, 2000.

[32] S. Maslov and K. Sneppen. Specificity and stability in topology of protein networks. Science 296: 910–913, 2002.

[33] M. McPherson, L. Smith-Lovin, and J. M. Cook. Birds of a feather: homophily in social networks. Annual Review of Sociology 27: 415–444, 2001.

[34] L. Adamic and N. Glance, The political blogosphere and the 2004 U.S. election: Divided they blog (2005).

[35] J. Park and M. E. J. Newman. The origin of degree correlations in the Internet and other networks. Phys. Rev. E 66: 026112, 2003.

[36] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. Annals of Combinatorics, 6: 125, 2002.

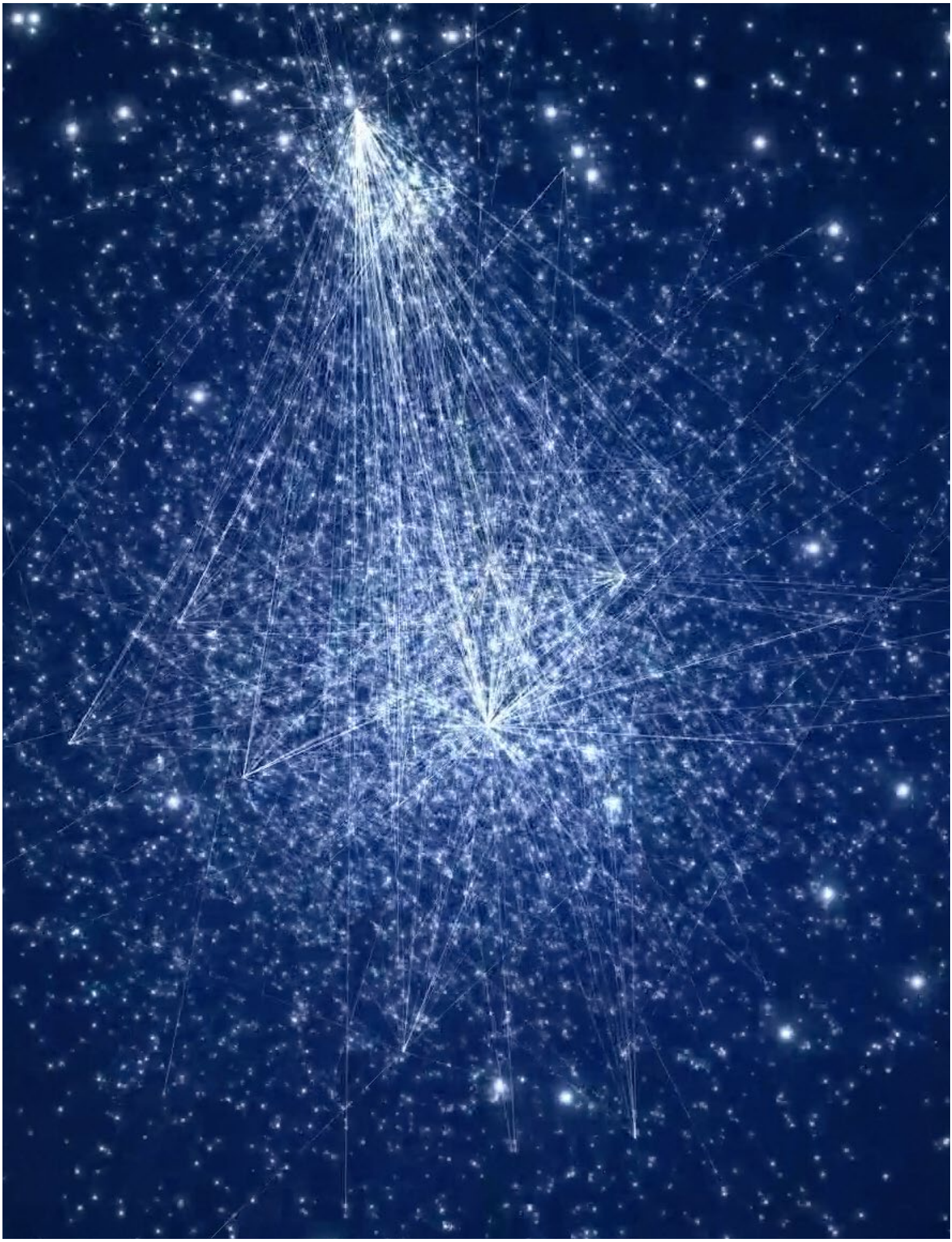
[37] Z. Burda and Z. Krzywicki. Uncorrelated random networks. Phys. Rev. E 67: 046118, 2003.

# CHAPTER 8

# NETWORK ROBUSTNESS



**Figure 8.0 (front cover)**
**Network representation by Mauro Martino**



# INTRODUCTION

Errors and failures can corrupt all human designs: the failure of a single component in your car's engine may force you to call for a tow truck or a wiring error in your computer chip can make your computer useless. Many natural and social systems, however, have a unique ability to sustain their basic functions even when several of their components fail. Indeed, there are countless protein misfolding errors or missed reactions in our cells, often without noticeable consequences; large organizations can function despite numerous missing employees. Understanding the origins of this robustness is important for many disciplines:

- Robustness is a central question in biology, which aims to understand how a cell or an organism functions under frequent internal errors and why some errors lead to diseases.

- It is of concern for social scientists and economists, who explore the stability of human societies and organizations in the face of famine, war, and changes in social and economic order.

- It is a key issue for ecologists and environmental scientists, who seek to estimate the chances that an ecosystem survives when faced with the disruptive effects of human activity.

- It is the ultimate goal in engineering, aiming to design communication systems, cars, or airplanes that can maintain a high readiness despite occasional component failures.

These biological, social and technological systems share a common feature: their functionality and robustness is guaranteed by densely interlinked networks. Indeed, cellular functions are encoded by intricate regulatory and metabolic networks; the society's resilience cannot be divorced from the interwoven social, professional, and communication web behind it; economic stability is guarded by a delicate network of financial and regulatory organizations; an ecosystem's survivability cannot be understood without a careful analysis of the food webs that sustain each species.


**Figure 8.1**
**Achilles' Heel of Complex Networks**

The cover of the 27 July, 2000 issue of *Nature*, highlighting the paper entitled *Attack and error tolerance of complex networks* that sparked the interest in network robustness [1].

Whenever nature seeks robustness, it resorts to networks to achieve it.

The purpose of this chapter is to explore the role networks play in ensuring the robustness of a complex system. We show that understanding the structure of the underlying network is essential if we want to quantify a system's ability to survive random failures or deliberate attacks. We also explore the role of these networks in the emergence of cascading failures, a damaging phenomenon frequently encountered in real systems. Most importantly, we show that the laws governing the error and attack tolerance of complex networks and the emergence of cascading failures are universal.

**Figure 8.2**
**Robust Robustness**

"Robust" comes from the latin Quercus Robur, meaning oak, the symbol of strength and longevity in the ancient world.

# PERCOLATION THEORY

Robustness requires us to understand the impact of node or link removal on the integrity of a network.The removal of a single node typically has only limited impact on a network's integrity Fig. 8.3a. The removal of multiple nodes, however, can break a network into isolated, non-communicating subgraphs Fig. 8.3c, d. Obviously, the more nodes we remove, the higher are the chances that we damage a network, prompting us to ask: How many nodes do we have to delete to fragment a network into isolated components? For example, what fraction of Internet routers must break down so that the Internet turns into isolated clusters of computers that are unable to communicate with each other? To answer these questions, we first introduce some concepts of percolation theory that offers the mathematical underpinnings of the network robustness problem.

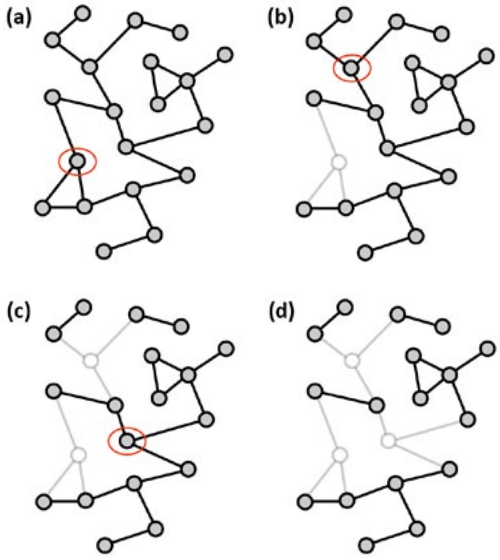



**Figure 8.3**
**The impact of node removal**

The gradual fragmentation of a small network following the breakdown of several nodes. In each panel we remove a new node (highlighted), together with its links. As the sequence of images indicates, while the removal of the first node has only limited impact on the network's integrity, the removal of the second node isolates two small clusters from the rest of the network and the removal of the third node fragments the network, breaking it into five non-communicating clusters of sizes $s$ = 2, 2, 2, 5, 6.

### PERCOLATION

Percolation theory is a highly developed subfield of statistical physics and mathematics [2, 3, 4]. A typical problem addressed by it is illustrated in Fig. 8.4, showing a square lattice with pebbles placed with probability $p$ at each intersection. Pebbles next to each other are considered connected, forming clusters of size two or more. Given that the position of each pebble is decided by chance, we ask:

- What is the size of the largest cluster?
- What is the average cluster size?

Obviously, the higher is $p$, the larger are the individual clusters. Percolation theory predicts, however, that the cluster size does not change continuously with $p$. Rather, for a wide range of $p$ values the lattice is populated with numerous tiny clusters Fig. 8.4a. If we increase $p$ beyond a critical value $p_c$, these small clusters grow rapidly until a single large cluster emerges rather suddenly. We call this the percolating cluster as it percolates through the lattice by reaching its ends. In other words, at $p_c$ we observe a phase transition from many small clusters to a percolating cluster that spans the whole lattice Fig. 8.4b.

To quantify the nature of this phase transition, we focus on several frequently measured quantities:

- The average cluster size, $\langle s \rangle$, represents the average size of all finite clusters observed for a given $p$. Percolation theory predicts that in the vicinity of $p_c$ it follows

$$\langle s \rangle \sim |p - p_c|^{-\gamma_p} . \tag{8.1}$$

In other words, the average cluster size diverges as we approach $p_c$ Fig. 8.4c.

The *order parameter*, $p_\infty$, represents the probability that a randomly chosen pebble belongs to the largest cluster. In the vicinity of the critical point $p_\infty$ follows

$$P_\infty \sim (p - p_c)^\beta . \tag{8.2}$$

Hence, as $p$ decreases towards $p_c$ the probability that a pebble belongs to the largest cluster drops zero Fig. 8.4d.

- The *correlation length*, $\varsigma$, represents the mean distance between two pebbles that belong to the same cluster. In the vicinity of $p_c$ it follows

$$\zeta \sim |p - p_c|^{-\nu} . \tag{8.3}$$

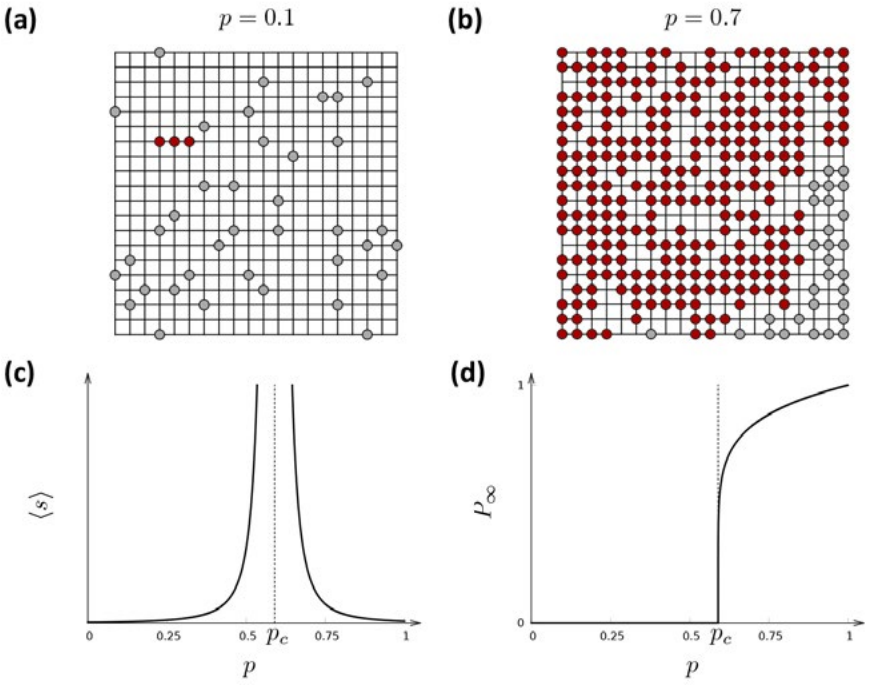



**Figure 8.4**
**Percolation**

A classical problem in percolation theory explores the random placement with probability $p$ of pebbles on a square lattice.

**(a)** For small $p$ most peebles are isolated. In this case the largest cluster has only three nodes, shown in red.

**(b)** For large $p$ most (but not all) pebbles belong to a single giant component, colored red. This giant component is called the percolating cluster, as it spans the whole lattice (See also Fig. 8.6).

**(c)** The average cluster size, $\langle s \rangle$, in function of p. As we approach pc from below, the numerous small clusters coalesce and $\langle s \rangle$ diverges. The same divergence is observed above pc, where to measure $\langle s \rangle$ we remove the largest component from the average. The plot shows schematically the divergence of $\langle s \rangle$ as described by Fig. 8.1. The same exponent $\gamma_p$ characterizes the divergence at both sides of the critical point.

**(d)** The $p$–dependence of the probability $p_\infty$ that a pebble belongs to the largest connected component. For $p < p_c$ all components are small, so $p_\infty$ is effectively zero. Once $p$ reaches $p_c$ a giant component emerges. Consequently beyond $p_c$ there is a finite probability that a node belongs to the largest component, as predicted by Fig. 8.2.

While at $p < p_c$ the distance between the peebles in the same cluster is finite, at $p_c$ the correlation length diverges. Therefore, at $p_c$, the linear size of the largest cluster becomes infinite, which is the reason it percolates the whole lattice.

The exponents $\gamma_p$, $\beta$, and $\nu$ are called critical exponents, as they characterize the system's behavior near the critical point $p_c$. Percolation theory predicts that these exponents are universal, meaning that they are independent of the nature of the lattice on which we place the peebles or the precise value of $p_c$. Therefore, if we place the peebles on a triangular or a hexagonal lattice, the behavior of $\langle s \rangle$, $P_\infty$, and $\zeta$ is characterized by the same $\gamma_p$, $\beta$, and $\nu$ exponents. Consider the following examples to better understand this universality:

- The exponents depend only on the dimension of the lattice. In two dimensions, the case illustrated in Fig. 8.4, we have $\gamma_p = 43/18$, $\beta = 5/36$, $\nu = 4/3$, while in three dimensions $\gamma_p = 1.80$, $\gamma = 0.41$, $\gamma = 0.88$ . For $d > 6$ we have $\gamma_p = 1$, $\gamma = 1$, $\gamma = 1/2$ [2], i.e. beyond $d = 6$ the exponents are independent of $d$.

- The value of $p_c$ is not universal, as it depends on the lattice type. For example, for a two-dimensional square lattice Fig. 8.4 we have $p_c \approx 0.593$, while for two-dimensional triangular lattice we have $p_c = 1/2$ (site percolation).

- The value of $p_c$ also changes with the dimension: for a square lattice we have $p_c \approx 0.593$ (d = 2); for a simple cubic lattice ($d = 3$) we have $p_c \approx 0.3116$. Therefore, in $d = 3$, we need to cover a smaller fraction of the nodes with pebbles to reach the percolation transition.

#### ROBUSTNESS AS AN INVERSE PERCOLATION TRANSITION

We can use percolation theory to describe the impact of node failures in networks, the phenomena of primary interest in robustness. For this we view a square lattice as a network whose nodes are the intersections Fig. 8.5. Next, we randomly remove an $f$ fraction of nodes, asking how their absence impacts the integrity of the lattice. If $f$ is small, the few missing nodes do little damage to the network. Increasing $f$, however, can remove chunks of nodes from the giant component. Finally, for sufficiently large $f$ the giant component breaks into tiny disconnected components.

Once again, the fragmentation process is not gradual, but it is characterized by a critical threshold $f_c$: for $f < f_c$ we continue to have a giant component, but once $f$ exceeds $f_c$, the giant component vanishes. This is illustrated by the $f$-dependence of $P_\infty$, representing the probability that a node is part of the giant component Fig. 8.5: $P_\infty$ is finite under $f_c$, but it drops to zero as we approach $f_c$. The critical exponents characterizing this breakdown, $\gamma_p$, $\beta$, $\mu$, are the same as those encountered in Eq. 8.1-8.3, as the two processes can be mapped into each other by choosing $f = 1 - p$. Furthermore, in ADVANCED TOPICS 8.A, we show that

random networks under random node failures share the same scaling exponents as infinite-dimensional percolation. This equivalence predicts that the value of the critical exponents for a random network are $\gamma_p = 1$, $\beta = 1$ and $\nu = 1$, the $d > 6$ values encountered earlier.

In summary, the breakdown of a network under random node removal is not a gradual process. In general, removing a small fraction of nodes has limited impact on a network's integrity. Once the number of removed nodes reaches a critical threshold, the network abruptly breaks into disconnected components. In other words, random node failures induce a phase transition from a connected to a fragmented state. We can use the tools of percolation theory to characterize this transition in both regular and in random networks.

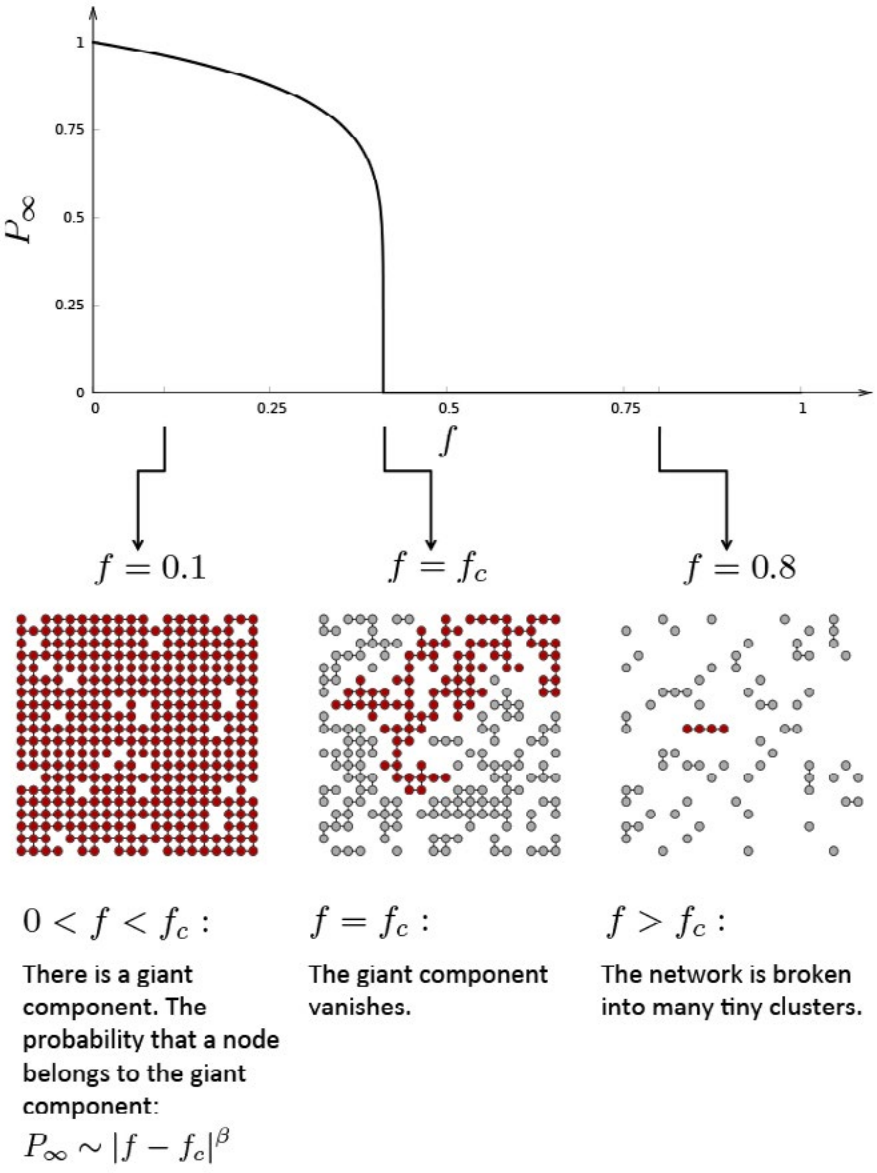



**Figure 8.5**
**Network breakdown as inverse percolation**

The consequences of node removal are accurately captured by the inverse of the percolation process discussed in **Fig. 8.4**. We start from a square lattice, viewed as a network whose nodes are the intersections. Next we randomly select and remove an $f$ fraction of nodes, measuring the size of the largest component formed by the remaining nodes, captured by $P_\infty$. The obtained networks are illustrated on the three bottom panels. Under each panel we list the characteristics of the corresponding phases.

# BOX 8.1

## PERCOLATION THEORY: A BRIEF HISTORY

Percolation theory began with a paper written by the mathematicians Simon Broadbent and John Hammersey in 1957, who proposed its name and formalized many of its mathematical concepts [5]. The theory rose to particular prominence in the 1960 and 70s with the development of critical phenomena in physics and the recognition that percolation offers an analytically treatable example of phase transitions. It also found important applications from oil exploration to transport phenomena in physics.

The spread of a fire in a forest is often used to illustrate the basic concepts of percolation theory. Let us assume that each pebble in Fig. 8.4 is a tree, hence the whole lattice describes a forest. If a tree catches fire, it ignites the neighboring trees; these, in turn ignite their neighbors. The fire continues to spread until no burning tree has a non-burning neighbor. The question we ask is the following: if we randomly ignite a tree, what fraction of the forest do we expect to burn down? And how long it takes the fire to burn out? The answer depends on the tree density, controlled by the parameter $p$. For small $p$ the forest consists of many small islands of trees ($p$ = 0.55 in Fig. 8.6), hence igniting any tree will only burn down the small cluster containing the igniting tree.

Consequently, the fire will die out quickly. For very large $p$ most trees belong to a single large cluster, hence the fire will rapidly sweep through much of the dense forest (see $p$ = 0.62 in Fig. 8.6). The simulations
indicate that there is a critical $p_c$, for which it takes extremely long time for the fire to end. This $p_c$ is the critical threshold of the percolation problem. Indeed, at $p = p_c$ the giant component just emerged through the union of many small clusters. Hence the fire has to follow a long and winding path to reach all clusters and all trees, a process that can be rather time consuming.

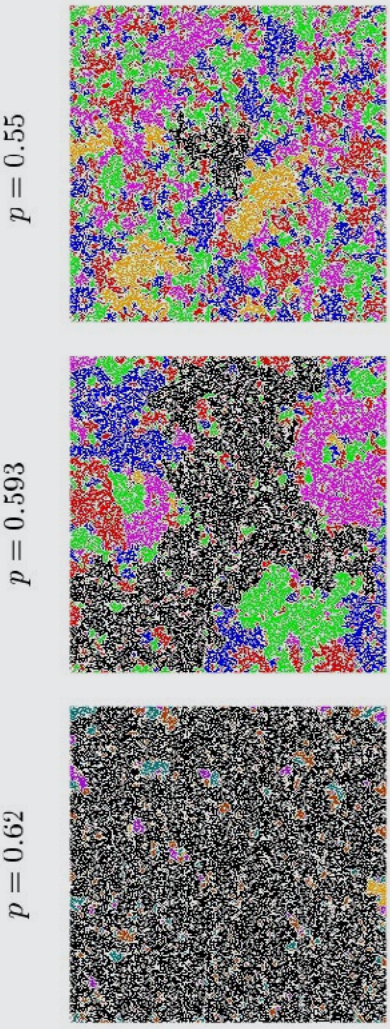



**Figure 8.6**
**Forrest Fire**

The emergence of the giant component on a square lattice as we change the occupation probability $p$. Each panel corresponds to a different p in the vicinity of $p_c$ . The largest cluster is shown in black. For $p < p_c$ the largest cluster is tiny, as seen on the top panel. If we view this as a forest, where the pebbles are trees, a fire can at most consume only a small fraction of the trees, hence it burns out quickly. Once $p$ reaches $p_c \approx 0.593$, however, the largest cluster percolates the whole lattice and the fire can "percolate" through the forest. Increasing $p$ beyond $p_c$ "fattens" the largest cluster, connecting more pebbles (trees) to it, as seen for $p$ = 0.62 on the bottom panel. Hence, the fire burns out quickly again.

# ROBUSTNESS OF SCALE-FREE NETWORKS

Percolation theory was developed either for regular lattices, whose nodes have identical degrees, or for random networks, whose nodes have comparable degrees. What happens, however, if the network is scale-free? Will the hubs affect the percolation transition? We can get a hint from a simulation testing the Internet's robustness to router failures [1]. We start from the router level map of the Internet and randomly select and remove nodes one-by-one. Percolation theory predicts that once the number of removed nodes reaches a critical value $f_c$, the Internet should fragment into many isolated subgraphs. The simulations indicate otherwise: the Internet refuses to break apart even under rather extensive random node removal. Instead, the size of the largest component decreases gradually, vanishing only in the vicinity of $f = 1$ Fig. 8.7a. Hence, the network behind the Internet shows an unusual robustness to random router failures: we must remove all nodes to destroy its giant component. This conclusion disagrees with percolation theory, which predicts that networks must fall apart after the removal of a finite fraction of nodes.

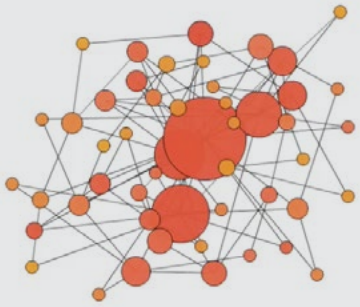
The behavior observed above is not unique to the Internet. Indeed,in Fig. 8.7b we show $P_\infty$ for a scale-free network with degree exponent $\gamma = 2.5$, observing a similar pattern: under random node removal the giant component vanishes only in the vicinity of $f = 1$, rather than collapsing at some finite $f_c$. This indicates that the robustness observed for the Internet is a property of the scale-free topology. The goal of this section is to uncover and quantify the source and the characteristics of this remarkable robustness.

### MALLOY-REED CRITERIA

To understand the origin of the anomalously high $f_c$ for the Internet and for a scale-free network we must first calculate $f_c$ for a network with an arbitrary degree distribution. To do so we rely on a simple observation: if a network has a giant component, then most nodes that belong to it must be connected to at least two other nodes Fig. 8.8. This leads to the Malloy-Reed criteria ADVANCED TOPICS 8.B, stating that the condition for the existence of a giant component is [6]

## MOVIE 8.1

**SCALE-FREE NETWORK UNDER NODE FAILURES**



To illustrate the robustness of a scale-free network we start from the network we constructed in Movie 4.1 using the Barabási-Albert model. Next we randomly select and remove nodes one-by-one. As the movie illustrates, despite the fact that we remove a significant fraction of the nodes, the network refuses to break apart. The origin of this robustness to random failures is the topic of SECTION 8.2. Visualization by Dashun Wang.

$$\kappa \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} > 2. \tag{8.4}$$

Consequently networks with $\kappa < 2$ must be fragmented into many disconnected components. The Malloy-Reed criteria links the network's integrity, as expressed by the presence or the absence of a giant component, to $\langle k \rangle$ and $\langle k^2 \rangle$, which depend only on the degree distribution $p_k$. To illustrate the predictive power of Fig. 8.4 we apply it to a random network, for which $\langle k^2 \rangle = \langle k \rangle (1 + \langle k \rangle)$. Hence, a random network has a giant component if

$$\kappa \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{\langle k \rangle (1 + \langle k \rangle)}{\langle k \rangle} = 1 + \langle k \rangle > 2. \tag{8.5}$$

or

$$\langle k \rangle > 1. \tag{8.6}$$

The condition coincides with Eq. 3.10, the necessary condition for the existence of a giant component.

### ROBUSTNESS OF SCALE-FREE NETWORKS

To understand the mathematical origin of the robustness observed in Fig. 8.8, we ask at what threshold $f_c$ will a scale-free network loose its giant component. To answer this we apply the Malloy-Reed criteria to a network with an arbitrary degree distribution. We find that the critical threshold follows [7] (ADVANCED TOPICS 8.C)

$$f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}. \tag{8.7}$$

The most remarkable prediction of Fig. 8.7 is that we can calculate the critical threshold $f_c$ depends only from $\langle k \rangle$ and $\langle k^2 \rangle$, quantities that are uniquely determined by the degree distribution $p_k$. Let us illustrate the utility of Fig. 8.7 by calculating the breakdown threshold of a random network. Using $\langle k^2 \rangle = \langle k \rangle (\langle k \rangle + 1)$, we obtain

$$f_c^{ER} = 1 - \frac{1}{\langle k \rangle}. \tag{8.8}$$

Hence, the denser is a random network, the higher is $f_c$, i.e. the more nodes we need to remove to break it apart. Eq. 8.8 also predicts that a random network always has a finite $f_c$, consequently it always breaks apart after the removal of a finite fraction of nodes.

Most important, Fig. 8.7 helps us understand the roots of the enhanced robustness observed in Fig. 8.7. Indeed, for scale-free networks with $\gamma < 3$ in the $N \to \infty$ limit the second moment diverges. If we insert $\langle k^2 \rangle \to \infty$ into Fig. 8.7, we find that $f_c$ converges to $f_c = 1$. This means that to fragment a scale-free network we must remove all of its nodes. In other

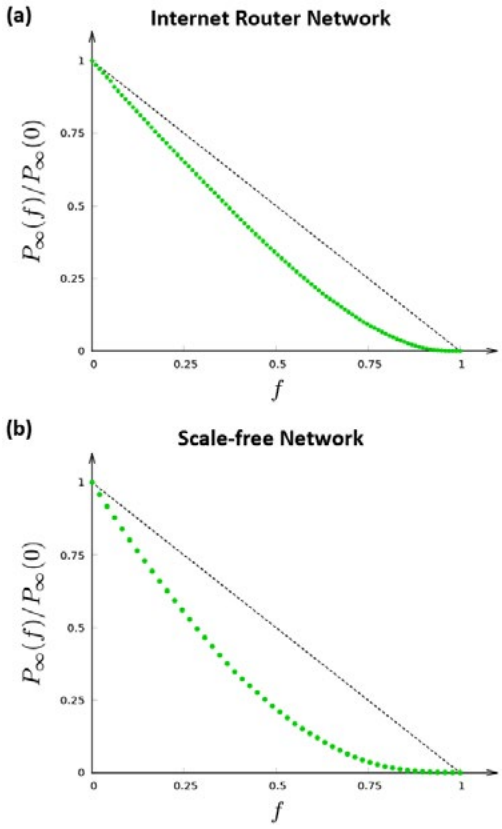



**Figure 8.7**
**Robustness of scale-free networks**

**(a)** The fraction of Internet routers that belong to the giant component after an $f$ fraction of routers are randomly removed. The ratio $P_\infty(f)/P_\infty(0)$ provides the relative size of the giant component. The simulations use the router level Internet topology of Table 4.1.

**(b)** The fraction of nodes that belong to the giant component after the removal of an $f$ fraction of nodes from a scalefree network with $\gamma = 2.5$, $N = 10,000$ and $k_{min} = 1$.

The plots indicate that the Internet and in general a scale-free network do not fall apart after the removal of a finite fraction of nodes. We need to remove almost all nodes (i.e. $f_c \simeq 1$) to fragment these networks.

words, the random removal of a finite fraction of its nodes does not break apart a large scale-free network. To further illustrate the roots of this anomaly we express $\langle k \rangle$ and $\langle k^2 \rangle$ in terms of the parameters characterizing a scale-free network: the degree exponent $\gamma$ and the minimal and maximal cutoffs, $k_{min}$ and $k_{max}$, obtaining (ADVANCED TOPICS 5.D)

$$f_c = \begin{cases} 1 - \dfrac{1}{\dfrac{\gamma-2}{3-\gamma} k_{\min}^{\gamma-2} k_{\max}^{3-\gamma} - 1} & 2 < \gamma < 3 \\ 1 - \dfrac{1}{\dfrac{\gamma-2}{\gamma-3} k_{\min} - 1} & \gamma > 3 \end{cases} \tag{8.9}$$

Eq. 8.9 predicts that Fig. 8.9:

- For $\gamma > 3$ the critical threshold $f_c$ depends only on $\gamma$ and $k_{min}$, hence $f_c$ is independent of the network size $N$. In this regime a scale-free network behaves like a random network: it falls apart after the removal of a finite fraction of its nodes.

- For $\gamma \leq 3$ the $k_{max}$ diverges for large $N$ (see Eq. 4.18). Therefore in the $N \to \infty$ limit Eq. 8.9 predicts $f_c \to 1$. Hence, to fragment an infinite scale-free network *we must remove all of its nodes*.

Eq. 8.6, 8.9 are the key results of this chapter, predicting that scale-free networks can withstand an arbitrary level of random failures without breaking apart. To understand the origin of this remarkable robustness we must inspect the role of the hubs. Random node failures by definition are blind to degree, affecting with the same probability a small or a large degree node. Yet, in a scale-free network we have far more small degree nodes than hubs. Therefore, random node removal will predominantly remove one of the numerous small nodes as the chances of removing one of the few large hubs is negligible. These small nodes contribute little to a network's integrity, hence their removal does not damage the network.

Returning to the airport analogy of Fig. 4.6, if we close a randomly selected airport, we will most likely be shutting down one of the numerous small airports. Its absence will be hardly noticed elsewhere in the world: you can still travel from New York to Tokyo, or from Los Angeles to Rio de Janeiro.

### ROBUSTNESS OF FINITE NETWORKS

Eq. 8.9 predicts that for a scale-free network $f_c$ converges to one only in the $k_{max} \to \infty$ (or $N \to \infty$) limit. While many networks of practical interest are very large, they are still finite, prompting us to ask if the observed anomaly is relevant for finite systems. We can address this by inserting Eq. 4.18 into Eq. 8.9, obtaining that $f_c$ depends on the network size $N$ as

**Figure 8.8**
**Malloy-Reed criteria**

To form a chain, each individual must hold the hand of two other individuals. Similarly, to have a giant component in a network, on average each of its nodes should have at least two neighbors. The Malloy-Reed criteria Fig. 8.4 exploits this property to help us calculate the critical point at which a network breaks apart.

$$f_c \simeq 1 - \frac{C}{N^{\frac{3-\gamma}{\gamma-1}}}. \qquad (8.10)$$

where $C$ collects all terms that do not depend on $N$. Eq. 8.10 indicates that the larger the size of a network, the closer will be its critical threshold to $f_c = 1$. To see how close fc can get in a real system to the theoretical $f_c = 1$, we calculate $f_c$ for the Internet. The router level map of the Internet has $\langle k^2 \rangle / \langle k \rangle = 37.94$ Table 4.1. Inserting this ratio into Eq. 8.7 we obtain $f_c = 0.972$. Therefore, we need to remove 97% of the routers to fragment the Internet into disconnected components. The probability that by chance 220,000 routers fail simultaneously, representing 97% of the $N = 228,263$ routers on the Internet, is effectively zero. This is one of the reasons the topology of the Internet is so robust to random failures.

A network displays enhanced robustness if its breakdown threshold deviates from the random network prediction Eq. 8.8, i.e. if

$$f_c > f_c^{ER}. \qquad (8.11)$$

Enhanced robustness has several ramifications:

- The inequality Eq. 8.11 is satisfied for all networks for which $\langle k^2 \rangle$ deviates from $\langle k \rangle$ ($\langle k \rangle + 1$). According to Fig. 4.8, for virtually all networks in our reference network list $\langle k^2 \rangle$ exceeds the random expectation. Hence the robustness predicted by Eq. 8.7 is not an isolated property of a few selected networks, but affects most networks of practical interest.

- Eq. 8.7 also predicts that the degree distribution of a network does not need to follow a strict power law to display enhanced robustness; all we need is a larger $\langle k^2 \rangle$ than expected for a random network of similar size.

- Finally, enhanced robustness is not limited to node removal, but emerges under link removal as well Fig. 8.10.

In summary, in this section we encountered a fundamental property of real networks: their robustness to random failures. Eq. 8.7 predicts that the breakdown threshold of a network depends only on its degree distribution through $\langle k \rangle$ and $\langle k^2 \rangle$. This predicts that random networks have a finite threshold, but for scale-free networks with $\gamma < 3$ the breakdown threshold converges to one. Therefore, we need to remove all nodes to break a scale-free network apart, indicating that these networks show enhanced robustness to random failures. The origin of this enhanced robustness is the large $\langle k^2 \rangle$ term. As for most real networks $\langle k^2 \rangle$ is larger than the random expectation, enhanced robustness is a generic property of many networks. This robustness is rooted in the fact that random failures affect mainly the numerous small nodes, which play a limited role in maintaing a network's integrity.

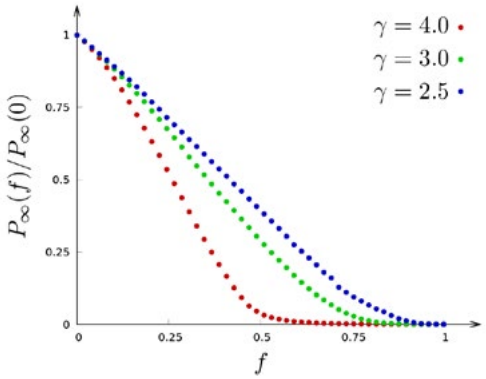



**Figure 8.9**
**Robustness and degree exponent**

The probability that a node belongs to the giant component after the removal of an $f$ fraction of nodes from scale-free networks with different degree exponent $\gamma$. For $\gamma = 4$ we observe a finite critical point $f_c$, as predicted by Eq. 8.9. For $\gamma < 3$, however, we have $f_c \to 1$. The networks were generated with the configuration model using $k_{min} = 2$ and $N = 10,000$.

| NETWORK | RANDOM FAILURES | RANDOM NETWORK | ATTACK |
|---|---|---|---|
| Internet | 0.92 | 0.84 | 0.16 |
| WWW | 0.88 | 0.85 | 0.12 |
| Power Grid | 0.61 | 0.63 | 0.20 |
| Mobile-Phone Call | 0.78 | 0.68 | 0.20 |
| Email | 0.92 | 0.69 | 0.04 |
| Science Collaboration | 0.92 | 0.88 | 0.27 |
| Actor Network | x | 0.99 | 0.55 |
| Citation Network | 0.96 | 0.95 | 0.76 |
| E. Coli Metabolism | 0.96 | 0.90 | 0.49 |
| Yeast Protein Interactions | 0.88 | 0.66 | 0.06 |

**Table 8.1**
**Breakdown thresholds under random failures and attacks**

The table shows the estimated fc for random failures (second column) and attacks (fourth column) for ten reference networks. The procedure for determining fc is described in ADVANCED TOPICS 10.X. The third column (random network) offers fc for a network whose $N$, $L$ coincides with the original values, but whose nodes are connected randomly to each other. Note that for most networks $f_c$ for random failures exceeds $f^{ER}$ for the corresponding random network, indicating that these networks display enhanced robustness, as defined in **Eq. 8.11**. Only the power grid lacks this property, a consequence of the fact that its degree distribution is exponential **Fig. 8.27e**.

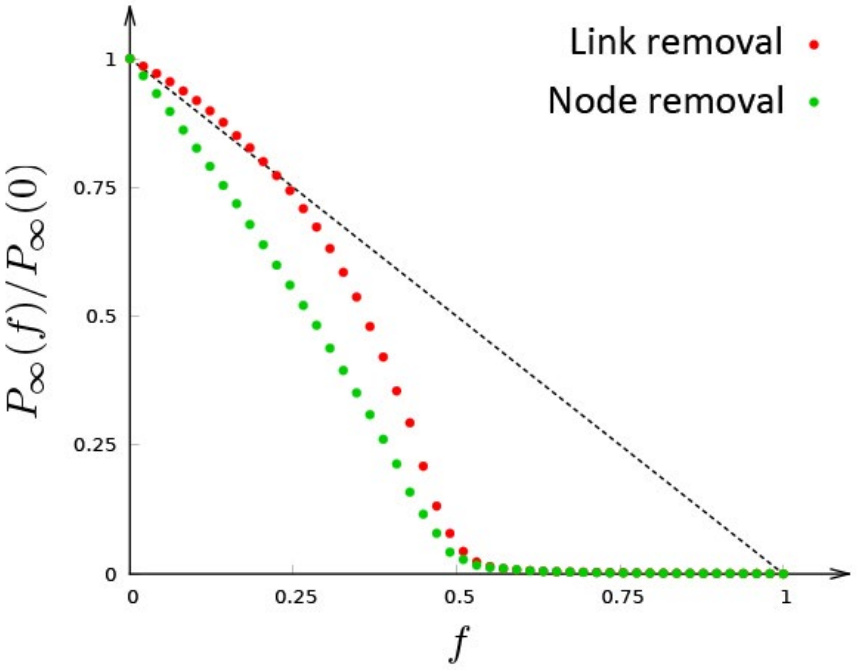



**Figure 8.10**
**Robustness and link remova**

Our focus on node removal prompts us to ask: what happens if we randomly remove the links rather than the nodes? That is, how robust are networks to link removal? The calculations predict that the critical threshold fc is the same for random link and node removal [7, 8]. To illustrate this,we compare the impact of random node and link removal on a random network with $\langle k \rangle = 2$. The plot indicates that the network falls apart at the same critical threshold $f_c \simeq 0.5$. The difference in the shape of the two curves is rooted in the fact that the removal of an f fraction of nodes leaves us with a smaller giant component than the removal of an f fraction of links. This is not unexpected: on average each node removes $\langle k \rangle$ links, hence the removal of an $f$ fraction of nodes is equivalent with removing an $f\langle k \rangle$ fraction of the links, which clearly makes more damage than the removal of an $f$ fraction of links.

# ATTACK TOLERANCE

The important role the hubs play in holding together a scale-free network prompts our next question: what if we do not remove the nodes randomly, but go after the hubs? That is, we first remove the highest degree node, followed by the node with the next highest degree and so on. The likelihood that nodes would break in this particular order under normal conditions is essentially zero. Instead this process mimics an attack on the network, as it assumes a detailed knowledge of the network topology, an ability to target the hubs, and a desire to deliberately cripple the network [1]. The removal of a single hub is unlikely to fragment a network, as the remaining hubs can still hold the network together. After the removal of a few hubs, however, large chunks of nodes start falling off Movie 8.2. If the attack continues, it can rapidly break the network into tiny clusters.

The impact of hub removal is quite obvious in the case of the scale-free network shown in Fig. 8.11: the critical point, which is absent under random failures (green curve), reemerges under attacks (red curve). Not only reemerges, but it has a remarkably low value. This indicates that the removal of a small fraction of the nodes, namely the system's hubs, is sufficient to break a scale-free mnetwork into tiny clusters. The goal of this section is to identify the origin of this attack vulnerability and to quantify its magnitude.

### CRITICAL THRESHOLD UNDER ATTACKS

As Fig. 8.11 indicates, an attack on a scale-free network has two consequences:

- The critical threshold, $f_c$, is smaller than $f_c = 1$, indicating that under attacks a scale-free network can be fragmented by the removal of a finite fraction of its hubs.

- The observed $f_c$ is remarkably low, indicating that we need to remove only a tiny fraction of the hubs to cripple the network.

To quantify this process we need to analytically calculate $f_c$ for a net-

## MOVIE 8.2

**SCALE-FREE NETWORK UNDER ATTACK**

During an attack we aim to inflict maximum damage on a network. We can do this by removing first the highest degree node, followed by the next highest degree, and so on. As the movie illustrates, it is sufficient to remove only a few hubs to break a scale-free network into disconnected components. Compare this with the network's refusal to break apart under random node failures, shown in MOVIE 8.1. Visualization by Dashun Wang.

work under attack. To do this we rely on the fact that hub removal changes the underlying network in two different ways [9]:

- It changes the maximum degree of the network from $k_{max}$ to $k'_{max}$ as all nodes with degree larger than $k'_{max}$ have been removed.

- The degree distribution of the network changes from $p_k$ to $p'_k$, as all nodes connected to the removed hubs will loose links, altering the degrees of the remaining nodes.

In ADVANCED TOPICS 8.E we combine these two changes and map the attack problem into the robustness problem discussed in the previous section. In other words, we can view an attack as random node removal from a network with adjusted $k'_{max}$ and $p'_k$. The calculations predict that the critical threshold $f_c$ for attacks on a scale-free network with degree exponent $\gamma$ is the solution of the equation [9, 10].

$$f_c^{\frac{2-\gamma}{1-\gamma}} = 2 + \frac{2-\gamma}{3-\gamma} k_{min} (f_c^{\frac{3-\gamma}{1-\gamma}} - 1). \qquad (8.12)$$

Fig. 8.12 shows the numerical solution of Eq. 8.12 in function of $\gamma$, leading to several conclusions:

- While $f_c$ for failures decreases monotonically with $\gamma$, $f_c$ for attacks has a complex non-monotonic behavior.

- $f_c$ for attacks is always smaller than $f_c$ for random failures.

- For large $\gamma$ a scale-free network behaves like a random network. As a random network lacks hubs, an attack on a random network will follow a scenario similar to random node removal. Numerical simulations support this expectation: Fig. 8.13 shows that a random network has a finite percolation threshold under both random failures and attack. The main difference is that $f_c$ for attacks is lower than $f_c$ for random failures.

- The failure and the attack thresholds converge to each other for large $\gamma$. Indeed, if $\gamma \to \infty$ then $p_k \to \delta(k - k_{min})$, meaning that all nodes have the same degree $k_{min}$. Therefore random failures and targeted attacks become indistinguishable in the $\gamma \to \infty$ limit, when $f_c \to 1 - 1/(k_{min} - 1)$.

The airport analogy helps us understand the fragility of scale-free networks to attacks: the closing of two hub airports, like Chicago's O'Hare Airport or the Atlanta International Airport for only a few hours would be headline news, altering travel throughout the US. Should some series of events lead to the simultaneous closure of the Atlanta, Chicago, Denver, and New York airports, the biggest hubs, air travel within the U.S. would come to a halt within hours.

In summary, while random node removal has difficulty fragmenting

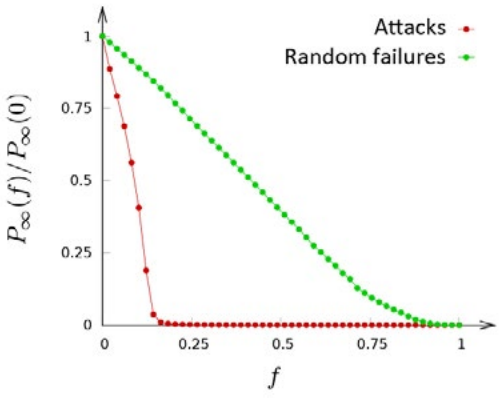



**Figure 8.11**
**Scale-free networks under attack**

The probability that a node belongs to the largest connected component in a scale-free network under attack (red) and under random failures (green). In the case of an attack the nodes are removed in a decreasing order of their degree: we first remove the biggest hub, followed by the next biggest and so on. In the case of failures, the order in which the nodes are chosen is random, independent of the node's degree. The plot illustrates the network's extreme fragility to attacks: $f_c$ is rather small, implying that the removal of only a few hubs can disintegrate the network. The initial network has a degree exponent $\gamma = 2.5$, $k_{min} = 2$ and $N = 10,000$.

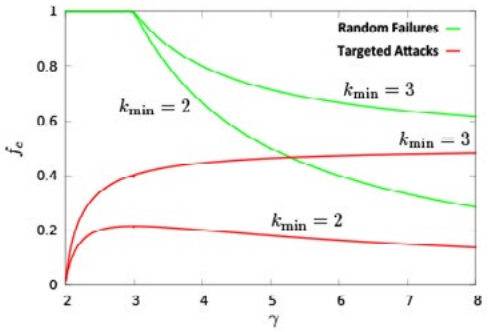



**Figure 8.12**
**Critical threshold under attack**

The dependence of the critical probability, $f_c$, on the degree exponent $\gamma$, for scale-free networks with $k_{min} = 2, 3$, as predicted by Eq. 8.12, for an attack (red curves) and by Eq. 8.7 for random failures (green curves). Note that Eq. 8.12 predicts that the attack threshold $f_c \to 0$ for $k_{min} = 2$ and $f_c \to 1/2$ for $k_{min} = 3$, in line with the behavior observed in the figure.

a scale-free network, an attack that targets the hubs can easily destroy a network. This fragility is bad news for the Internet, as it indicates that it is inherently vulnerable to deliberate attacks. It can be good news in medicine, as the vulnerability of bacteria to the removal of its hub proteins offers avenues to design drugs that target these hubs, potentially destroying the organism.

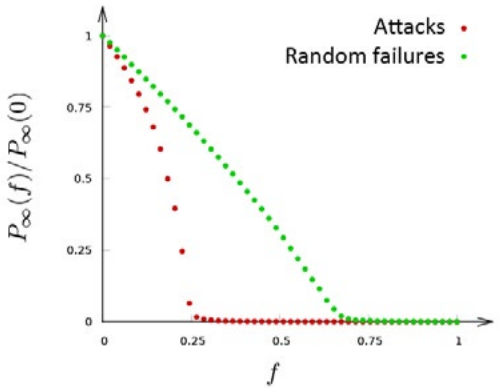



**Figure 8.13**
**Attack and failures in random networks**

The fraction of nodes that belong to the giant component in a random (i.e. Erdős-Rényi) network if an $f$ fraction of nodes are removed randomly (random failure, green) and in decreasing order of their degree (attacks, red). Both curves indicate the existence of a finite threshold, in contrast with scale-free networks, for which $f_c \to 1$ under random failures. The simulations were performed for random networks with $N = 10,000$ and $\langle k \rangle = 3$.

# BOX 8.2

## PAUL BARAN AND THE INTERNET

In 1959 RAND, a Californian think-tank, has assigned Paul Baran, a young engineer at that time, to develop a communication system that can survive a Soviet nuclear attack. As a nuclear strike handicaps all equipment within the range of the detonation, Baran had to design a system whose users outside this range would not lose contact with one another. He described the communication system of his time as a "hierarchical structure of a set of stars connected in the form of a larger star," offering an early description of what we would call today a scale-free network. He concluded that this topology is too centralized to be viable under attack. He also discarded the hub-and-spoke topology shown in Fig. 8.14a noting that "The centralized network is obviously vulnerable as destruction of a single central node destroys communication between the end stations." Baran decided that the ideal survivable architecture was a distributed mesh-like network, shown in Fig. 8.14C, which is sufficiently redundant, so that even if some of its nodes break down, alternative paths can maintain the connection between the remaining nodes. Baran's ideas were ignored by the military, so when the Internet was born a decade later, it relied on distributed protocols that allowed each node to decide where to link. This decentralized approach paved the way to the emergence of a scale-free Internet, rather than the uniform mesh-like topology envisioned by Baran. Consequently the Internet today resembles more the decentralized structure B, the one that Baran wanted to avoid, than the distributed topology C he preferred.

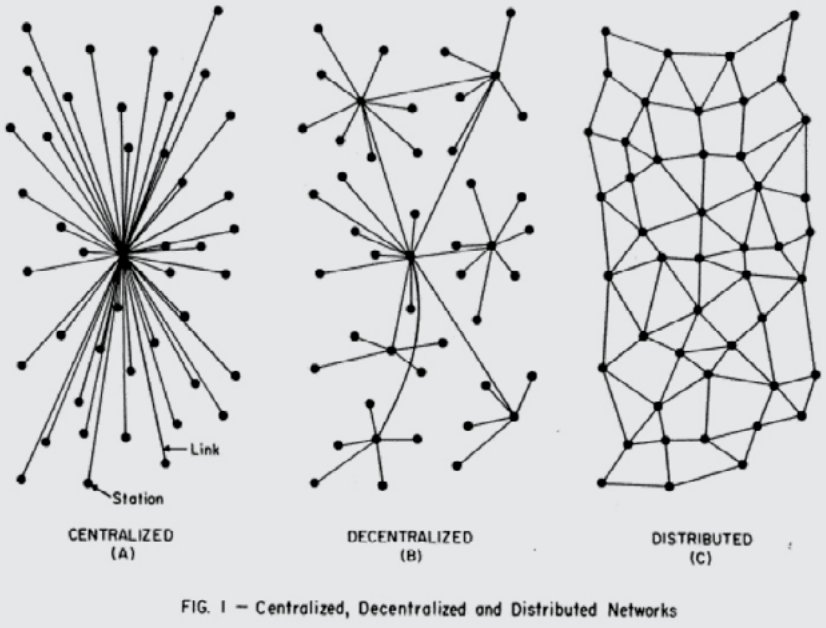



FIG. 1 — Centralized, Decentralized and Distributed Networks

**Figure 8.14**
**Baran's Network**

Possible network configurations, described by Paul Baran in his 1959 report.

# CASCADING FAILURES

Throughout this chapter we assumed that the nodes in a network fail independently of each other. In reality, the activity of each node in a network depends on the activity of its neighboring nodes. Consequently the failure of one node can often induce the failure of the nodes connected to it. Let us consider a few examples:

- **Blackouts (power grid)**
  As electricity travels with the speed of light, after a node or link failure the electric currents are instantaneously reorganized on the rest of the power grid. For example, on August 10, 1996, a hot day in Oregon, a line carrying 1,300 megawatts sagged close to a tree and snapped. Because electricity cannot be stored, the current it carried was automatically shifted to two lower voltage lines with 115 and 230 kilovolt capacity. These were not designed to carry the excess current, so they also failed. Seconds later the excess current lead to the malfunction of thirteen generators, causing a blackout in eleven U.S. states and two Canadian provinces [11].

- **Denial of service attacks (Internet)**
  If a router malfunctions, responding too slowly or failing to transmit the packets received by it, the Internet protocols will alert the neighboring routers, which will re-route the packets, using alternative routes to avoid the troubled equipment. Consequently a failed router can place a significant burden on other routers, potentially inducing cascading failures in the form of a series of denial of service attacks distributed throughout the Internet [12].

- **Financial Crises**
  Cascading failures are common in economic systems as well. For example, the drop in the house prices in 2008 in the U.S. lead to a global financial meltdown that is considered the worst crisis since the 1930s Great Depression. In other words, the impact of the housing bubble has spread along the links of the financial network, inducing a cascade of failures throughout the economy, leading to failed banks,

**Figure 8.15**
**Domino effect**

In general we call the domino effect a sequence of events that is induced by a local change, yet it propagates through the whole system. The phenomena is similar to the fall of a series of dominos induced by the fall of the first domino. The domino effect represents perhaps the simplest illustration of cascading failures, the topic of this section.

# BOX 8.3

**NORTHEAST BLACKOUT OF 2003**

One of the largest blackouts in North America took place on August 14, 2003, just before 4:10 p.m. Its cause was a software bug in the alarm system at a control room of the *FirstEnergy* Corporation in Ohio. Missing the alarm, the operators were unaware of the need to redistribute the power after an overloaded transmission line hit a tree. Consequently a normally manageable local failure stared a cascading failure that shut down more than 508 generating units at 265 power plants, leaving without electricity an estimated 10 million people in Ontario and 45 million people in eight U.S. states.

**Figure 8.16**
**The 2003 Power Outage**

Canadian and USA states affected by the August 14, 2003 blackout, illustrating how a local failure can turn into a major global event.

companies and even nations [13, 14, 15].

Despite covering different domains, these examples have several common characteristics. First, the initial failure had only limited impact on the network structure. Second, the initial failure did not stay localized, but it spread along the links of the network, inducing additional failures. Eventually, multiple nodes, one after the other, failed to carry out their normal functions. Each of these systems experienced cascading failures, a dangerous phenomena in many networks [16]. In this section we discuss the empirical patterns governing such cascading failures. The modeling of these events in the topic of the next section.

### EMPIRICAL RESULTS

Cascading failures are well documented in the case of the power grid, information systems and tectonic motion, offering detailed statistics about their frequency and magnitude.

- **Blackouts**

  A blackout, also called a power outage or power failure, is a loss of the electric power in some area. It can be caused by failures at power stations, damage to electric transmission lines, substations, a short circuit, and so on. When the operating limits of a component is exceeded, it is typically automatically disconnected to protect it. In other words, a component can "fail" in the sense that it is not available to transmit power. Such failure redistributes the power previously carried by the failed component to other components, altering power flows, frequency, voltage and phase, and inducing changes in the operation of the control, monitoring and alarm systems. These changes can in turn disconnect other components as well, in some cases starting an avalanche of failures.

  A frequently recorded measure of blackout size is the energy unserved. Fig. 8.17a shows the probability distribution $p(s)$ of energy unserved in all North American blackouts between 1984 and 1998. Electrical engineers approximate the obtained distribution with a power law [17],

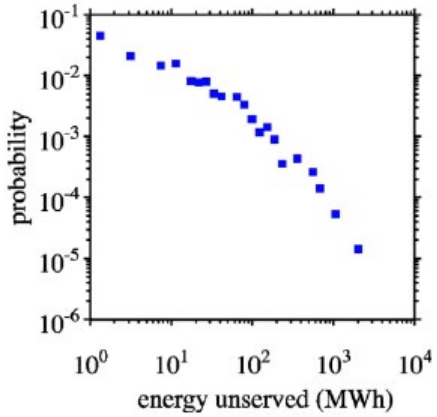
$$p(s) \sim s^{-\alpha}, \tag{8.13}$$

  where the estimated value of the avalanche exponent $\alpha$ is listed in Table 8. 2 for several countries. The power law nature of this distribution indicates that most blackouts are rather small, affecting only a few consumers. These coexists, however, with occasional major blackouts, where millions of consumers lose power BOX 8.3.

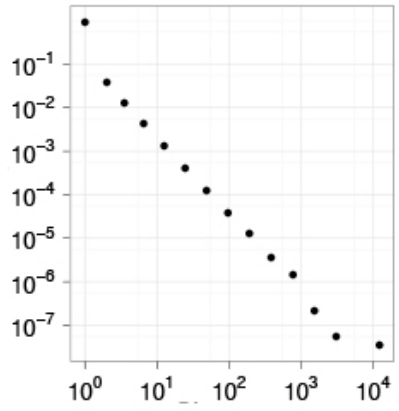
- **Information cascades**

  Modern communication systems, from email to mobile phones, Facebook or Twitter, allow for the cascade-like spreading of information along the links of the social network. As the events pertaining to the spreading process often leave digital traces, these platforms al-

a




b



c

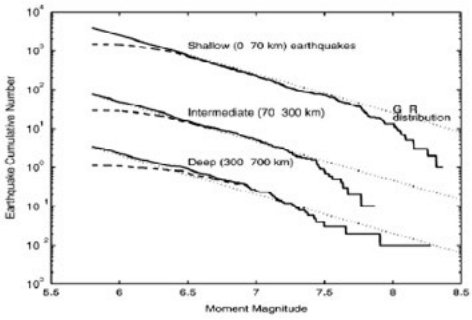



**Figure 8.17**
**Cascade size distributions**

**(a)** The distribution of energy loss for all North American blackouts between 1984 and 1998, as documented by the North American Electrical Reliability Council. The distribution is typically fitted to **Eq. 8.13**. The reported exponents for different countries are listed in **Table 8.2**. After [17].

**(b)** The distribution of cascade sizes on Twitter. While most tweets cause no reaction, bringing the average cascade size down to 1.14, a tiny fraction of tweets are shared thousands of times. Overall the retweet numbers are well approximated with **Eq. 8.13** with $\alpha \simeq 1.75$. After [18].

**(c)** The distribution of earthquake sizes recorded between 1977 and 2000. The dotted line indicates the power law fit **Eq. 8.13** used by seismologists to characterize the distribution. After [19].

low researchers to detect and explore the underlying cascades. The micro-blogging service Twitter has been particularly useful in this context. On Twitter the network of who follows whom can be reconstructed by crawling the follower graph behind the service. As users frequently share web-content using URL shorteners, the tracking of a spreading/sharing process is relatively straightforward. A study tracking 74 million such events over a two month interval in 2009 traced the diffusion of each URL from a particular seed node through its reposts until the end of a cascade Fig. 8.18.

As Fig. 8.17b indicates, the cascade size distribution follows the power-law Eq. 8.13 with an avalanche exponent $\alpha \simeq 1.75$. This indicates that the vast majority of posted URLs do not spread at all, a claim also supported by the fact that the average cascade size is only 1.14. Yet, a small fraction of URLs are reposted thousands of times.

- **Earthquakes**

  Most geological fault surfaces are irregular and sticky, not permitting a smooth slide against each other. Once a fault has locked, the continued relative motion of the plates increases the stress, accumulating an increasing amount if strain energy around the fault surface. This continues to accumulate until the stress is sufficient to break through the asperity, resulting in a sudden slide that releases the stored energy and causes an earthquake. Earthquakes can be also induced by the natural rupture of geological faults, by volcanic activity, landslides, mine blasts and even nuclear tests. Each year around 500,000 earthquakes are detected with instrumentation. Only about 100,000 of these are sufficiently strong to be felt by humans. Seismologists approximate the distribution of earthquake sizes with the power law Eq. 8.13 with $\alpha \approx 1.67$ Fig. 8.17c. Earthquakes are rarely considered a manifestly network phenomenon, given the difficulty of mapping out the precise interdependencies in the Earth's crust that causes them. Yet, the resulting cascading failures bear many similarities to network based cascading events, suggesting common mechanisms.

The power-law distribution Eq. 8.13 followed by blackouts, information cascades and earthquakes indicates that most cascading failures are relatively small.

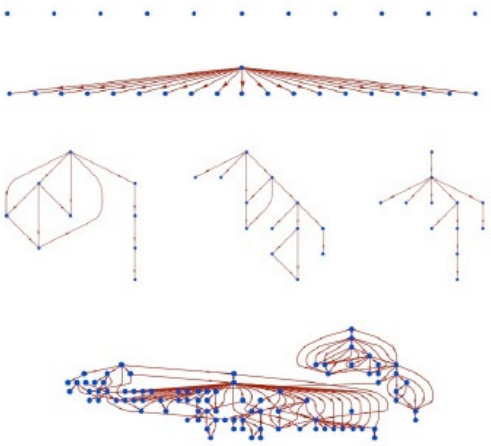


**Figure 8.18**
**Information Cascades**

Examples of information cascades on Twitter. Nodes denote Twitter users, the top node corresponding to the individual who first posted a certain shortened URL. The links correspond to those who retweeted it. The observed cascades capture the heterogeneity of information avalanches: most URLs are not retweeted at all (shown as individual nodes in the figure), but some are part of major retweet avalanches, like the one seen at the bottom panel. After [18].

| SOURCE | EXPONENT | S |
|---|---|---|
| Power grid (North America) | 2.0 | Power |
| Power grid (Sweden) | 1.6 | Energy |
| Power grid (Norway) | 1.7 | Power |
| Power grid (New Zealand) | 1.6 | Energy |
| Power grid (China) | 1.8 | Energy |
| Twitter Cascades | 1.75 | Retweets |
| Earthquakes | 1.67 | Energy |

**Table 8.2**
**Avalanche exponents in real systems.**

The reported avalanche exponents characterizing the power law distribution Eq. 8.13 of energy loss in various countries [17], twitter cascades [18] and earthquake sizes [19]. The third column indicates the nature of the measured cascade size s, corresponding to power or energy not served, the number of retweets generated by a typical tweet and earthquake magnitudes.

These capture the loss of electricity in a few houses, tweets of little interest to most users, or earthquakes so small that are detected only by sensitive instruments. Eq. 8.13 also predicts that these numerous small events coexist with a few exceptionally large events, like blackouts leaving millions without power, messages retweeted by hundreds or earthquakes causing major loss of human life. Examples of such major cascades include the 2003 power outage in North America BOX 8.3, the tweet *Iran Election Crisis: 10 Incredible YouTube Videos http://bit.ly/vPDLo* that was shared 1,399 times [20], or the January 2010 earthquake in Haiti, with over 200,000 victims Fig. 8.19. What is particularly intriguing as we compare these systems is that the avalanche exponents reported by electrical engineers, media researches and seismologists are so close to each other: they are all between 1.6 and 2 Table 8.2.

Cascading failures are documented in many other environments:

- The consequences of bad weather or mechanical failures can cascade through airline schedules, delaying flights whose schedule is apparently unrelated to the original cause, in some cases stranding thousands of passengers BOX 8.5 [21].

- The extinction of a species can cascade through an ecosystem, inducing the extinction of numerous species and altering the habitat of others [22, 23, 24, 25].

- Cascading failures are common in international trade, when the shortage of a particular component cripples supply chains, affecting the availability of a numerous products. For example, the 2011 floods in Thailand have resulted in a chronic shortage of car components that disrupted the production chain of more than 1,000 automotive factories. Thanks to these cascading events the damage was not limited to the flooded factories, resulting in insurance claims reaching $20 billion [26].

In summary, cascading effects are observed in systems of rather different nature. Their size distribution is well approximated with a power law, implying that most cascades are too small to be noticed; a few, however, are huge, having global impact. The goal of the next section is to understand the origin of these phenomena.

**Figure 8.19**
**Earthquake Damage**

Devastation caused in Port-au-Prince, Haiti, by a magnitude 7 earthquake that hit the island on January 12, 2010. After *http://en.wikipedia.org/wiki/File:Haiti_earthquake_damage.jpg*

# BOX 8.4

## CASCADING CONGESTION

In the U.S. flight delays have an economic impact of over $40 billions per year [27], caused by the need for enhanced operations, passenger loss of time, decreased productivity and missed business and leisure opportunities. A flight delay is defined as the time difference between the expected and actual departure/arrival times of a flight. Airline schedules include a buffer period between consecutive flights to allow for potential delays. When a delay exceeds this buffer, subsequent flights that rely on the same aircraft, flight crew or gate, are also delayed. Consequently the impact of a delay can propagate in a cascade-like fashion through the system.

A study found that while most flights in 2010 were on time, 37.5% arrived or departed late [21]. The delay distribution has a broad tail, similar to Eq. 8.13, implying that while most flights were delayed by just a few minutes, a few were hours behind schedule. These long delays induce correlated delay patterns, a signature of cascading congestions in the air transportation system Fig. 8.20.

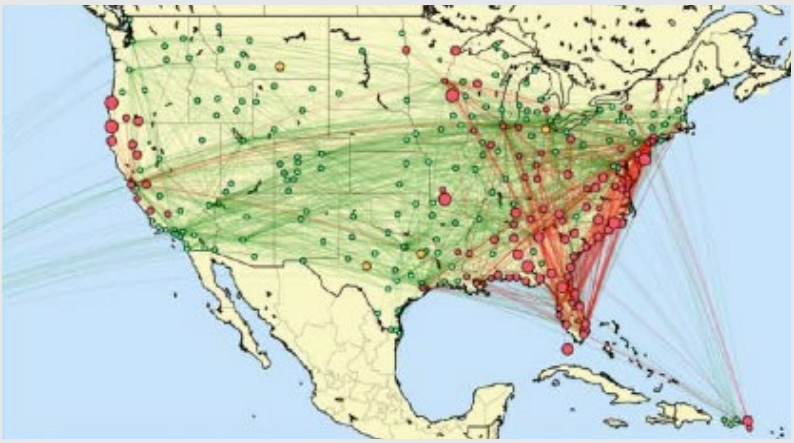


**Figure 8.20**
**Clusters of congested airports**

A U.S. aviation map showing the congested airports as red nodes, while those with normal traffic as green nodes. The lines correspond to the direct flights between them on March 12, 2010. Congested airports form a correlated cluster, a manifestation of cascading delays in air travel. After [21].

# MODELING CASCADING FAILURES

The unfolding of a cascading event depends on the structure of the network on which it propagates, the nature of the propagation process, and the breakdown criteria of each individual component. The empirical results discussed in the previous section indicate that despite the diversity of these factors, the statistics of cascading processes is universal, being independent of the particularities of the system. The purpose of this section is to understand the mechanisms governing cascading phenomena and to explain the power-law nature of the observed cascade size distributions.

Numerous models have been proposed to capture the dynamics of cascading events [17, 28, 29, 30, 31, 32, 33, 34]. While these models differ in the degree of fidelity they employ to capture specific phenomena, they indicate that systems that develop cascades share three key ingredients:

**(i)** Each system is characterized by some flow over a network, like the flow of electric current in the power grid or the transport of information in communication systems.

**(ii)** Each component has a local breakdown rule that determines when it contributes to a cascade, either by failing or choosing to pass on a piece of information.

**(iii)** Each system has a mechanism to redistribute the traffic or flow to other nodes upon the failure or the activation of a component.

Next, we discuss two models that offer an increasing level of abstraction and with that an increased ability to predict the characteristics of cascading failures.

### FAILURE PROPAGATION MODEL

Introduced to model the spread of ideas and opinions, the failure propagation model [29] is used to describe both information cascades and cascading failures [34]. Consider a network with an arbitrary degree distribution $p_k$, where each node contains an agent. Each agent $i$ can be in the

state *0* (active or healthy) or 1 (inactive or failed), and is characterized by a breakdown threshold $\varphi_i \equiv \varphi$.

All agents are initially in state 0. At time $t = 0$ one agent is switched to state 1, capturing for example an initial failure or the birth of new information. In each subsequent time step, we randomly pick an agent and update its state following a simple threshold rule:

- If the selected agent *i* is in state 0, it inspects the state of its $k_i$ neighbors. The agent *i* adopts state *1* (i.e. it also fails) if at least a $\varphi$ fraction of its $k_i$ neighbors are in state *1*, otherwise it retains its original state 0.

- If the selected agent *i* is in state *1*, it does not change its state.

Depending on the local network topology, an initial perturbation can die out immediately, failing to induce the failure of any other node. It can also lead to the failure of additional nodes, as illustrated in Fig. 8.21a, b. To characterize the dynamics of the model we focus on two quantities:

**(i)** The probability that a global cascade is triggered by a single node, a global cascade is defined as a sequence of failures that involves a finite fraction of all nodes Fig. 8.21c.

**(ii)** The expected size distribution of the observed cascades Fig. 8.21d.

Both quantities depend on the average degree $\langle k \rangle$ of the network and the threshold $\varphi$. The simulations document three regimes with distinct avalanche characteristics:

- **Subcritical Regime**
  If $\langle k \rangle$ is high, changing the state of a node is unlikely to move other nodes over their threshold, as the unflipped nodes have many neighbors that did not yet flip. In this regime cascades die out quickly and their sizes follow an exponential distribution. The system is subcritical, unable to support large global cascades (blue symbols in Fig. 8.21c, d).

- **Supercritical Regime**
  If $\langle k \rangle$ is very small, the flipping of a single node can put several of its neighbors over the threshold, triggering a global cascade. In this case virtually any perturbation induces a major breakdown, making the system supercritical (black symbols in Fig. 8.21c, d).

- **Critical Regime**
  At the boundary of the subcritical and supercritical regime the avalanches have widely different sizes. Numerical simulations indicate that in this regime the avalanche sizes s follow Eq. 8.13 (green, red symbols in Fig. 8.21d), with $\alpha = 3/2$ if the underlying network is random.

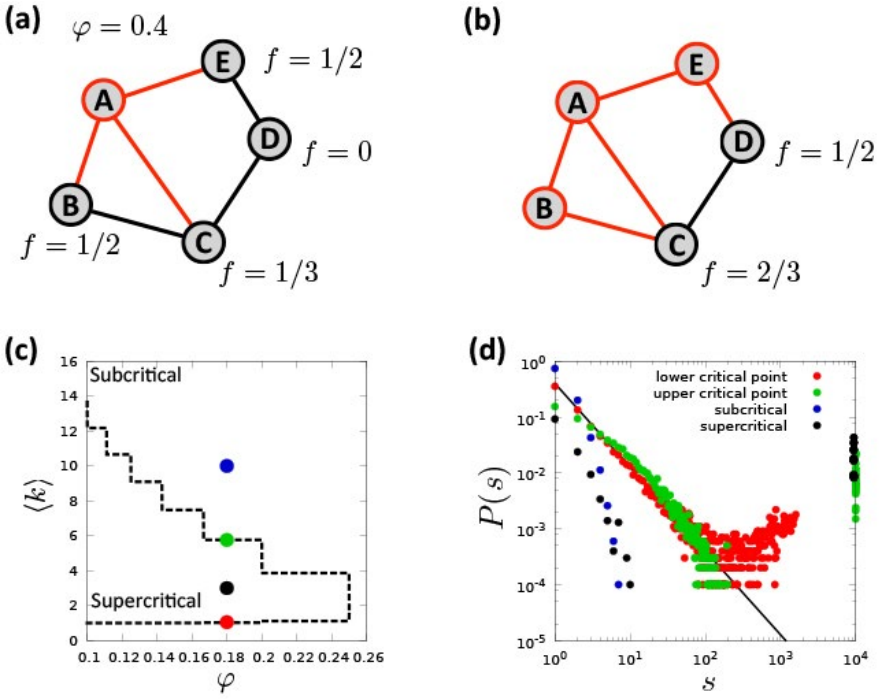



**Figure 8.21**
**Failure propagation model**

**(a)** The emergence of a cascade in a small network when each node has the same breakdown threshold $\varphi = 0.4$. Initially all nodes are in state 0, shown as black circles. After node A changes its state to 1 (red circle), its neighbors *B* and *E* will have a fraction $f = 1/2 > 0.4$ of their neighbors in state *1*. Consequently they also fail, changing their state to 1, as shown in (b).

**(b)** In the next time step *C* and *D* will also fail, having $f > 0.4$. Consequently at the end the cascade sweeps the whole network, reaching a size $s = 5$. One can check that if we initially flip node *B*, it will not induce an avalanche, as none of its neighbors pass the threshold $\varphi$.

**(c)** The phase diagram of the failure propagation model in terms of the threshold function $\varphi$ and the network's average degree $\langle k \rangle$. The dashed line encloses the region of the $(\langle k \rangle, \varphi)$ plane in which the cascades can propagate in a random graph.

**(d)** Cascade size distributions for $N = 10{,}000$ and $\varphi = 0.18$, $\langle k \rangle = 1.05$ (red), $\langle k \rangle = 3.0$ (black), $\langle k \rangle = 5.76$ (green) and $\langle k \rangle = 10.0$ (blue). At the lower critical point we observe a power law $p(s)$ with exponent $\alpha = 3/2$ . In the supercritical regime we have only a few small avalanches, as most cascades are global. In the upper critical and subcritical regime we see only small avalanches. After [29].

In ADVANCED TOPICS 8.F we discuss two additional models describing cascading failures:

- The overload model is designed to capture power grid failures. Its distinguishing feature is a global flow process: as the electric current redistributes itself throughout the power grid, the model assumes that each failure instantaneously increases the load of all nodes. This is different from the local spread of failures in the failure propagation model.

- The self-organized critical model aims to model only the behavior of a system in the critical regime.

Despite there differences these two models predict the same avalanche exponent ($\alpha = 3/2$ for a random network) in the critical regime as the failure propagation model. The fact that the three models with rather different propagation dynamics and failure mechanisms predict similar scaling laws and avalanche exponents suggests that the underlying phenomena is universal, i.e. model independent.

### BRANCHING MODEL

Given the complexity of the models discussed above, it is hard to analytically predict their scaling behavior. To understand the origin of the power-law nature of the observed $p(s)$ and to calculate the avalanche exponent $\alpha$, we turn to the branching model. This is perhaps the simplest model that still captures the basic features of cascading events.

The model builds on the observation that the history of a cascading failure (avalanche) can be described as a branching process. Let us designate the node whose failure triggers the avalanche as the root of a tree.

The branches of the tree are the nodes whose failure was directly triggered by this initial failure. For example, in Fig. 8.21 a, b, the breakdown of node *A* starts the avalanche, hence *A* is the root of the tree. The failure of *A* leads to the failure of *B* and *E*, which are the two branches of the tree. Subsequently *E* induces the failure of *D* and *B* leads to the failure of *C* Fig. 8.22a.

The branching model captures the essential features of this avalanche propagation process Fig. 8.22. The model starts with a single active node. In the next time step each active node produces *k* offsprings, where *k* is selected from a $p_k$ distribution. If a node selects $k = 0$, that branch dies out permanently Fig. 8.22b. If it selects $k > 0$, it will have *k* new active sites. The size of an avalanche corresponds to the size of the tree when all active sites died out Fig. 8.22c.

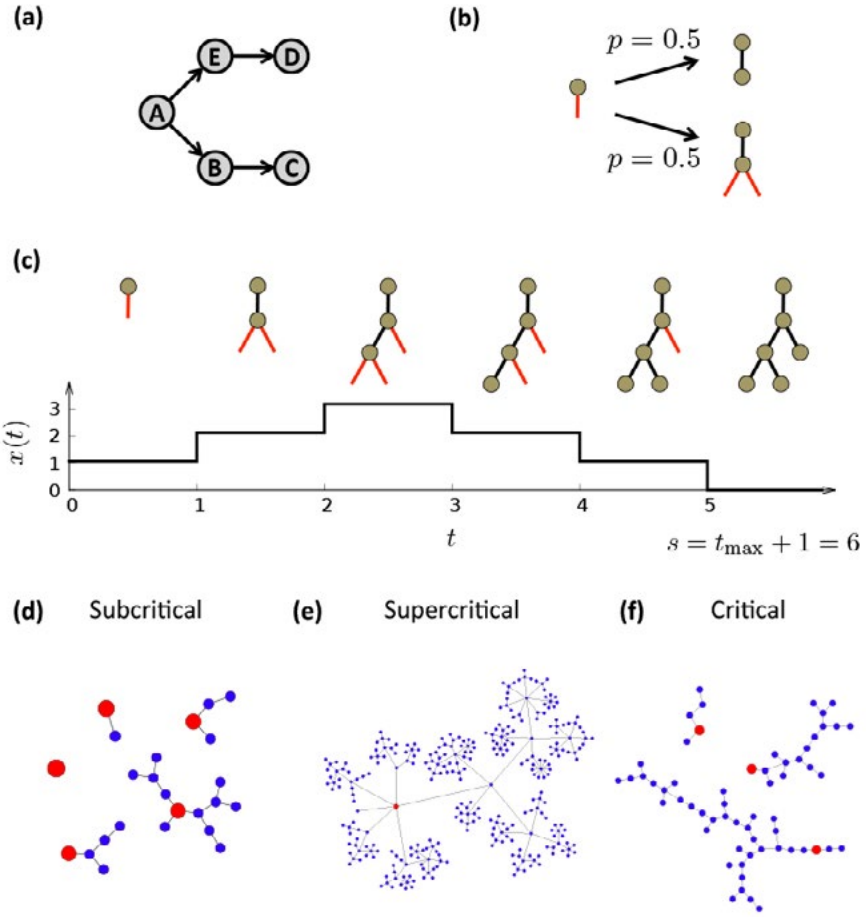



**Figure 8.22**
**Branching process**

**(a)** The branching process describing the propagation of the failure shown in **Fig. 8.20a,b**. The perturbation starts from node *A*, whose failure flips *B* and *E*, which in turn flip *C* and *D*, respectively.

**(b)** An elementary branching process. Each active link (red link) can become inactive with probability $p_0 = 1/2$ (top) or give birth to two new active links with probability $p_2 = 1/2$ (bottom).

**(c)** The number of active sites, *x(t)*, in function of time t. A nonzero *x(t)* means that the avalanche persists. When *x(t)* becomes zero, we loose all active sites and the avalanche ends. In the image this happens at $t = 5$, hence the size of the avalanche is $t_{max} + 1 = 6$. An exact mapping between the branching model and a one dimensional random walk helps us calculate the avalanche exponent. Consider a branching process starting from a stub with one active end. When the active site becomes inactive, it decreases the number of its active sites, i.e. $x \to x - 1$. When the active site branches, creates two active sites, i.e. $x \to x + 1$. This maps the avalanche size s to the time it takes for the walk that starts at $x = 1$ to reach $x = 0$ for the first time. This is a much studied process in random walk theory, predicting that the return time distribution follows a power law with exponent 3/2 [32]. For branching process corresponding to scale-free $p_k$, the avalanche exponent depends on γ, as predicted by **Fig. 8.15d, f**.

Typical avalanches generated by the branching model in the: subcritical *(d)*, supercritical *(e)* and critical regime *(f)*. The red node in each cascade marks the root of the tree, i.e. the first perturbation. In *d* and *f* we show multiple trees, while in *e* we show only one, as each tree grows indefinitely.

The branching model predicts the same three phases as those observed in the cascading failures model. These phases are determined by $\langle k \rangle$ of $p_k$:

- **Subcritical regime**
  If $\langle k \rangle < 1$, on average each branch has less then one offspring. Consequently each tree will terminate quickly Fig. 8.22d. In this regime the avalanche sizes follow an exponential distribution.

- **Supercritical regime**
  If $\langle k \rangle > 1$, on average each branch has more than one offspring. Consequently the tree will continue to grow indefinitely Fig. 8.22e. This captures the supercritical phase, when all avalanches are global.

- **Critical regime**
  If $\langle k \rangle = 1$, on average each branch has exactly one offspring. Consequently some trees are large; others die out shortly Fig. 8.21. Numerical simulations indicate that in this regime the avalanche size distribution follows a power law.

The branching model can be solved analytically, allowing us to predict the avalanche size distribution for an arbitrary $p_k$. If $p_k$ is bounded, *e.g.* it follows a binomial or exponential form, the calculations predict $\alpha = 3/2$. If, however, $p_k$ is scale-free, then the avalanche exponent depends on the power-law exponent $\gamma$ as Fig. 8.23 [31, 32]

$$\alpha = \begin{cases} 3/2, & \gamma \geq 3 \\ \gamma/(\gamma-1), & 2 < \gamma < 3 \end{cases}. \qquad (8.15)$$

We can revisit Table 8.2 in the light of Eq. 8.15, to confirm that the empirically observed avalanche exponents are all between 1.5 and 2, as predicted by Eq. 8.15.

In summary, numerous models capture the dynamics of cascading failures. These models differ in their realism as well as the number and the nature of their tuning parameters. Yet, their predictions are consistent with each other:

- They predict the existence of a critical state, in which the avalanche sizes follow a power law. The value of the avalanche exponent $\alpha$ depends on the degree exponent of the network on which the avalanche propagates, as predicted by Eq. 8.15.

- They predict the existence of a subcritical regime, in which all perturbations die out immediately, and a supercritical regime, when most perturbations sweep the whole system.

Note that a detailed modeling of cascading failures should also account for the fact that nodes and links have different capacities to carry traffic [33]. Such models are best discussed in the context of weighted networks.

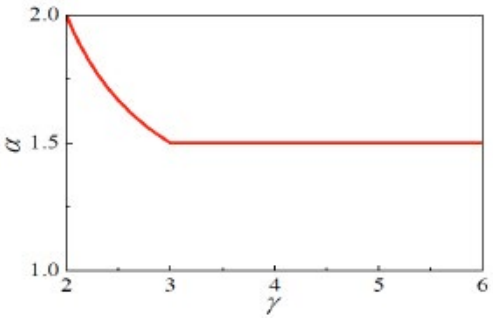



**Figure 8.23**
**The avalanche exponent is universal**

The dependence of the avalanche exponent $\alpha$ on the degree exponent $\gamma$ of the network on which the avalanche propagates, according to **Eq. 8.15**. The plot indicates that between $2 < \gamma < 3$ the avalanche exponent depends on the exponent of $p_k$. Beyond $\gamma = 3$, however, the avalanches behave as they would be spreading on a random network.

# BUILDING ROBUSTNESS

Can we enhance a network's robustness? In this section we take advantage of the insights we gained in the previous sections to design networks that are simultaneously robust to random failures and attacks. We also discuss mechanisms proposed to stop a cascading failure, allowing us to enhance a system's dynamical robustness. Finally, we apply the developed tools to the power grid, linking its robustness to its reliability.

## DESIGNING ROBUST TOPOLOGIES

The coexistence of robustness to random failures and fragility to attacks of scale-free networks prompts us to ask: could we design networks that are simultaneously robust to attacks and random failures [35, 36, 37, 38]? This appears to be a conflicting desire. For example, the hub-and-spoke network of Fig. 8.24a is robust to random failures, as only the failure of its central node can break the network into isolated components. Therefore, the probability that a random failure will fragment the network is $1/(N-1)$, which is negligible for large $N$. At the same time this network is rather vulnerable to attacks, as the removal of a single node, its central hub, will break the network into isolated nodes.

We can enhance this network's robustness to both failures and attacks by connecting its peripheral nodes Fig. 8.24b. There is a price, however, for this enhanced robustness: it requires us to double the number of links. If we define the cost to build and maintain a network to be proportional to its average degree $\langle k \rangle$, the cost of the network of Fig. 8.24b is 24/7, which is double of the cost 12/7 of the network of Fig. 8.24a. The increased cost helps us refine our question: can we maximize the robustness of a network to both random failures and targeted attacks, without changing the cost, i.e. keeping $\langle k \rangle$ constant?

To enhance a network's robustness against random failures we can increase its percolation threshold $f_c$, which denotes the moment when the network falls apart. As $f_c$ depends only on $\langle k \rangle$ and $\langle k^2 \rangle$ according to Eq. 8.7, the degree distribution which maximizes $f_c$ needs to maximize $\langle k^2 \rangle$ for a fixed $\langle k \rangle$. This is achieved by a bimodal distribution, whose nodes

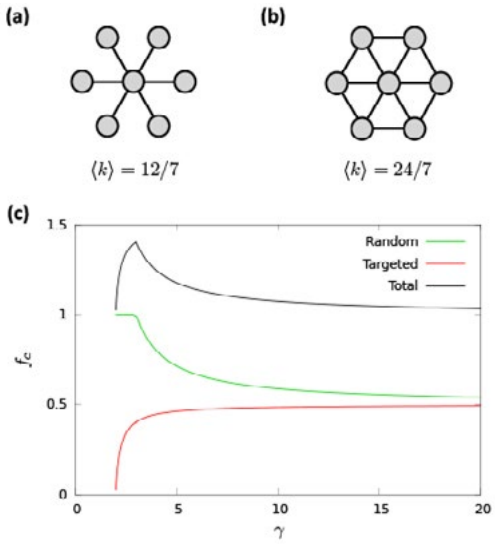



**Figure 8.24**
**Enhancing Robustness**

**(a)** A hub-and-spoke network is robust to random failures but has a low tolerance to an attack that removes its central hub.

**(b)** By connecting some of the small degree nodes, the reinforced network has a higher tolerance to targeted attacks. Yet, the cost, captured by the total number of links the network needs to maintain, i.e. $\langle k \rangle$, is higher in the reinforced network.

**(c)** Random, $f_c^{rand}$, targeted $f_c^{targ}$ and total $f_c^{tot}$ percolation thresholds for scale-free networks in function of the degree exponent $\gamma$. The plot is shown for $k_{min} = 3$.

have either degree $k_{min}$ or $k_{max}$, the two extreme values allowed in the respective network.

In a similar spirit, if we wish to optimize the network topology against both random failures and attacks, we search for topologies that maximize the sum Fig. 8.24c

$$f_c^{tot} = f_c^{rand} + f_c^{t\arg}. \tag{8.16}$$

A combination of analytical arguments and numerical simulations indicate that this too is best achieved by the bimodal degree distribution [35, 36, 37, 38]

$$p_k \equiv (1-r)\delta(k-k_{\min}) + r\delta(k-k_{\max}) \tag{8.17}$$

describing a network in which an $r$ fraction of nodes have degree $k_{max}$ and the remaining $(1 - r)$ fraction have degree $k_{min}$. As we show in ADVANCED TOPICS 8.G, the maximum of $f_c^{tot}$ is obtained when $r = 1/N$, i.e. when there is a single node with degree $k_{max}$ and the remaining nodes have degree $k_{min}$. The value of $k_{max}$ depends on the system size as (ADVANCED TOPICS 8.G)

$$k_{\max} = AN^{2/3}. \tag{8.18}$$

In other words, a network that is robust to both random failures and attacks has a single hub with degree Eq. 8.18, while the rest of the nodes have the same degree $k_{min}$. This configuration is obviously robust against random failures as the chance of removing the central hub is rather small. The obtained network may appear to be vulnerable to an attack that removes its $k_{max}$ hub, but it is not necessarily so. Indeed, the network's giant component is held together by both the central hub as well as by the many nodes with degree $k_{min}$, that for $k_{min} > 1$ form a giant component on their own. Hence while the removal of the $k_{max}$ hub causes a major time loss, the remaining low degree nodes are robust against subsequent targeted removal Fig. 8.25.

### HALTING CASCADING FAILURES

How can we avoid cascading failures? The first instinct is to reinforce the network through the addition of new links. If that is feasible, in some system may solve the problem. In others additional links could worsen the situation, offering more routes for the failure to spread. The true problem with reinforcement is that in most real systems the time frame needed to establish a new link is much larger than the timeframe of a cascading failure. For example thanks to regulatory, financial and legal barriers, building a new power line can take up to two decades. In contrast, a cascading failure can sweep the power grid in a few seconds. There is no way we can reinforce the network during this short time frame.

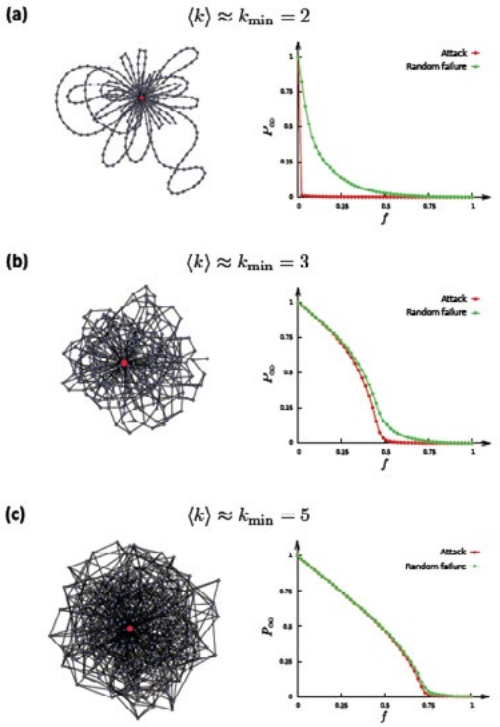



**Figure 8.25**
**Optimizing attack and failure tolerance**

The figure illustrates the optimal network topologies predicted by Eq. 8.16 and Eq. 8.17, consisting of a single hub of size Eq. 8.18 and the rest of the nodes have the same degree kmin determined by $\langle k \rangle$. The left panels illustrate the network topology for $N = 300$; the right panels show the failure/attack curves for $N = 10000$.

**(a)** For small $\langle k \rangle$ the hub plays a key role in holding the network together. Once we remove this central hub, given that $\langle k \rangle$ is small, the network breaks apart. Hence the attack and error curves are rather different.

**(b)** For larger $\langle k \rangle$ a giant component exists even without the central hub. Hence while the hub enhances the system's robustness to random failures, it is no longer essential for the network. In this case both the attack and error fc are large.

**(c)** For even larger $\langle k \rangle$ the error and the attack curves are indistinguishable, as the network is robust even without its central hub.

# BOX 8.5

**ROBUSTNESS, REDUNDANCY, RESILIENCE**

Redundancy and resilience are concepts deeply linked to robustness. It is useful, therefore, to clarify the relationship between them.

### Robustness

A system is robust if it can maintain its basic functions in the presence of internal and external errors. In a network context, robustness refers to the system's ability to carry out its basic functions, even when some of its nodes and links may be missing.

### Resilience

A system is resilient when it can adapt to internal and external errors by changing its method of operations while continuing to function. Hence, resilience is a dynamical property that requires a shift in the system's core activities.

### Redundancy

Redundancy implies the presence of parallel functions and components that, if needed, can replace a missing function or component. Networks show considerable redundancy in their ability to navigate information between two nodes, thanks to the multiple independent paths between most node pairs. For example, if you live in the United States, your local senator offers you a short path to the President. Yet, you may also reach to the president through many other, often equally short chains of acquaintances. A similar redundancy is built into the Internet: if a router fails, the packets normally handled by it are rerouted along alternative routes.

In a counterintuitive fashion, the impact of cascading failures can be lowered through selective node and link removal [39]. To do so we note that each cascading failure has two parts:

(i) Initial failure results in the breakdown of the first node or link, representing the source of the subsequent cascade.

(ii) Propagation, when the initial failure induces the breakdown of additional nodes and it starts cascading through the network.

In real networks the time interval between (i) and (ii) is much shorter than the time scale over which new nodes and links could be added to reinforce the network. Yet, simulations indicate that the size of a cascade can be reduced if we intentionally remove additional, well selected nodes, right after the initial failure, but before the failure could propagate. Even though the intentional removal of a node or a link increases the damage to the network, the removal of a well chosen component can suppress the cascade propagation. The mechanism is similar to the method used by firefighters, who set a controlled fire in the fire-line to consume the fuel in the path of a wildfire. The implementation of this procedure depends on the details of the spreading and failure mechanism, but simulations indicate that we can limit the size of the cascades if we remove nodes with small loads and links with large excess load in the vicinity of the initial failure.

A dramatic manifestation of the potentially positive effects of further damage is provided by the *Lazarus effect*, the ability to revive a bacteria that is unable to grow through the knockout of a few well selected genes [40] Fig. 8.26. Therefore, in a counterintuitive fashion, controlled damage can be beneficial to a network facing cascading failures.

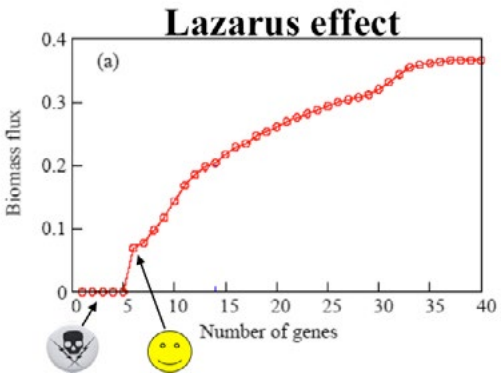



**Figure 8.26**
**Lazarus effect in bacteria**

A bacteria's growth is limited by its ability to generate biomass, the molecules the bacteria needs to build its cell wall, DNA and other cellular components. If some key genes are missing, preventing the bacteria from generating the necessary biomass, it cannot multiply and will likely die. Genes in whose absence the biomass flux is zero are called *essential*. The plot above shows the biomass flux for a mutant of E. *Coli*, a bacteria frequently studied by biologists. The mutant is missing an essential gene, hence its biomass flux is zero, as shown on the vertical axis. Consequently, it cannot multiply. Yet, the removal of five additional genes can turn on the biomass flux. Consequently, in a counterintuitive fashion, we can revive a dead organism, through the removal of further genes, a phenomena called *Lazarus effect* [40].

### CASE STUDY: ESTIMATING ROBUSTNESS

The European power grid is an ensemble of more than twenty national power grids consisting of over 3,000 generators and substations (nodes) and 200,000 km of transmission lines Fig. 8.27a-d. The degree distribution of this network can be approximated with Fig. 8.27e [41, 42]

$$p_k = \frac{e^{-k/\langle k\rangle}}{\langle k\rangle} \tag{8.19}$$

indicating that its topology is characterized by a single parameter, $\langle k\rangle$. As we showed in SECTION 5.5, such $p_k$ emerges in growing networks that lack preferential attachment. By determining $\langle k\rangle$ separately for each national power grid, we can predict the critical threshold $f_c$ for attacks, using the tools SECTION 8.3. As shown in Fig. 8.27f, for national power grids with $\langle k\rangle > 1.5$ there is a reasonable agreement between the observed and the predicted $f_c$ (Group 1). However, for national power grids with $\langle k\rangle < 1.5$ (Group 2) the predicted $f_c$ underestimates the real $f_c$, indicating that these national networks are more robust to attacks than expected based on their degree distribution. As we show next, this enhanced robustness correlates with the reliability of the respective national networks.

To test the relationship between robustness and reliability, we use several quantities collected for each power failure: (1) energy not supplied; (2) total loss of power; (3) average interruption time, measured in minutes per year. The measurements indicate that Group 1 networks, for which the real and the theoretical $f_c$ agree, represent two thirds of the full network size and share almost as much power and energy as the Group 2 networks. Yet, this group accumulates more than five times the average interruption time, more than two times the recorded power losses and almost four times the undelivered energy. Hence, Group 1 networks are more fragile than the Group 2 networks. This result offers direct evidence that networks that are topologically more robust are also more reliable. Note that this finding is rather counterintuitive: one would expect the denser networks to be more robust. We find, however, that the sparser power grids show enhanced robustness.

In summary, the results of this section indicate that a better understanding of the network topology is essential to develop strategies to improve robustness. We can improve robustness by either designing network topologies that are simultaneously robust to both random failures and attacks, or by designing interventions that limit the spread of cascading failures.

Our ability to design robust networks would suggest that we should redesign the topology of the Internet and the power grid to enhance their robustness [43]. Given the chance, this could indeed be achieved. Yet, these infrastructural networks were built incrementally over decades, following the self-organized growth process described in the previous chapters. Given the enormous cost of each link and node, it is unlikely that we would ever be given a chance to redesign them. In general the design principles of robust networks should be enforced only if robustness is an absolute criteria, like in the case of the wiring diagram of an airplane, whose failure could be fatal.

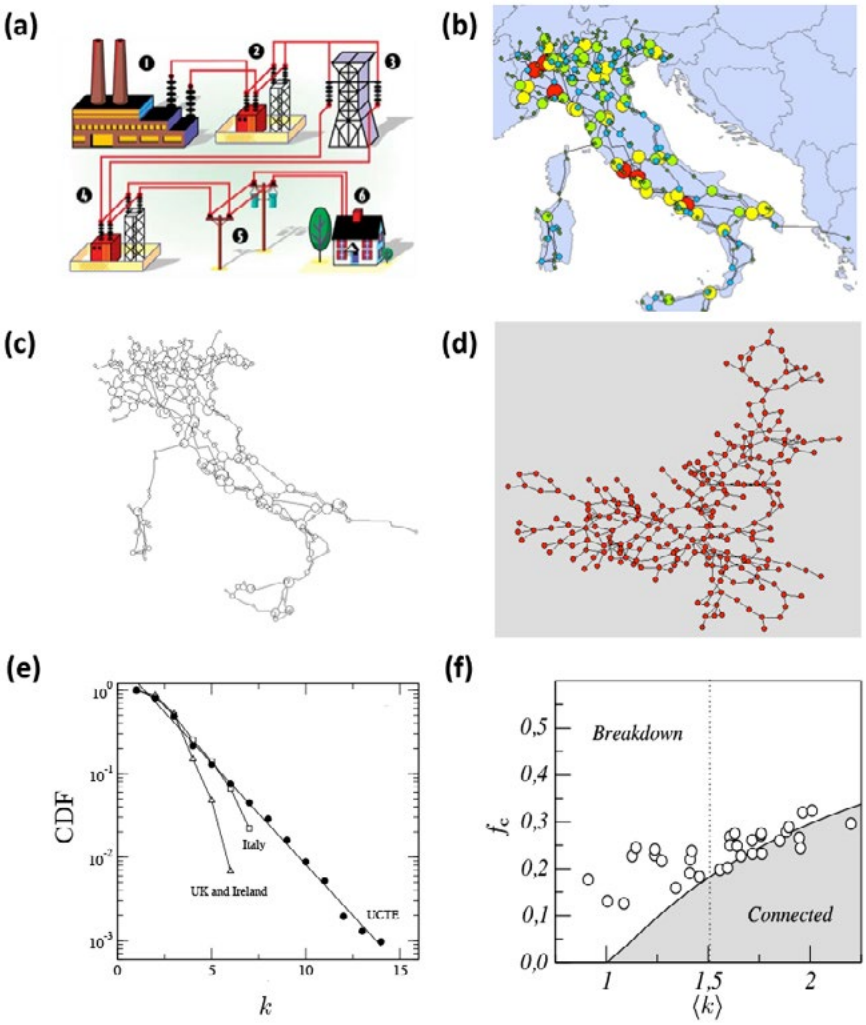



**Figure 8.27**
**The power grid**

**(a)** From a power engineer's perspective a power grid is a complex machinery co sisting of (1) power generators, (2) switching units, (3) the high voltage transmission grid, (4) transformers, (5) low voltage lines, (6) consumers, like households or busines es. When we study the network behind the power grid, many of these details are i nored. The necessary procedure to arrive to the network topologies amenable for study is illustrated in (b)-(d) for the Italian power grid.

**(b)** The power grid with the details of production and consumption. Once we strip these off, we obtain the spatial network shown in (c). Once the spatial information is also removed, we arrive to the network show in (d), which is the typical object of study at the network level.

**(e)** The cumulative degree distribution of the European power grid. The plot shows the data for the full network (UCTE) and separately for Italy, UK, and Ireland, indicating that the national grid's $p_k$ also follows the exponential (8.19).

**(f)** Phase space ($f_c$, $k$) for exponential uncorrelated networks under attack, where $f_c$ is the fraction of hubs we must remove to fragment the network. The continuous curve corresponds to the critical boundary for attacks, below which the network has retains its giant component. The plot also shows the estimated $f_c(\langle k \rangle)$ for attacks from the thirty-three national power grids within EU, shown as circles. The plot allows us to distinguish two classes of power grids. For countries with $\langle k \rangle > 1.5$ (Group 1), the analytical prediction for $f_c$ agrees with the numerically observed values. However, for countries with $\langle k \rangle > 1.5$ (Group 2) the analytical prediction underestimates the numerically observed values. Therefore, Group 2 national grids show enhanced robustness to attacks, which means that they are more robust than expected for a random network with the same degree sequence. Reliability measures indicate that the power grids in the robust Group 2 countries are more reliable. After [41].

# SUMMARY

The terrorist attacks of September 11, 2001 offered a vivid illustration of the important role hubs play in attacks. Indeed, the targets of the attack were not chosen at random: the World Trade Center in New York, the Pentagon, and the White House (an intended target) in Washington DC are the hubs of America's economic, military, and political power [44]. Yet, while causing a human tragedy far greater than any other event America has experienced since the Vietnam war, the attacks failed at their main goal: to topple the network. They did offer, however, an excuse to start new wars, like the Iraq and the Afghan wars, hence inducing a series of cascading events whose impact was far more devastating than the 9/11 terrorist attacks themselves. Yet, all networks, ranging from the economic to the military and the political web, survived. Hence, we can view 9/11 as a tale of robustness and network resilience. The roots of this robustness were uncovered in this chapter: real networks have a whole hierarchy of hubs. Taking out any one of them is not sufficient to topple the underlying network.

Network robustness represents good news for most complex systems. Indeed, there are uncountable errors in our cells, from misfolding proteins to the late arrival of a transcription factor. Yet, the robustness of the underlying cellular network helps our cells to carry on their normal functions. Network robustness also explains why we rarely notice the effect of router errors on the Internet or why the disappearance of a species does not lead to an immediate environmental catastrophe.

This topological robustness has its price, however: a fragility against attacks. As we showed in this chapter, the simultaneous removal of hubs will break any network. This is bad news for the Internet, as it allows crackers to design strategies that can harm this vital communication system. It is bad news for economic systems, as it indicates that hub removal can cripple the whole economy, vividly illustrated by the 2009 financial meltdown. Yet, it is good news for drug design, as it suggests that an accurate map of cellular networks can help us design drugs that can kill unwanted bacteria or cancer cells.

The main message of this chapter is simple: network topology, robustness, and fragility cannot be separated from one other. Rather, each complex system has its own Achilles' Heel: the networks behind them are robust to random failures but vulnerable to attacks. When considering robustness, we cannot ignore the fact that most systems have numerous controls and feedback loops that help them survive in the face of errors and failures. Internet protocols were designed to 'route around the trouble', guiding the traffic to avoid routers that malfunction; cells have numerous mechanisms to dismantle faulty proteins and to shut down malfunctioning genes. This chapter documented a new contribution to error tolerance: the structure of most complex systems preferred by nature offers them an enhanced error and failure tolerance. By the virtue of their topology only, real systems display a high degree of topological robustness.

The robustness of scale-free networks prompts us to ask: is this enhanced robustness the reason why many real networks are scalefree? Perhaps real systems have developed a scale-free architecture to satisfy their need for robustness. If this hypothesis is correct we should be able to set robustness as an optimization criteria and obtain a scale-free network. Yet, as we showed in SECTION 8.6, a network optimized for robustness has a hub-and-spoke topology. Its degree distribution is bimodal, rather than a power law. This suggests that robustness is not the force that drives the development of real networks. Rather, networks are scale-free thanks to growth and preferential attachment. It so happens that scale-free networks also have enhanced robustness. Yet, they are not the most robust networks we could design.

Finally, note that enhanced robustness does not require a network to be scale-free. Indeed, Eq. 8.7 links $f_c$ to $\langle k^2 \rangle$, hence any network whose $\langle k^2 \rangle$ is larger than expected for a random network will display enhanced robustness. Of course, if a network is scale-free with $\gamma < 3$, yet automatically displays enhanced robustness.

The results of this chapter allow us to formulate our next law:

### ACHILLES' HEEL

Scale-free networks are robust to random failures and fragile to attacks. Let us revisit the three criteria that prompts us to call this statement a network law:

#### A. Quantitative formulation

Eq. 8.7 indicates that the critical threshold of a scale-free network, capturing its response to random failures, converges to $f_c = 1$, implying an enhanced robustness to random node deletion. As we showed in SECTION 8.3, the finite threshold re-emergences under attacks.

#### B. Universality

Enhanced robustness is present in all networks whose $\langle k^2 \rangle$ is higher than expected in a random network. According to Table 4.1, this is true for most real networks.

## BOX 8.7

### NETWORK ROBUSTNESS: BRIEF HISTORY

The systematic study of network robustness within network science started with a paper published in Nature Fig. 8.1 by Réka Albert, Hawoong Jeong and Albert-László Barabási [1], that discovered the robustness of scale-free networks to random failures and their fragility to attacks. Yet, when it comes to our understanding of network robustness, particularly important were the contributions of Shlomo Havlin and his collaborators, like Reuven Cohen, who showed that the percolation threshold of a scale-free network is determined by the first two moments of the degree distribution.

**Figure 8.28**
**Shlomo Havlin**

A statistical physicist at Bar Ilan University in Israel, Havlin has played an important role in the development of network science. He derived (8.7), demonstrating that the origin of the enhanced robustness is rooted in the convergence of the the percolation threshold to one. His contributions to the field are diverse, from discovering the selfsimilar nature of real networks [45] to introducing the study of layered interdependent networks [46].

### C. Non-random origins

The phenomena discussed in this chapter is obviously absent from random networks, that have a finite threshold against both failures and attacks Fig. 8.13.

## BOX 8.7

**AT A GLANCE: NETWORK ROBUSTNESS**

**Malloy-Reed criteria:**

A giant component exists if

$$\kappa \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} > 2$$

**Critical threshold for random failures:**

$$f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}$$

Random Network: $f_c^{ER} = 1 - \frac{1}{\langle k \rangle}$

**Enhanced robustness:** $f_c > f_c^{ER}$

**Critical threshold for attacks:**

$$f_c^{\frac{2-\gamma}{1-\gamma}} = 2 + \frac{2-\gamma}{3-\gamma} k_{\min} (f_c^{\frac{3-\gamma}{1-\gamma}} - 1)$$

**Size distribution for cascading failures:**

$$p(s) \sim s^{-\alpha}$$

$$\alpha = \begin{cases} 3/2 & \gamma > 3 \\ \frac{\gamma}{\gamma - 1} & 2 < \gamma < 3 \end{cases}$$

# ADVANCED TOPICS 8.A
# RANDOM NETWORKS AND PERCOLATION

### RANDOM NETWORKS AND INFINITE DIMENSIONAL LATTICE

The percolation transition observed when nodes are randomly removed from a random network is characterized by the same critical exponent as percolation in $d > 6$ dimensions. This equivalence is illustrated by the following two-step argument.

i. The Cayley tree, a regular branching tree shown in Fig. 8.29a, is a frequently used as a model of an infinite dimensional lattice. Indeed, in $d$-dimensions the volume of a hypersphere of radius $r$ is proportional to $r^d$, whereas its surface is $r^{d-1}$. Hence, in general, we have $surface \propto volume^{1-1/d}$. For example, in $d = 2$ the area (volume) of a circle of radius $r$ is $\pi r^2$ and its circumference (surface) is $2\pi r$. In $d = 3$, the volume depends on $r$ as $r^3$ whereas the area increases as $r^2$. In the $d \to \infty$ limit the surface of a $d$-dimensional hypersphere is proportional to its volume, as $1/d$ becomes negligible. This proportionality is valid for the Cayley tree: the number of nodes in the outer layer of a Cayley tree (surface) equals the number of nodes inside the tree (volume). Hence, we can view the Cayley tree as an infinite dimensional object Fig. 8.29a.

ii. At the same time the Cayley tree captures the local topology of a random network. Indeed, in a very large random network the probability of finding loops is negligible. Hence, locally the network is a tree. To see this consider a cluster of three nodes occupied with pebbles on a $d$-dimensional cubic lattice (red nodes in $d = 2$, Fig. 8.29b). If we add an additional pebble, for it to be part of the cluster, it has to be adjacent to at least one of the three previous pebbles. We can place the new pebble in $3(2d - 2)$ possible spots so that it does not form a loop (green sites); only one of the site closes the loop (red site). Therefore, the probability that the four pebbles form a loop decreases as,

$$\frac{1}{3(2d-2)+1} \qquad (8.19b)$$

which is negligible in the $d \to \infty$ limit [2]. Consequently, locally the nodes in a random network form a tree, well approximated by the Cayley tree of Fig. 8.29a.

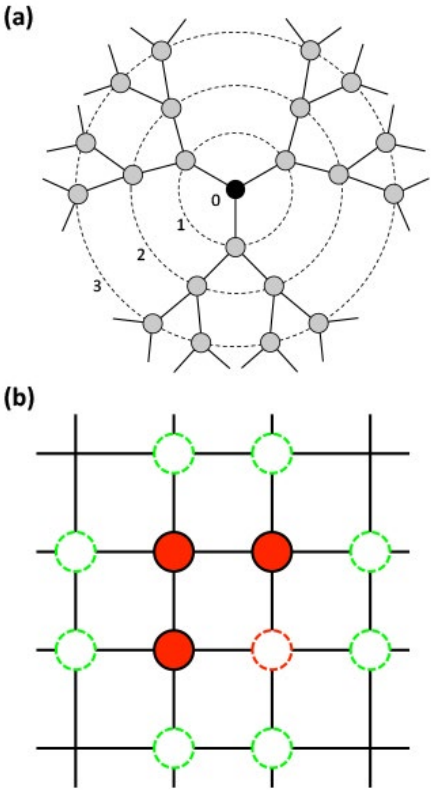



**Figure 8.29**
**Infinite dimensional percolation**

**(a)** The Cayley tree represents a model of an infinite dimensional lattice, as the number of nodes on its surface is proportional to the total number of nodes within the tree, corresponding to its volume. This is a property of an infinite dimensional lattice as well.

**(b)** If we form a four-node cluster in a square lattice, the likelihood that they form a loop is negligible. Indeed, we have positions for the new node to form a cluster (green dashed circles); one (red dashed) leads to a loop.

In summary, percolation in a random network is in the same universa ity class as percolation on a Cayley tree, which in turn is in the universality class of infinite dimensional percolation [2]. Therefore, $\gamma_p = 1$, $\beta = 1$ and $\gamma = 1$, the percolation critical exponents for $d = \infty$, characterize the behavior of the clusters near the critical point $f_c$ when an fraction of nodes are randomly selected and removed from a random network.

### PERCOLATION EXPONENTS FOR SCALE-FREE NETWORKS

To understand how a scale-free network breaks apart as we approach the threshold Eq. 8.7, we need to determine the critical exponents $\gamma_p$, $\beta$ and $\nu$. The calculations show that the scale-free property alters the value of these exponents, leading to systematic deviations from the exponents discussed in SECTION 8.1 that characterize random networks.

Let us start with the probability $P_\infty$ that a randomly selected node belongs to the giant component. According to Eq. 8.2 this follows a power law near $p_c$ (or $f_c$ in the case of node removal). The calculations predict that for a scale-free network the exponent $\beta$ depends on the degree exponent $\gamma$ as [7, 47, 48, 49, 50]

$$\beta = \begin{cases} \dfrac{1}{3-\gamma} & 2 < \gamma < 3, \\ \dfrac{1}{\gamma-3} & 3 < \gamma < 4, \\ 1 & \gamma > 4. \end{cases} \tag{8.20}$$

Hence, while for a random network (captured by the $\gamma > 4$ regime) we have $\beta = 1$, for most scale-free networks of practical interest $\beta > 1$. Therefore, the collapse of the giant component is steeper at the critical point in a scale-free network than in a random network.

The exponent describing the average component size near $p_c$ follows [47]

$$\gamma_p = \begin{cases} 1 & \gamma > 3 \\ -1 & 2 < \gamma < 3. \end{cases} \tag{8.21}$$

The negative $\gamma_p$ for $\gamma < 3$ may appear surprising. Note, however, that for $\gamma < 3$ we always have a giant component. Hence, the divergence Fig. 8.1 cannot be observed in this regime. For a random graph of arbitrary degree distribution the size distribution of the finite clusters follows [47, 49, 50]

$$n_s \sim s^{-\tau} e^{-s/s^*}. \tag{8.22}$$

Here, $n_s$ is the number of clusters of size $s$ and $s^*$ is the crossover cluster size. At criticality

$$s^* \sim |q - q_c|^{-\sigma}. \tag{8.23}$$

The pertinent critical exponents are

$$\tau = \begin{cases} \dfrac{5}{2} & \gamma > 4 \\ \dfrac{2\gamma - 3}{\gamma - 2} & 2 < \gamma < 4, \end{cases}$$

$$\sigma = \begin{cases} \dfrac{3-\gamma}{\gamma-2} & 2 < \gamma < 3 \\ \dfrac{\gamma-3}{\gamma-2} & 3 < \gamma < 4 \\ \dfrac{1}{2} & \gamma > 4. \end{cases} \tag{8.24}$$

Once again, the random network values $\tau = 5/2$ and $\sigma = 1/2$ are recovered for $\gamma > 4$. Finally, the exponent $\nu$ governs the average size of the finite components, obeying the scaling relation $\nu = (3 - \tau)/\sigma$. Hence,

$$\nu = \begin{cases} \dfrac{\gamma-3}{\gamma-2}, & 3 < \gamma < 4 \\ \dfrac{3-\gamma}{\gamma-2}, & 2 < \gamma < 3. \end{cases} \tag{8.25}$$

In summary, the exponents describing the breakdown of a scale-free network depend on the degree exponent $\gamma$. This is true even in the range $3 < \gamma < 4$, where the percolation transition occurs at a finite threshold $f_c$. The mean-field behavior predicted for percolation in infinite dimensions, capturing the response of a random network to random failures, is recovered for $\gamma > 4$.

# ADVANCED TOPICS 8.B
# MALLOY-REED CRITERIA

The purpose of this section is to derive the Malloy-Reed criteria BOX 8.7, which allows us to calculate the percolation threshold of an arbitrary network [6]. For a giant component to exist each node that belongs on average to it must be connected to at least two other nodes Fig. 8.8. Therefore, the average degree $k_i$ of a randomly chosen node i that is part of the giant component should be at least 2. Denote with $p(k_i \mid i \leftrightarrow j)$ the joint probability that a node in a network with degree $k_i$ is connected to an arbitrary node $j$ that is part of the giant component. This conditional probability allows us to determine the expected degree of node $i$ as

$$\langle k_i \mid i \leftrightarrow j \rangle = \sum_{k_i} k_i P(k_i \mid i \leftrightarrow j) = 2 \,. \tag{8.26}$$

In other words, $\langle k_i \mid i \leftrightarrow j \rangle$ should be equal to two, the condition for node $i$ to be part of the giant component. We can write the probability appearing in the sum Eq. 8.26 as

$$P(k_i \mid i \leftrightarrow j) = \frac{P(k_i, i \leftrightarrow j)}{P(i \leftrightarrow j)} = \frac{P(i \leftrightarrow j \mid k_i) p(k_i)}{P(i \leftrightarrow j)} \,. \tag{8.27}$$

where we used Bayes' theorem in the last term. For a network with degree distribution $p_k$, in the absence of degree correlations, we can write

$$P(i \leftrightarrow j) = \frac{2L}{N(N-1)} = \frac{\langle k \rangle}{N-1} \qquad P(i \leftrightarrow j \mid k_i) = \frac{k_i}{N-1} \,. \tag{8.28}$$

which expresses the fact that we can choose between $N-1$ nodes to link to, each with probability $1/(N-1)$ and that we can try this $k_i$ times. We can now return to Eq. 8.26,

$$\sum_{k_i} k_i P(k_i \mid i \leftrightarrow j) = \sum_{k_i} k_i \frac{P(i \leftrightarrow j \mid k_i) p(k_i)}{P(i \leftrightarrow j)} = \sum_{k_i} k_i \frac{k_i p(k_i)}{\langle k \rangle} = \frac{\sum_{k_i} k_i^2 p(k_i)}{\langle k \rangle} \tag{8.29}$$

With that we arrive at the Malloy-Reed criteria Eq. 8.4, indicating that the condition to have a giant component is

$$\kappa \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} > 2 \,. \tag{8.30}$$

# ADVANCED TOPICS 8.C CRITICAL THRESHOLD UNDER RANDOM FAILURES

The purpose of this section is to derive Eq. 8.7, that provides the critical threshold for random node removal [7, 50]. The random removal of an $f$ fraction of nodes has two consequences:

- It alters the degree of some nodes, as nodes that were previously connected to the removed nodes will lose some links [$k \to k' \leq k$].
- It changes the degree distribution, as the neighbors of the missing nodes will have an altered degree [$p_k \to p'_{k'}$].

To be specific, after we randomly remove an f fraction of nodes, a node with degree $k$ turns into a node with degree $k'$ with probability

$$\binom{k}{k'} f^{k-k'}(1-f)^{k'} \quad k' \leq k \,. \tag{8.31}$$

The first $f$-dependent term in Eq. 8.31 accounts for the fact that the selected node lost $(k - k')$ links, each with probability $f$; the next term accounts for the fact that node removal leaves $k'$ links untouched, each with probability $(1 - f)$.

The probability that we have a degree-$k$ node in the original network is $p_k$; the probability that we have a new node with degree $k'$ in the new network is

$$p'_{k'} \simeq \sum_{k=k'}^{k \equiv k'} p_k \binom{k}{k'} f^{k-k'}(1-f)^{k'}. \tag{8.32}$$

Let us assume that we know $\langle k \rangle$ and $\langle k^2 \rangle$ for the original degree distribution $p_k$. Our goal is to calculate $\langle k' \rangle$, $\langle k'^2 \rangle$ for the new degree distribution $p'_{k'}$, obtained after we randomly removed an $f$ fraction of the nodes. For this we write

$$\begin{aligned}\langle k'\rangle_f &= \sum_{k'=0}^{\infty} k' p'_{k'} \\ &= \sum_{k'=0}^{\infty} k' \sum_{k=k'}^{\infty} p_k \left(\frac{k!}{k'!(k-k')}\right) f^{k-k'}(1-f)^{k'} \\ &= \sum_{k'=0}^{\infty} k' \sum_{k'=k'}^{\infty} p_k \frac{k!}{k'!(k-k')!} f^{k-k'}(1-f)^{k'}.\end{aligned} \tag{8.33}$$

The sum above is performed over the triangle shown in Fig. 8.30. We can check that we are performing the same sum if we change the order of summation together with the limits of the sums as

$$= \sum_{k'=0}^{\infty} \sum_{k=k'}^{\infty} = \sum_{k=0}^{\infty} \sum_{k'=0}^{k}. \tag{8.34}$$

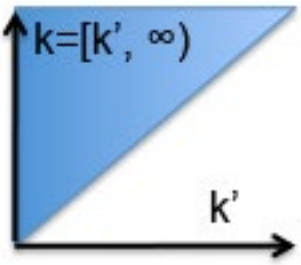



**Figure 8.30**
**The integration domain**

In **Eq. 8.34** we changed the integration order, *i.e.* the order of the two sums. We could do so because both sums are defined over the triangle shown in blue in the figure.

Hence we obtain

$$\begin{aligned}\langle k'\rangle_f &= \sum_{k=0}^{\infty} k' \sum_{k'=0}^{k} p_k \frac{k!}{(k'-1)!(k-k')!} f^{k-k'}(1-f)^{k'-1}(1-f) \\ &= \sum_{k=0}^{\infty} (1-f) k p_k \sum_{k'=0}^{k} \frac{(k-1)!}{(k'-1)!(k-k')!} f^{k-k'}(1-f)^{k'-1} \\ &= \sum_{k=0}^{\infty} (1-f) k p_k \sum_{k'=0}^{k} \left(\frac{k-1}{k'-1}\right) f^{k-k'}(1-f)^{k'-1} \\ &= \sum_{k=0}^{\infty} (1-f) k p_k \\ &= (1-f)\langle k\rangle.\end{aligned} \tag{8.35}$$

This connects $\langle k'\rangle$ to the original $\langle k\rangle$ after the random removal of an $f$ fraction of nodes. We perform a similar calculation for $\langle k'^2\rangle$:

$$\begin{aligned}\langle k'2\rangle_f &= \langle k'(k'-1)+k'\rangle_f \\ &= \langle k'(k'-1)\rangle_f + \langle k'\rangle_f \\ &= \sum_{k'=0}^{\infty} k'(k'-1) p'_{k'} + \langle k'\rangle_f.\end{aligned} \tag{8.36}$$

Again, we change the order of the sums Fig. 8.30

$$\begin{aligned}\langle k'(k'-1)\rangle_f &= \sum_{k'=0}^{\infty} k'(k'-1) p'_{k'} \\ &= \sum_{k'=0}^{\infty} k'(k'-1) \sum_{k=k'}^{\infty} p_k \left(\frac{k}{k'}\right) f^{k-k'}(1-f)^{k'} \\ &= \sum_{k'=0}^{\infty} k'(k'-1) \sum_{k'=0}^{k} p_k \frac{k!}{k'!(k-k')!} f^{k-k'}(1-f)^{k'} \\ &= \sum_{k'=0}^{\infty} \sum_{k'=0}^{k} p_k \frac{k!}{(k'-2)!(k-k')!} f^{k-k'}(1-f)^{k'-2}(1-f)^2 \\ &= \sum_{k=0}^{\infty} (1-f)^2 k(k-1) p_k \sum_{k'=0}^{k} \frac{(k-2)!}{(k'-2)!(k-k')!} f^{k-k'}(1-f)^{k'-2} \\ &= \sum_{k=0}^{\infty} (1-f)^2 k(k-1) p_k \sum_{k'=0}^{k} \left(\frac{k-2}{k'-2}\right) f^{k-k'}(1-f)^{k'-2} \\ &= \sum_{k=0}^{\infty} (1-f)^2 k(k-1) p_k \\ &= (1-f)^2 \langle k(k-1)\rangle.\end{aligned} \tag{8.37}$$

Hence we obtain

$$\begin{aligned}\left\langle k'^2\right\rangle_f &= \left\langle k'(k'-1)+k'\right\rangle_f \\ &= \left\langle k'(k'-1)\right\rangle_f + \left\langle k'\right\rangle_f \\ &= (1-f)^2\left\langle k(k-1)\right\rangle + (1-f)\langle k\rangle \\ &= (1-f)^2\left(\left\langle k^2\right\rangle - \langle k\rangle\right) + (1-f)\langle k\rangle \\ &= (1-f)^2\left\langle k^2\right\rangle - (1-f)^2\langle k\rangle + (1-f)\langle k\rangle \\ &= (1-f)^2\left\langle k^2\right\rangle - \left(-f^2+2f-1+1-f\right)\langle k\rangle \\ &= (1-f)^2\left\langle k^2\right\rangle + f(1-f)\langle k\rangle .\end{aligned} \tag{8.38}$$

which connects $\langle k'^2\rangle$ to the original $\langle k^2\rangle$ after the random removal of an $f$ fraction of nodes. Let us put the results Eq. 8.35 and Eq. 8.38 together:

$$\langle k'\rangle_f = (1-f)\langle k\rangle \tag{8.39}$$

$$\langle k'\rangle_f = (1-f)^2\langle k^2\rangle + f(1-f)\langle k\rangle \tag{8.40}$$

According to the Malloy-Reed criteria the breakdown threshold is given by the equality

$$\kappa \equiv \frac{\langle k'^2\rangle_f}{\langle k'\rangle_f} = 2 \tag{8.41}$$

Inserting Eq. 8.38 and Eq. 8.40 into Eq. 8.41 we obtain our final result Eq. 8.7 in SECTION 8.3,

$$f_c = 1 - \frac{1}{\frac{\langle k^2\rangle}{\langle k\rangle} - 1} \tag{8.42}$$

providing the breakdown threshold of networks with arbitrary $p_k$ under random node removal.

# ADVANCED TOPICS 8.D BREAKDOWN OF A FINITE SCALE-FREE NETWORK

The goal of this section is to determine the dependence Eq. 8.10 of the breakdown threshold of a scale-free network on the network size $N$. We start by calculating the $m^{th}$ moment of a power-law distribution

$$\langle k^m \rangle = (\gamma-1)k_{\min}^{\gamma-1}\int_{k_{min}}^{k_{\max}} k^{m-\gamma}dk \qquad = \frac{(\gamma-1)}{(m-\gamma+1)}k_{\min}^{\gamma-1}[k^{m-\gamma+1}]_{k_{min}}^{k_{max}} \tag{8.43}$$

As discussed in CHAPTER 4, we have

$$k_{max} = k_{min}N^{\frac{1}{\gamma-1}} \tag{8.44}$$

$$\langle k^m \rangle = \frac{(\gamma-1)}{(m-\gamma+1)}k_{min}^{\gamma-1}[k_{\max}^{m-\gamma+1} - k_{min}^{m-\gamma+1}] \tag{8.45}$$

To calculate $f_c$ we determine the ratio

$$\kappa \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{(2-\gamma)}{(3-\gamma)}\frac{k_{\max}^{3-\gamma} - k_{\min}^{3-\gamma}}{k_{\max}^{2-\gamma} - k_{\min}^{2-\gamma}}, \tag{8.46}$$

which in the $N \to \infty$ (and hence the $k_{max} \to \infty$) limit depends on $\gamma$ as

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle} = \left|\frac{2-\gamma}{3-\gamma}\right| \begin{cases} k_{min} & \gamma > 3 \\ k_{\max}^{3-\gamma}k_{\min}^{\gamma-2} & 3 > \gamma > 2 \\ k_{max} & 2 > \gamma > 1 \end{cases} \tag{8.47}$$

The breakdown threshold is given by

$$f_c = 1 - \frac{1}{\kappa - 1}, \tag{8.48}$$

where $\kappa$ is given by Eq. 8.46. Inserting Eq. 8.43 into Eq. 8.42 and Eq. 8.47, we obtain

$$f_c \simeq 1 - \frac{C}{N^{\frac{3-\gamma}{\gamma-1}}}, \tag{8.49}$$

which is Eq. 8.10, providing the dependence of $f_c$ on $N$.

# ADVANCED TOPICS 8.E
# THRESHOLD UNDER ATTACK

The goal of this section is to explore how a scale-free network responds to attack, by deriving Eq. 8.12. In other words, we calculate $f_c$ for an uncorrelated scale-free networks, generated by the configuration model with $p_k = c \cdot k^{-\gamma}$ where $k = k_{min},..., k_{max}$ and $c \approx (\gamma - 1)/(k_{min}^{-\gamma+1} - k_{max}^{-\gamma+1})$.

The removal of an $f$ fraction of nodes in a decreasing order of their degree (hub removal) has two effects [9, 50]:

**(i)** The maximum degree of the network changes from $k_{max}$ to $k'_{max}$.

**(ii)** The links connected to the removed hubs are also removed, changing the degree distribution of the remaining network.

The resulting network is still uncorrelated, therefore we can use the Molloy-Reed criteria to determine the existence of a giant component. We start by considering the impact of (i). The new upper cutoff, $k'_{max}$, is given by

$$f = \int_{k_{\max}}^{k'_{\max}} p_k dk = \frac{\gamma - 1}{\gamma - 1} \frac{k'^{-\gamma+1}_{\max} - k^{-\gamma+1}_{\max}}{k^{-\gamma+1}_{\min} - k^{-\gamma+1}_{\max}}, \tag{8.50}$$

If we assume that $k_{max} \gg k'_{max}$ and $k_{max} \gg k_{min}$ (true for large scale-free networks with natural cutoff), we can ignore the $k_{max}$ terms, obtaining

$$f = \left( \frac{k'_{\max}}{k_{\min}} \right)^{-\gamma+1}, \tag{8.51}$$

which leads to

$$k'_{max} = k_{min} f^{\frac{1}{1-\gamma}}. \tag{8.52}$$

Eq. 8.51 provides the new maximum degree of the network after we remove an $f$ fraction of the hubs.

Next, we turn to (ii), accounting for the fact that hub removal also removes the links connected to these hubs, changing the degree distribution $p_k \rightarrow p'_k$. In the absence of degree correlations we can assume that the links of the removed hubs connect to randomly selected stubs. Consequently, we calculate the fraction of links removed 'randomly', $\tilde{f}$, as a consequence of removing an $f$ fraction of the hubs:

$$\tilde{f} = \frac{\int_{k_{max'}}^{k_{max}} kp_k dk}{\langle k \rangle} = \frac{1}{\langle k \rangle} c \int_{k_{max'}}^{k_{max}} k^{-\gamma+1} dk \qquad (8.53)$$

$$= \frac{1}{\langle k \rangle} \frac{1-\gamma}{2-\gamma} \frac{k'_{max}{}^{-\gamma+2} - k_{max}^{-\gamma+2}}{k_{min}{}^{-\gamma+1} - k_{max}^{-\gamma+2}}$$

Ignoring the $k_{max}$ again and using $\langle k \rangle \approx \frac{\gamma-1}{\gamma-2} k_{min}$ we obtain

$$\tilde{f} = \left( \frac{k'_{max}}{k_{min}} \right)^{-\gamma+2}. \qquad (8.54)$$

Using Eq. 8.49 we obtain:

$$\tilde{f} = f^{\frac{2-\gamma}{1-\gamma}}. \qquad (8.55)$$

For $\gamma \rightarrow 2$ we have $\tilde{f} \rightarrow 1$, which means that the removal of a tiny fraction of the hubs removes all links, potentially destroying the network. The reason is that for $\gamma = 2$ the hubs dominate the network. The degree distribution of the remaining network is

$$p'_{k'} = \sum_{k=k_{min}}^{k'_{max}} \binom{k}{k'} \tilde{f}^{k-k'} (1-\tilde{f})^{k'} p_k. \qquad (8.56)$$

Note that we obtained the same degree distribution as Eq. 8.27 in ADVANCED TOPICS 5.B. This means that now we can use the calculation method developed for random node removal. To be specific, we calculate $\kappa$ for a scale-free network with $k_{min}$ and $k'_{max}$ using Eq. 8.45:

$$\kappa = \frac{2-\gamma}{3-\gamma} \frac{k'^{3-\gamma}_{max} - k^{3-\gamma}_{min}}{k'^{2-\gamma}_{max} - k^{2-\gamma}_{min}}. \qquad (8.57)$$

Substituting into this Eq. 8.57 we have

$$\kappa = \frac{2-\gamma}{3-\gamma} \frac{k^{3-\gamma}_{min} f^{(3-\gamma)/(1-\gamma)} - k^{3-\gamma}_{min}}{k^{2-\gamma}_{min} f^{(2-\gamma)/(1-\gamma)} - k^{2-\gamma}_{min}} = \frac{2-\gamma}{3-\gamma} k_{min} \frac{f^{(3-\gamma)/(1-\gamma)} - 1}{f^{(2-\gamma)/(1-\gamma)} - 1}. \qquad (8.58)$$

After simple transformations we obtain:

$$f_c^{\frac{2-\gamma}{1-\gamma}} = 2 + \frac{2-\gamma}{3-\gamma} k_{min} \left( f_c^{\frac{3-\gamma}{1-\gamma}} - 1 \right). \qquad (8.59)$$

which is Eq. 8.12 in SECTION 8.3.

# ADVANCED TOPICS 8.F
# MODELING CASCADING FAILURES

In this section we discuss two additional cascading failure models, that together with the Failure Propagation Model and the Branching Model discussed in SECTION 8.5, help illustrate the universality of the mechanisms governing cascading failures.

### OVERLOAD MODEL

The overload model was proposed to capture the emergence of large blackouts [17]. The model has $N$ identical nodes (components), each node $j$ assigned an initial load $L_j$, which is a random variable uniformly distributed between $L_{min}$ and $L_{max}$. A node fails when its load exceeds a preassigned threshold, $L_{fail}$ assumed to be the same for all nodes. When a node fails, a fixed amount of power $P$ is transferred to all other nodes in the network. Hence the impact of each failure is global, affecting not only the neighbors of the failed node, but all other nodes.

This mimics the fact that after each node or link failure the electric currents rearrange themselves globally. Hence, the impact of a local failure is not limited to the failed node's or link's direct neighbors, but can alter the current flowing through all nodes and links. Consequently, the model's behavior is independent of the network topology: the system behaves as if it would be fully connected.

To begin a cascade, we assume an initial disturbance that adds to the load of each component an additional load $P$. Some nodes with high initial loads $L_i$ may fail and each such failure distributes an additional load $P$ to the remaining nodes, potentially causing further failures Fig. 8.31a, b. The model's behavior is captured by the phase diagram of Fig. 8.31c, predicting three regimes:

- **Subcritical regime**
  If the initial load $L_{min}$ is under a global threshold, then most local perturbations die out, hence we do not observe global avalanches. In this regime the avalanche size distribution is bounded Fig. 8.31d.

- **Supercritical regime**
  If the initial load $L_{min}$ is over the threshold, then the perturbations propagate, resulting in avalanches that involve most nodes. In this regime the avalanche size distribution is again bounded and bimodal, capturing the coexistence of only very large and very small avalanches.

- **Critical regime**
  At the boundary of the subcritical and the supercritical regime, the avalanche size distribution follows a power law. The observed avalanche exponent is $\alpha = 3/2$.

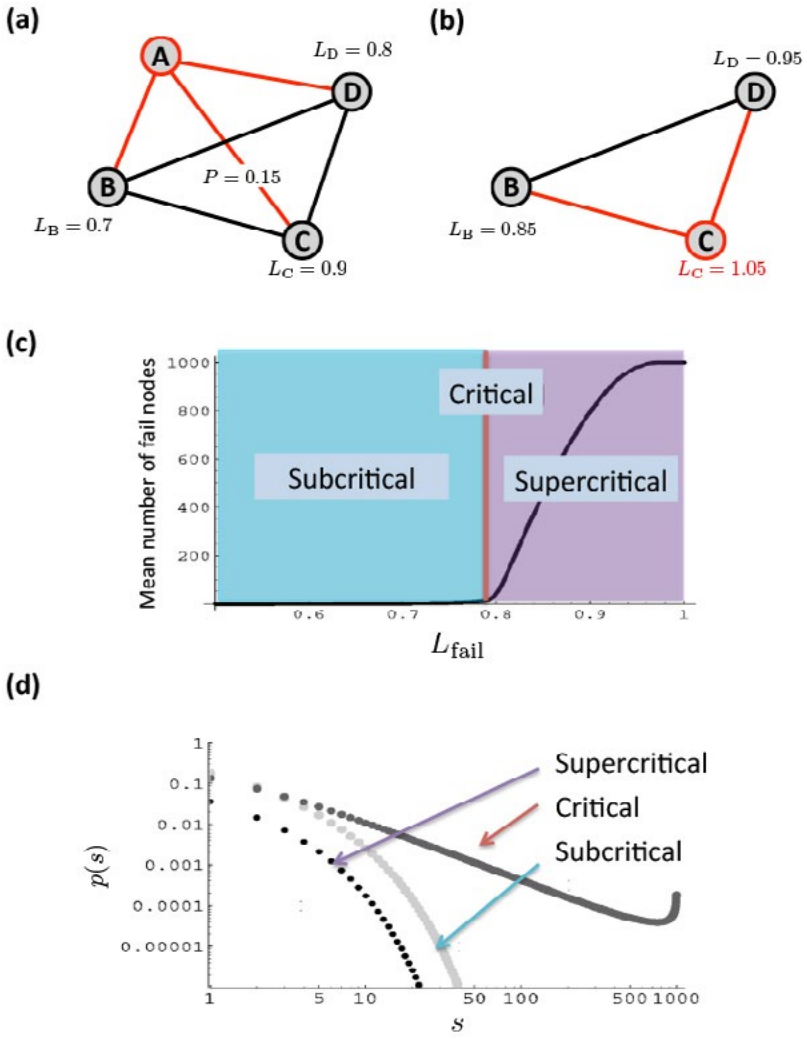



**Figure 8.31**
**Avalanche sizes in the overload**

**(a)** To illustrate the dynamics of the overload model, we start with four nodes, each assigned an initial load $L_i$. When node $A$ is perturbed by increasing its load above the threshold $L_{fail} = 1$, it fails and redistributes a fixed amount of power $P = 0.15$ to all other nodes in the network. Hence, the remaining nodes will now have a load $L_i = L_i + P$.

**(b)** The extra power will cause node $C$ to also fail, as now $L_c = 1.05 > L_{fail}$. That event adds another $P$ power to the remaining nodes, prompting them to fail too, allowing the avalanche to sweep the full network.

**(c)** Phase Diagram: Mean number of failed components in function of $L_{fail}$ in the overload model. At the critical loading $L_{fail} = 0.8$ the model changes its behavior, supporting large avalanches.

**(d)** Log-log plot showing the distribution of the number of components failed for three values of $L_{fail}$. Note the power-law scaling of $p(s)$ for the critical load $L = 0.8$.

After [17].

### SANDPILE MODEL

A common feature of the cascade and the overload models is that the empirically documented power-law behavior appears only in the critical regime. Therefore, we need to tune the model parameters to reach this critical state. Outside this critical regime the avalanche sizes follow an exponential distribution. This raises an important question: how does nature drive these systems to criticality? Attempts to answer this question have lead to a family of models collectively

called self-organized critical (SOC) models [51]. These models do not require external tuning, but self-organize to a critical state. Probably the best known of these is the sandpile model [52] that mimics the dropping of single grains on a plane. When the pile gets too high on a site, it topples by moving its grains to the neighboring sites. The model is typically defined on a lattice. With interest in cascading failures, its network version has been explored [30, 31]. The model starts from a network with an arbitrary wiring diagram and evolves following these steps:

1. Each node i is given a prescribed threshold $z_i (\le k_i)$. We denote with $[z_i]$ the smallest integer not smaller than $z_i$ ($z_i \le k_i$).

2. At each time step a grain is added at a randomly chosen node i, so that the height of the node i increases by one ($h_i \to h_i + 1$).

3. If the height of node *i* reaches or exceeds $z_i$, it becomes unstable and $z_i$ grains on the node topple to $z_i$ randomly chosen adjacent nodes among the $k_i$ neighbors of node *i*. Therefore, $h_i \to h_i + 1 - z_i$ and $h_i \to h_i + 1$ for $z_i$ neighboring nodes of *j*.

4. If this toppling causes any of the adjacent nodes to become unstable, $z_i$ subsequent toppings follow until there is no unstable node left. This process defines an avalanche.

We repeat the steps 2–4 and determine the avalanche size *s* in each case, where *s* represents the number of toppling events in a given avalanche. The analytical calculations indicate that for a random network the avalanche size distribution follows [30, 31]

$$p(s) \sim s^{-3/2}. \tag{8.60}$$

For a scale-free network the distribution depends on γ as:

$$p(s) \sim \begin{cases} s^{-\gamma/(\gamma-1)} & 2 < \gamma < 3 \\ s^{-3/2(\ln s)^{-1/2}} & \gamma = 3 \\ s^{-3/2} & \gamma > 3 \end{cases} . \tag{8.61}$$

In summary, the four cascading failure models discussed in this chapter predict a critical regime where the avalanche size distribution follows a power law. A summary of the avalanche exponents obtained for these models is provided in Table 8.3.

| MODELS | $\alpha_{ER}$ | $\alpha_{SF}$ |
|---|---|---|
| Failure Propagation Model | 1.5 | |
| Overload Model | 1.5 | - |
| BTW Sandpile Model | 1.5 | γ/(γ-1) |
| Branching Process | 1.5 | γ/(γ-1) |

**Table 8.3**
**The avalanche exponents in the explored models**

The avalanche exponents for the four models supporting cascading failures. Here $\alpha_{ER}$ corresponds to the avalanche exponent if the underlying network is random; $\alpha_{SF}$ describes avalanches propagating on a scale-free network.

# ADVANCED TOPICS 8.G
# ATTACK AND ERROR TOLERANCE OF REAL NETWORKS

In this section we explore the attack and error curves for the ten reference networks discussed in Tables 4.1 and Eq. 8.2. The corresponding curves are shown in Fig. 8.33. Their inspection reveals several patterns, confirming the results discussed in this chapter:

- For all networks the error and attack curves separate, confirming the Achilles' Heel property: real networks are robust to random failures but are fragile to attacks.

- The degree of separation between the error and attach curves depends on the underlying degree heterogeneity and the average degree of each network. For example, for the citation and the actor networks we observe a very large $f_c$ for attacks, at of 0.5 and 0.75, respectively. This is mainly because these networks have an unusually large $\langle k \rangle$, with $\langle k \rangle = 20.8$ for citations and $\langle k \rangle = 83.7$ for the actor network. Their high robustness to attacks is attributed to their high link density.

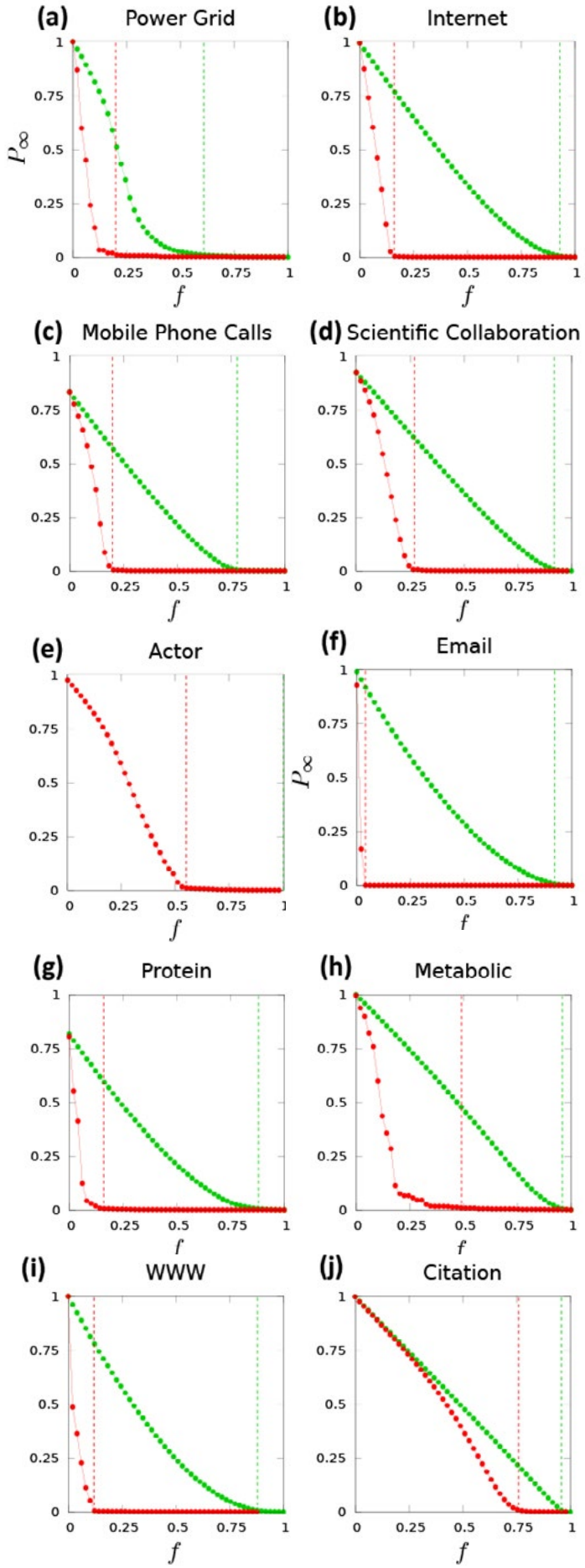



**Figure 8.32**

**Error and attack curves for reference networks**

The error (green) and attack (red) curves for the ten reference networks listed in **Table 4.1**. The green vertical line corresponds to the estimated $f_c$ for errors, while the red line corresponds to $f_c$ for attacks. We estimated $f_c$ as the point where the giant component first drops below 1% of its original size. This procedure in most systems accurately captures the point where $P_\infty$ drops to zero. The only exception is for the metabolic network, for which $f <$ 0.25, but a small cluster persists, pushing the reported $f_c$ to $f_c \approx 0.5$. Another way to detect the critical point fc is to plot the size of the second largest component in function of the fraction of deleted nodes $f$. For infinite networks $S_2$ diverges at $f_c$; for finite networks we do not observe a true phase transition, but $S_2$ has a maximum at $f_c$. Hence we can estimate the critical point $f_c$ by searching for this maximum.

# ADVANCED TOPICS 8.H THE OPTIMAL DEGREE DISTRIBUTION

The purpose of this section is to derive the bimodal distribution that simultaneously optimizes a network's topology against attacks and failures, as discussed in SECTION 8.6 [36]. Let us assume, as we did in Eq. 8.17, that the degree distribution is bimodal, consisting of two delta functions:

$$p_k = (1-r)\delta(k-k_{\min}) + r\delta(k-k_{\max}). \tag{8.62}$$

First, we calculate the total threshold, $f^{\text{tot}}$, as a function of $r$ and $k_{max}$ for a fixed $\langle k\rangle$. To obtain analytical expressions for $f_c^{rand}$ and $f_c^{\text{targ}}$, we start by calculating the moments of the bimodal distribution Eq. 8.60,

$$\begin{aligned} \langle k\rangle &= (1-r)k_{\min} + rk_{\max} \\ \langle k^2\rangle &= (1-r)k_{\min}^2 + rk_{\max}^2 = \frac{(\langle k\rangle - rk_{\max})^2}{1-r} + rk_{\max}^2 . \end{aligned} \tag{8.63}$$

Inserting these into Eq. 8.7 we obtain

$$f_c^{rand} = \frac{\langle k\rangle^2 - 2r\langle k\rangle k_{\max} - 2(1-r)\langle k\rangle + rk_{\max}}{\langle k\rangle^2 - 2r\langle k\rangle k_{\max} - (1-r)\langle k\rangle + rk_{\max}} \tag{8.64}$$

To determine the threshold for targeted attack, we must consider the fact that we have only two types of nodes, an $r$ fraction of nodes having the larger degree $k_{max}$ and the remaining $(1 - r)$ fraction has the smaller degree $k_{min}$. Hence hub removal can either remove all hubs (case (i)), or only some fraction of them (case (ii)):

**(i)** $f_c^{\text{targ}} > r$. In this case all hubs have been removed, hence all nodes left after the targeted attack have degree $k_{min}$. We therefore obtain

$$f_c^{t\,\text{arg}} = r + \frac{1-r}{\langle k\rangle - rk_{\max}}\left\{\langle k\rangle \frac{\langle k\rangle - rk_{\max} - 2(1-r)}{\langle k\rangle - rk_{\max} - (1-r)} - rk_{\max}\right\}. \tag{8.65}$$

**(ii)** $f_c^{\text{targ}} < r$. In this case the removed nodes are all from the higher degree group, leaving behind some $k_{max}$ nodes. Hence we obtain

$$f_c^{targ} = \frac{\langle k\rangle^2 - 2r\langle k\rangle k_{\max} + rk_{\max}^2 - 2(1-r)\langle k\rangle}{k_{\max}(k_{\max}-1)(1-r)}. \tag{8.66}$$

With the thresholds Eq. 8.64 - Eq. 8.66, we can now evaluate the total threshold $f_c^{\text{tot}}$ given by Eq. 8.16. We can obtain an expression for the optimal value of $k_{max}$ as a function of $r$ by determining the value of $k$ for which $f_c^{\text{tot}}$ is maximal. Using Eq. 8.64 and Eq. 8.66, we find that for small $r$ the optimal value of $k_{max}$ can be approximated by

$$k_{\max} \sim \left\{\frac{2\langle k\rangle^2(\langle k\rangle-1)^2}{2\langle k\rangle-1}\right\}^{1/3} r^{-2/3} \equiv Ar^{-2/3}. \tag{8.67}$$

Using this result and Eq. 8.14, for small $r$

$$f_c^{tot} = f_c - \frac{3\langle k\rangle}{A^2} r^{1/3} + O(r^{2/3}). \tag{8.68}$$

Thus $f$ tot $c$ approaches the theoretical maximum when $r$ approaches zero. For a network of $N$ nodes, the maximum value of $f_c^{\text{tot}}$ is obtained when $r$ = $1/N$, being the smallest value consistent with having at least one node of degree $k_{max}$. Given this $r$ the equation determining the optimal $k_{max}$, representing the size of the central hubs, is [36]

$$k_{\max} = AN^{2/3}, \tag{8.69}$$

where A is defined in Eq. 8.67.

# HOMEWORK

1. We have seen in SECTION 8.2 that the value of $p_c$ decreases with th lattice dimension: for a simple cubic lattice, representing the three dimensional version of a square lattice, we have $p_c = 0.2488$, less than half of $p_c = 1/2$ for two dimensional square lattice.

   Can you offer an intuitive explanation why does $p_c$ decrease with the lattice dimension?

# BIBLIOGRAPHY

[1] R. Albert, H. Jeong, and A.-L. Barabási. Attack and error tolerance of complex networks. Nature 406: 378, 2000.

[2] D. Stauffer and A. Aharony, Introduction to Percolation Theory. Taylor and Francis. London, 1994.

[3] A. Bunde and S. Havlin. Fractals and Disordered Systems. Springer, 1996.

[4] B. Bollobás, O. Riordan. Percolation. Cambridge University Press. Cambridge, 2006.

[5] S. Broadbent and J. Hammersley. Percolation processes I. Crystals and mazes. Proceedings of the Cambridge Philosophical Society 53: 629, 1957.

[6] M. Molloy and B. Reed. Random Structures and Algorithms 6:161, 1995.

[7] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin. Resilience of the Internet to random breakdowns. Phys. Rev. Lett. 85: 4626, 2000.

[8] D. S. Callaway, M. E. J. Newman, S. H. Strogatz. and D. J. Watts. Network robustness and fragility: Percolation on random graphs. Phys. Rev. Lett. 85: 5468–5471, 2000.

[9] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin. Breakdown of the Internet under intentional attack. Phys. Rev. Lett 86: 3682, 2001.

[10] B. Bollobás and O. Riordan. Robustness and Vulnerability of Scale-Free Random Graphs. Internet Mathematics 1:2003.

[11] D.N. Kosterev, C.W. Taylor and W.A. Mittlestadt. Model Validation of the August 10, 1996 WSCC System Outage. IEEE Transactions on Power

Systems 14:1999.

[12] C. Labovitz, A. Ahuja and F. Jahasian. Experimental Study of Internet Stability and Wide-Area Backbone Failures. Proceedings of IEEE FTCS, Madison, WI, 1999.

[13] A. G. Haldane and R. M. May. Systemic risk in banking ecosystems. Nature 469: 351-355, 2011.

[14] T. Roukny, H. Bersini, H. Pirotte, G. Caldarelli and S. Battiston. Default Cascades in Complex Networks: Topology and Systemic Risk. Scientific Reports 3: 2759, 2013.

[15] G. Tedeschi, A. Mazloumian, M. Gallegati, and D. Helbing. Bankruptcy cascades in interbank markets. PLoS One 7: e52749, 2012.

[16] D. Helbing. Globally networked risks and how to respond. Nature 497:2013.

[17] I. Dobson, B. A. Carreras, V. E. Lynch and D. E. Newman. Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. CHAOS 17: 026103, 2007.

[18] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts. Everyone's an influencer: quantifying influence on twitter. In Proceedings of thefourth ACM international conference on Web search and data mining (WSDM '11). ACM, New York, NY, USA, 65-74, 2011.

[19] Y. Y. Kagan. Accuracy of modern global earthquake catalogs. Phys. Earth Planet. Inter. 135:173, 2003.

[20] M. Nagarajan, H. Purohit, and A. P. Sheth. A Qualitative Examination of Topical Tweet and Retweet Practices. In ICWSM, 2010.

[21] P. Fleurquin, J.J. Ramasco and V.M. Eguiluz. Systemic delay propagation in the US airport network. Scientific Reports 3: 1159, 2013.

[22] B. K. Ellis, J. A. Stanford, D. Goodman, C. P. Stafford, D.L. Gustafson, D. A. Beauchamp, D. W. Chess, J. A. Craft, M. A. Deleray, and B. S. Hansen. Long-term effects of a trophic cascade in a large lake ecosystem. PNAS. 108: 1070, 2011

[23] V. R. Sole, M. M. Jose. Complexity and fragility in ecological networks. Proc. R. Soc. Lond. B 268:2039, 2001.

[24] F. Jordán and I. Scheuring. Can keystones help in background extinction? preprint, 2000.

[25] S.L. Pimm and P. Raven. Biodiversity: Extinction by numbers. Nature 403: 843, 2000.

[26] World Economic Forum, Building Resilience in Supply Chains. World Economic Forum, 2013.

[27] Joint Economic Committee of US Congress. Your flight has been delayed again: Flight delays cost passengers, airlines and the U. S. economy billions. Available at http://www.jec.senate.gov, May 22. 2008.

[28] I. Dobson, B. A. Carreras, and D. E. Newman. A loadingdependent model of probabilistic cascading failure. Probability in the Engineering and Informational Sciences 19: 15, 2005.

[29] D. Watts. A simple model of global cascades on random networks. PNAS 99: 5766, 2002.

[30] K.-I. Goh, D.-S. Lee, B. Kahng, and D. Kim. Sandpile on scale-free networks. Phys. Rev. Lett. 91:148701, 2003.

[31] D.-S. Lee, K.-I. Goh, B. Kahng, and D. Kim. Sandpile avalanche dynamics on scale-free networks. Physica A, 338: 84, 2004.

[32] M. Ding and W. Yang. Distribution of the first return time in fractional Brownian motion and its application to the study of onoff intermittency. Phys. Rev. E. 52: 207-213, 1995.

[33] A. E. Motter and Y.-C. Lai. Cascade-based attacks on complex networks. Physical Review E 66: 065102, 2002.

[34] Z. Kong and Edmund M. Yeh. Resilience to Degree-Dependent and Cascading Node Failures in Random Geometric Networks. IEEE Transactions on Information Theory 56: 5533, 2010.

[35] G. Paul, S. Sreenivas, an and H. E. Stanley. Resilience of complex networks to random breakdown. Phys. Rev. E 72, 056130, 2005.

[36] G. Paul, T. Tanizawa, S. Havlin, and H. E. Stanley. Optimization of robustness of complex networks. European Physical Journal B 38: 187–191, 2004.

[37] A.X. C. N. Valente, A. Sarkar, and H. A. Stone. Two-peak and three-peak optimal complex networks. Phys. Rev. Lett. 92: 118702, 2004.

[38] T. Tanizawa, G. Paul, R. Cohen, S. Havlin, and H. E. Stanley. Optimization of network robustness to waves of targeted and random attacks. Phys. Rev. E 71: 047101, 2005.

[39] A.E. Motter. Cascade control and defense in complex networks. Phys. Rev. Lett. 93: 098701, 2004.

[40] A. Motter, N. Gulbahce, E. Almaas, A.-L. Barabási. Predicting synthetic rescues in metabolic networks. Molecular Systems Biology 4: 1-10,

2008.

[41] R.V. Sole, M. Rosas-Casals, B. Corominas-Murtra, and S. Valverde. Robustness of the European power grids under intentional attack. Phys. Rev. E 77: 026102, 2008.

[42] R. Albert, I. Albert, and G.L. Nakarado. Structural Vulnerability of the North American Power Grid. Phys. Rev. E 69: 025103 R, 2004.

[43] C.M. Schneider, N. Yazdani, N.A.M. Araújo, S. Havlin and H.J. Herrmann. Towards designing robust coupled networks. Scientific Reports 3: 1969, 2013.

[44] A.-L. Barabási. Linked: The New Science of Networks. Plume, New York, 2002.

[45] C.M. Song, S. Havlin, H.A and Makse. Self-similarity of complex networks. Nature 433:392, 2005.

[46] S.V. Buldyrev, R. Parshani, G. Paul, H.E. Stanley and S. Havlin. Catastrophic cascade of failures in interdependent networks. Nature 464: 08932, 2010.

[47] R. Cohen, D. ben-Avraham and S. Havlin. Percolation critical exponents in scale-free networks. Phys. Rev. E 66: 036113, 2002.

[48] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. Anomalous percolation properties of growing networks. Phys. Rev. E 64: 066110, 2001.

[49] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E 64: 026118, 2001.

[50] R. Cohen and S. Havlin. Complex Networks: Structure, Robustness and Function. Cambridge University Press. Cambridge, UK, 2010. [51] P. Bak. How Nature Works: The Science of Self-Organized Criticality. New York, Copernicus, 1996.

[52] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality: an explanation of noise. Phys. Rev. Lett. 59: 381, 1987.