important consequence of nonlinear gravitational processes if the initial conditions are gaussian, and is a potentially powerful signature to exploit in statistical tests of this class of models; see Fig. 1.

The information needed to fully specify a non-gaussian field (or, in a wider context, the information needed to define an image[8]) resides in the complete set of Fourier phases. Unfortunately, relatively little is known about the behaviour of Fourier phases in the nonlinear regime of gravitational clustering[9–14], but it is essential to understand phase correlations in order to design efficient statistical tools for the analysis of clustering data. A first step on the road to a useful quantitative description of phase information is to represent it visually. We do this using colour, as shown in Fig. 2. To view the phase coupling in an N-body simulation, we Fourier-transform the density field; this produces a complex array containing the real (R) and imaginary (I) parts of the transformed 'image', with the pixels in this array labelled by wavenumber **k** rather than position **x**. The phase for each wavenumber, given by $\phi = \arctan(I/R)$, is then represented as a hue for that pixel.

The rich pattern of phase information revealed by this method (see Fig. 3) can be quantified, and related to the gravitational dynamics of its origin. For example, in our analysis of phase coupling[5] we introduced a quantity D_k:

$$D_k \equiv \phi_{k+1} - \phi_k \qquad (4)$$

This quantity measures the difference in phase of modes with neighbouring wavenumbers in one dimension. We refer to D_k as the phase gradient. To apply this idea to a two-dimensional simulation, we simply calculate gradients in the x and y directions independently. Because the difference between two circular random variables is itself a circular random variable, the distribution of D_k should initially be uniform. As the fluctuations evolve waves begin to collapse, spawning higher-frequency modes in phase with the original[15]. These then interact with other waves to produce the non-uniform distribution of D_k seen in Fig. 3.

It is necessary to develop quantitative measures of phase information that can describe the structure displayed in the colour representations. In the beginning, the phases ϕ_k are random and so are the D_k obtained from them. This corresponds to a state of minimal information, or in other words, maximum entropy. As information flows into the phases, the information content must increase and the entropy decrease. This can be quantified by defining an information entropy for the set of phase gradients[5]. We construct a frequency distribution, $f(D)$, of the values of D_k obtained from the whole map. The entropy is then defined as

$$S(D) = -\int f(D) \log[f(D)] dD \qquad (5)$$

where the integral is taken over all values of D, that is, from 0 to 2π. The use of D, rather than ϕ itself, to define entropy is one way of accounting for the lack of translation invariance of ϕ, a problem that was missed in previous attempts to quantify phase entropy[16]. A uniform distribution of D is a state of maximum entropy (minimum information), corresponding to gaussian initial conditions (random phases). This maximal value of $S_{max} = \log(2\pi)$ is a characteristic of gaussian fields. As the system evolves, it moves into states of greater information content (that is, lower entropy). The scaling of S with clustering growth displays interesting properties[5], establishing an important link between the spatial pattern and the physical processes driving clustering growth. This phase information is a unique 'fingerprint' of gravitational instability, and it therefore also furnishes statistical tests of the presence of any initial non-gaussianity[17–19]. □

1. Saunders, W. *et al.* The density field of the local Universe. *Nature* **349,** 32–38 (1991).
2. Shectman, S. *et al.* The Las Campanas redshift survey. *Astrophys. J.* **470,** 172–188 (1996).
3. Smoot, G. F. *et al.* Structure in the COBE differential microwave radiometer first-year maps. *Astrophys. J.* **396,** L1–L4 (1992).
4. Peebles, P. J. E. *The Large-scale Structure of the Universe* (Princeton Univ. Press, Princeton, 1980).
5. Chiang, L.-Y. & Coles, P. Phase information and the evolution of cosmological density perturbations. *Mon. Not R. Astron. Soc.* **311,** 809–824 (2000).
6. Guth, A. H. & Pi, S.-Y. Fluctuations in the new inflationary universe. *Phys. Rev. Lett.* **49,** 1110–1113 (1982).
7. Bardeen, J. M., Bond, J. R., Kaiser, N. & Szalay, A. S. The statistics of peaks of Gaussian random fields. *Astrophys. J.* **304,** 15–61 (1986).
8. Oppenheim, A. V. & Lim, J. S. The importance of phase in signals. *Proc. IEEE* **69,** 529–541 (1981).
9. Ryden, B. S. & Gramann, M. Phase shifts in gravitationally evolving density fields. *Astrophys. J.* **383,** L33–L36 (1991).
10. Scherrer, R. J., Melott, A. L. & Shandarin, S. F. A quantitative measure of phase correlations in density fields. *Astrophys. J.* **377,** 29–35 (1991).
11. Soda, J. & Suto, Y. Nonlinear gravitational evolution of phases and amplitudes in one-dimensional cosmological density fields. *Astrophys. J.* **396,** 379–394 (1992).
12. Jain, B. & Bertschinger, E. Self-similar evolution of gravitational clustering: is $n = 1$ special? *Astrophys. J.* **456,** 43–54 (1996).
13. Jain, B. & Bertschinger, E. Self-similar evolution of gravitational clustering: N-body simulations of the $n = -2$ spectrum. *Astrophys. J.* **509,** 517–530 (1998).
14. Thornton, A. L. *Colour Object Recognition Using a Complex Colour Representation and the Frequency Domain.* Thesis, Univ. Reading (1998).
15. Shandarin, S. F. & Zel'dovich, Ya. B. The large-scale structure: turbulence, intermittency, structures in a self-gravitating medium. *Rev. Mod. Phys.* **61,** 185–220 (1989).
16. Polygiannikis, J. M. & Moussas, X. Detection of nonlinear dynamics in solar wind and a comet using phase-correlation measures. *Sol. Phys.* **158,** 159–172 (1995).
17. Ferreira, P. G., Magueijo, J. & Górski, K. M. Evidence for non-Gaussianity in the COBE DMR 4-year sky maps. *Astrophys. J.* **503,** L1–L4 (1998).
18. Pando, J., Valls-Gabaud, D. & Fang, L. Evidence for scale-scale correlations in the cosmic microwave background radiation. *Phys. Rev. Lett.* **81,** 4568–4571 (1998).
19. Bromley, B. C. & Tegmark, M. Is the cosmic microwave background really non-gaussian? *Astrophys. J.* **524,** L79–L82 (1999).
20. Matarrese, S., Verde, L. & Heavens, A. F. Large-scale bias in the universe: bispectrum method. *Mon. Not. R. Astron. Soc.* **290,** 651–662 (1997).
21. Scoccimarro, R., Couchman, H. M. P. & Frieman, J. A. The bispectrum as a signature of gravitational instability in redshift space. *Astrophys. J.* **517,** 531–540 (1999).
22. Verde, L., Wang, L., Heavens, A. F. & Kamionkowski, M. Large-scale structure, the cosmic microwave background, and primordial non-Gaussianity. *Mon. Not. R. Astron. Soc.* **313,** 141–147 (2000).
23. Stirling, A. J. & Peacock, J. A. Power correlations in cosmology: Limits on primordial non-Gaussian density fields. *Mon. Not. R. Astron. Soc.* **283,** L99–L104 (1996).
24. Foley, J. D. & Van Dam, A. *Fundamentals of Interactive Computer Graphics* (Addison-Wesley, Reading, Massachusetts, 1982).
25. Melott, A. L. & Shandarin, S. F. Generation of large-scale cosmological structures by gravitational clustering. *Nature* **346,** 633–635 (1990).
26. Beacom, J. F., Dominik, K. G., Melott, A. L., Perkins, S. P. & Shandarin, S. F. Gravitational clustering in the expanding Universe—controlled high resolution studies in two dimensions. *Astrophys. J.* **372,** 351–363 (1991).

....

Error and attack tolerance of complex networks

Réka Albert, Hawoong Jeong & Albert-László Barabási

Department of Physics, 225 Nieuwland Science Hall, University of Notre Dame, Notre Dame, Indiana 46556, USA

....

Many complex systems display a surprising degree of tolerance against errors. For example, relatively simple organisms grow, persist and reproduce despite drastic pharmaceutical or environmental interventions, an error tolerance attributed to the robustness of the underlying metabolic network[1]. Complex communication networks[2] display a surprising degree of robustness: although key components regularly malfunction, local failures rarely lead to the loss of the global information-carrying ability of the network. The stability of these and other complex systems is often attributed to the redundant wiring of the functional web defined by the systems' components. Here we demonstrate that error tolerance is not shared by all redundant systems: it is displayed only by a class of inhomogeneously wired networks,

called scale-free networks, which include the World-Wide Web[3-5], the Internet[6], social networks[7] and cells[8]. We find that such networks display an unexpected degree of robustness, the ability of their nodes to communicate being unaffected even by unrealistically high failure rates. However, error tolerance comes at a high price in that these networks are ==extremely vulnerable to attacks== (that is, to the selection and removal of a few nodes that play a vital role in maintaining the network's connectivity). Such error tolerance and attack vulnerability are generic properties of communication networks.

The increasing availability of topological data on large networks, aided by the computerization of data acquisition, had led to great advances in our understanding of the generic aspects of network structure and development[9-16]. The existing empirical and theoretical results indicate that complex networks can be divided into two major classes based on their connectivity distribution $P(k)$, giving the probability that a node in the network is connected to k other nodes. The first class of networks is characterized by a $P(k)$ that peaks at an average $\langle k \rangle$ and decays exponentially for large k. The most investigated examples of such exponential networks are the random graph model of Erdös and Rényi[9,10] and the small-world model of Watts and Strogatz[11], both leading to a fairly homogeneous network, in which each node has approximately the same number of links, $k \simeq \langle k \rangle$. In contrast, results on the World-Wide Web (WWW)[3-5], the Internet[6] and other large networks[17-19] indicate that many systems belong to a class of inhomogeneous networks, called scale-free networks, for which $P(k)$ decays as a power-law, that is $P(k) \sim k^{-\gamma}$, free of a characteristic scale. Whereas the probability that a node has a very large number of connections ($k \gg \langle k \rangle$) is practically prohibited in exponential networks, highly connected nodes are statistically significant in scale-free networks (Fig. 1).

We start by investigating the robustness of the two basic connectivity distribution models, the Erdös–Rényi (ER) model[9,10] that produces a network with an exponential tail, and the scale-free model[17] with a power-law tail. In the ER model we first define the N nodes, and then connect each pair of nodes with probability p. This algorithm generates a homogeneous network (Fig. 1), whose connectivity follows a Poisson distribution peaked at $\langle k \rangle$ and decaying exponentially for $k \gg \langle k \rangle$.
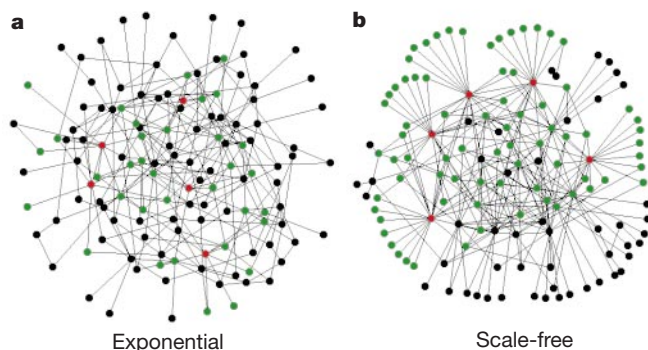
The inhomogeneous connectivity distribution of many real networks is reproduced by the scale-free model[17,18] that incorporates two ingredients common to real networks: growth and preferential attachment. The model starts with m_0 nodes. At every time step t a new node is introduced, which is connected to m of the already-existing nodes. The probability Π_i that the new node is connected to node i depends on the connectivity k_i of node i such that $\Pi_i = k_i / \Sigma_j k_j$. For large t the connectivity distribution is a power-law following $P(k) = 2m^2/k^3$.

The interconnectedness of a network is described by its diameter d, defined as the average length of the shortest paths between any two nodes in the network. The diameter characterizes the ability of two nodes to communicate with each other: the smaller d is, the shorter is the expected path between them. Networks with a very large number of nodes can have quite a small diameter; for example, the diameter of the WWW, with over 800 million nodes[20], is around 19 (ref. 3), whereas social networks with over six billion individuals
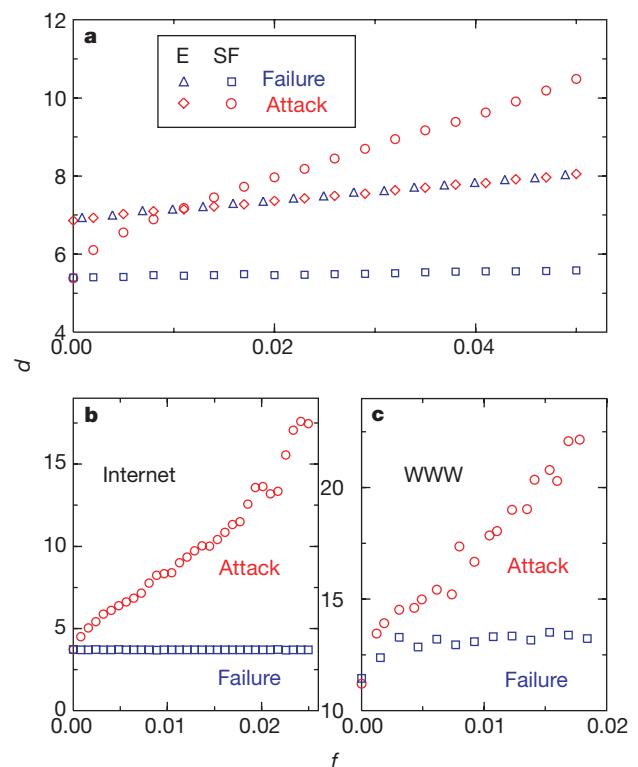

Figure 2 Changes in the diameter d of the network as a function of the fraction f of the removed nodes. **a**, Comparison between the exponential (E) and scale-free (SF) network models, each containing $N = 10,000$ nodes and 20,000 links (that is, $\langle k \rangle = 4$). The blue symbols correspond to the diameter of the exponential (triangles) and the scale-free (squares) networks when a fraction f of the nodes are removed randomly (error tolerance). Red symbols show the response of the exponential (diamonds) and the scale-free (circles) networks to attacks, when the most connected nodes are removed. We determined the f dependence of the diameter for different system sizes ($N = 1,000; 5,000; 20,000$) and found that the obtained curves, apart from a logarithmic size correction, overlap with those shown in **a**, indicating that the results are independent of the size of the system. We note that the diameter of the unperturbed ($f = 0$) scale-free network is smaller than that of the exponential network, indicating that scale-free networks use the links available to them more efficiently, generating a more interconnected web. **b**, The changes in the diameter of the Internet under random failures (squares) or attacks (circles). We used the topological map of the Internet, containing 6,209 nodes and 12,200 links ($\langle k \rangle = 3.4$), collected by the National Laboratory for Applied Network Research ⟨http://moat.nlanr.net/Routing/rawdata/⟩. **c**, Error (squares) and attack (circles) survivability of the World-Wide Web, measured on a sample containing 325,729 nodes and 1,498,353 links[3], such that $\langle k \rangle = 4.59$.

Figure 1 Visual illustration of the difference between an exponential and a scale-free network. **a**, The exponential network is homogeneous: most nodes have approximately the same number of links. **b**, The scale-free network is inhomogeneous: the majority of the nodes have one or two links but a few nodes have a large number of links, guaranteeing that the system is fully connected. Red, the five nodes with the highest number of links; green, their first neighbours. Although in the exponential network only 27% of the nodes are reached by the five most connected nodes, in the scale-free network more ==than 60% are reached, demonstrating the importance of the connected nodes in the scale-free network== Both networks contain 130 nodes and 215 links ($\langle k \rangle = 3.3$). The network visualization was done using the Pajek program for large network analysis: ⟨http://vlado.fmf.==uni-lj.si/pub==/networks/pajek/pajekman.htm⟩.

are believed to have a diameter of around six[21]. To compare the two network models properly, we generated networks that have the same number of nodes and links, such that $P(k)$ follows a Poisson distribution for the exponential network, and a power law for the scale-free network.

To address the error tolerance of the networks, we study the changes in diameter when a small fraction f of the nodes is removed. The malfunctioning (absence) of any node in general increases the distance between the remaining nodes, as it can eliminate some paths that contribute to the system's interconnectedness. Indeed, for the exponential network the diameter increases monotonically with f (Fig. 2a); thus, despite its redundant wiring (Fig. 1), it is increasingly difficult for the remaining nodes to communicate with each other. This behaviour is rooted in the homogeneity of the network: since all nodes have approximately the same number of links, they all contribute equally to the network's diameter, thus the removal of each node causes the same amount of damage. In contrast, we observe a drastically different and surprising behaviour for the scale-free network (Fig. 2a): the diameter remains unchanged under an increasing level of errors. Thus even when as many as 5% of

the nodes fail, the communication between the remaining nodes in the network is unaffected. This robustness of scale-free networks is rooted in their extremely inhomogeneous connectivity distribution: because the power-law distribution implies that the majority of nodes have only a few links, nodes with small connectivity will be selected with much higher probability. The removal of these 'small' nodes does not alter the path structure of the remaining nodes, and thus has no impact on the overall network topology.

An informed agent that attempts to deliberately damage a network will not eliminate the nodes randomly, but will preferentially target the most connected nodes. To simulate an attack we first remove the most connected node, and continue selecting and removing nodes in decreasing order of their connectivity k. Measuring the diameter of an exponential network under attack, we find that, owing to the homogeneity of the network, there is no substantial difference whether the nodes are selected randomly or in decreasing order of connectivity (Fig. 2a). On the other hand, a drastically different behaviour is observed for scale-free networks. When the most connected nodes are eliminated, the diameter of the scale-free network increases rapidly, doubling its original value if 5% of the nodes are removed. This vulnerability to attacks is rooted in the inhomogeneity of the connectivity distribution: the connectivity is maintained by a few highly connected nodes (Fig. 1b), whose removal drastically alters the network's topology, and
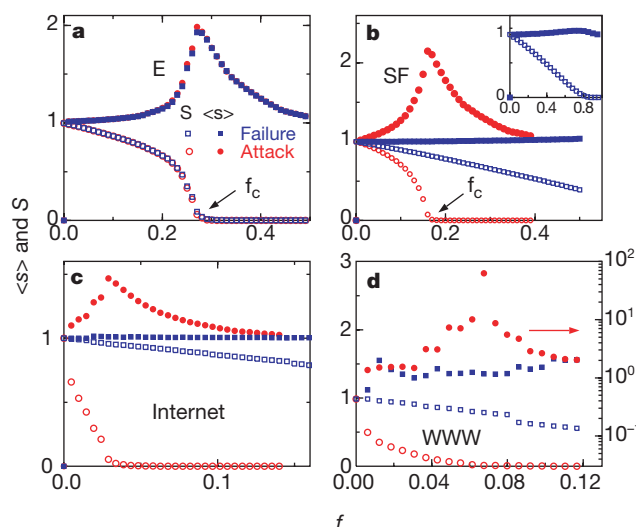

Figure 3 Network fragmentation under random failures and attacks. The relative size of the largest cluster S (open symbols) and the average size of the isolated clusters $\langle s \rangle$ (filled symbols) as a function of the fraction of removed nodes f for the same systems as in Fig. 2. The size S is defined as the fraction of nodes contained in the largest cluster (that is, $S = 1$ for $f = 0$). **a**, Fragmentation of the exponential network under random failures (squares) and attacks (circles). **b**, Fragmentation of the scale-free network under random failures (blue squares) and attacks (red circles). The inset shows the error tolerance curves for the whole range of f, indicating that the main cluster falls apart only after it has been completely deflated. We note that the behaviour of the scale-free network under errors is consistent with an extremely delayed percolation transition: at unrealistically high error rates ($f_{max} \simeq 0.75$) we do observe a very small peak in $\langle s \rangle$ ($\langle s_{max} \rangle \simeq 1.06$) even in the case of random failures, indicating the existence of a critical point. For **a** and **b** we repeated the analysis for systems of sizes $N = 1{,}000$, $5{,}000$ and $20{,}000$, finding that the obtained S and $\langle s \rangle$ curves overlap with the one shown here, indicating that the overall clustering scenario and the value of the critical point is independent of the size of the system. **c**, **d**, Fragmentation of the Internet (**c**) and WWW (**d**), using the topological data described in Fig. 2. The symbols are the same as in **b**. $\langle s \rangle$ in **d** in the case of attack is shown on a different scale, drawn in the right side of the frame. Whereas for small f we have $\langle s \rangle \simeq 1.5$, at $f_c^w = 0.067$ the average fragment size abruptly increases, peaking at $\langle s_{max} \rangle \simeq 60$, then decays rapidly. For the attack curve in **d** we ordered the nodes as a function of the number of outgoing links, k_{out}. We note that while the three studied networks, the scale-free model, the Internet and the WWW have different γ, $\langle k \rangle$ and clustering coefficient[11], their response to attacks and errors is identical. Indeed, we find that the difference between these quantities changes only f_c and the magnitude of d, S and $\langle s \rangle$, but not the nature of the response of these networks to perturbations.
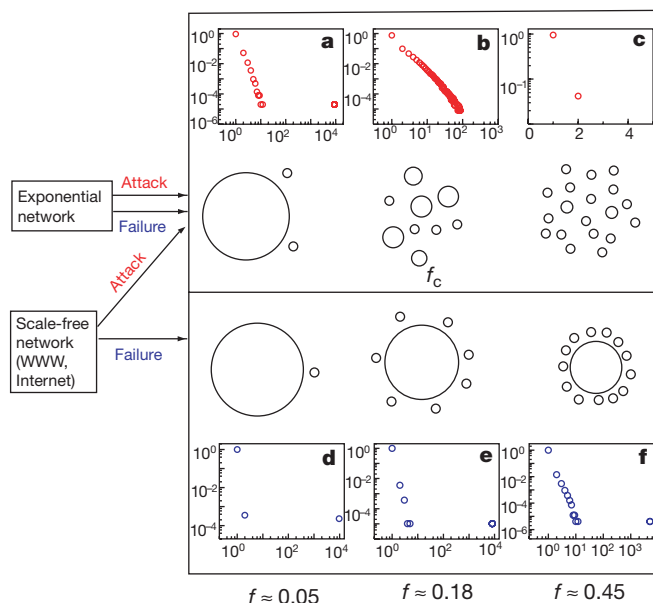

Figure 4 Summary of the response of a network to failures or attacks. **a–f**, The cluster size distribution for various values of f when a scale-free network of parameters given in Fig. 3b is subject to random failures (**a–c**) or attacks (**d–f**). Upper panels, exponential networks under random failures and attacks and scale-free networks under attacks behave similarly. For small f, clusters of different sizes break down, although there is still a large cluster. This is supported by the cluster size distribution: although we see a few fragments of sizes between 1 and 16, there is a large cluster of size 9,000 (the size of the original system being 10,000). At a critical f_c (see Fig. 3) the network breaks into small fragments between sizes 1 and 100 (**b**) and the large cluster disappears. At even higher f (**c**) the clusters are further fragmented into single nodes or clusters of size two. Lower panels, scale-free networks follow a different scenario under random failures: the size of the largest cluster decreases slowly as first single nodes, then small clusters break off. Indeed, at $f = 0.05$ only single and double nodes break off (**d**). At $f = 0.18$, the network is fragmented (**b**) under attack, but under failures the large cluster of size 8,000 coexists with isolated clusters of sizes 1 to 5 (**e**). Even for an unrealistically high error rate of $f = 0.45$ the large cluster persists, the size of the broken-off fragments not exceeding 11 (**f**).

decreases the ability of the remaining nodes to communicate with each other.

When nodes are removed from a network, clusters of nodes whose links to the system disappear may be cut off (fragmented) from the main cluster. To better understand the impact of failures and attacks on the network structure, we next investigate this fragmentation process. We measure the size of the largest cluster, S, shown as a fraction of the total system size, when a fraction f of the nodes are removed either randomly or in an attack mode. We find that for the exponential network, as we increase f, S displays a threshold-like behaviour such that for $f > f_c^e \simeq 0.28$ we have $S \simeq 0$. Similar behaviour is observed when we monitor the average size $\langle s \rangle$ of the isolated clusters (that is, all the clusters except the largest one), finding that $\langle s \rangle$ increases rapidly until $\langle s \rangle \simeq 2$ at f_c^e, after which it decreases to $\langle s \rangle = 1$. These results indicate the following breakdown scenario (Fig. 3a). For small f, only single nodes break apart, $\langle s \rangle \simeq 1$, but as f increases, the size of the fragments that fall off the main cluster increases, displaying unusual behaviour at f_c^e. At f_c^e the system falls apart; the main cluster breaks into small pieces, leading to $S \simeq 0$, and the size of the fragments, $\langle s \rangle$, peaks. As we continue to remove nodes ($f > f_c^e$), we fragment these isolated clusters, leading to a decreasing $\langle s \rangle$. Because the ER model is equivalent to infinite dimensional percolation[22], the observed threshold behaviour is qualitatively similar to the percolation critical point.

However, the response of a scale-free network to attacks and failures is rather different (Fig. 3b). For random failures no threshold for fragmentation is observed; instead, the size of the largest cluster slowly decreases. The fact that $\langle s \rangle \approx 1$ for most f values indicates that the network is deflated by nodes breaking off one by one, the increasing error level leading to the isolation of single nodes only, not clusters of nodes. Thus, in contrast with the catastrophic fragmentation of the exponential network at f_c^e, the scale-free network stays together as a large cluster for very high values of f, providing additional evidence of the topological stability of these networks under random failures. This behaviour is consistent with the existence of an extremely delayed critical point (Fig. 3) where the network falls apart only after the main cluster has been completely deflated. On the other hand, the response to attack of the scale-free network is similar (but swifter) to the response to attack and failure of the exponential network (Fig. 3b): at a critical threshold $f_c^{sf} \simeq 0.18$, smaller than the value $f_c^e \simeq 0.28$ observed for the exponential network, the system breaks apart, forming many isolated clusters (Fig. 4).

Although great efforts are being made to design error-tolerant and low-yield components for communication systems, little is known about the effect of errors and attacks on the large-scale connectivity of the network. Next, we investigate the error and attack tolerance of two networks of increasing economic and strategic importance: the Internet and the WWW.

Faloutsos et al.[6] investigated the topological properties of the Internet at the router and inter-domain level, finding that the connectivity distribution follows a power-law, $P(k) \sim k^{-2.48}$. Consequently, we expect that it should display the error tolerance and attack vulnerability predicted by our study. To test this, we used the latest survey of the Internet topology, giving the network at the inter-domain (autonomous system) level. Indeed, we find that the diameter of the Internet is unaffected by the random removal of as high as 2.5% of the nodes (an order of magnitude larger than the failure rate (0.33%) of the Internet routers[23]), whereas if the same percentage of the most connected nodes are eliminated (attack), d more than triples (Fig. 2b). Similarly, the large connected cluster persists for high rates of random node removal, but if nodes are removed in the attack mode, the size of the fragments that break off increases rapidly, the critical point appearing at $f_c^I \simeq 0.03$ (Fig. 3b).

The WWW forms a huge directed graph whose nodes are documents and edges are the URL hyperlinks that point from one document to another, its topology determining the search engines' ability to locate information on it. The WWW is also a scale-free network: the probabilities $P_{out}(k)$ and $P_{in}(k)$ that a document has k outgoing and incoming links follow a power-law over several orders of magnitude, that is, $P(k) \sim k^{-\gamma}$, with $\gamma_{in} = 2.1$ and $\gamma_{out} = 2.45$[3,4,24]. Since no complete topological map of the WWW is available, we limited our study to a subset of the web containing 325,729 nodes and 1,469,680 links ($\langle k \rangle = 4.59$) (ref. 3). Despite the directedness of the links, the response of the system is similar to the undirected networks we investigated earlier: after a slight initial increase, d remains constant in the case of random failures and increases for attacks (Fig. 2c). The network survives as a large cluster under high rates of failure, but the behaviour of $\langle s \rangle$ indicates that under attack the system abruptly falls apart at $f_c^w = 0.067$ (Fig. 3c).

In summary, we find that scale-free networks display a surprisingly high degree of tolerance against random failures, a property not shared by their exponential counterparts. This robustness is probably the basis of the error tolerance of many complex systems, ranging from cells[8] to distributed communication systems. It also explains why, despite frequent router problems[23], we rarely experience global network outages or, despite the temporary unavailability of many web pages, our ability to surf and locate information on the web is unaffected. However, the error tolerance comes at the expense of attack survivability: the diameter of these networks increases rapidly and they break into many isolated fragments when the most connected nodes are targeted. Such decreased attack survivability is useful for drug design[8], but it is less encouraging for communication systems, such as the Internet or the WWW. Although it is generally thought that attacks on networks with distributed resource management are less successful, our results indicate otherwise. The topological weaknesses of the current communication networks, rooted in their inhomogeneous connectivity distribution, seriously reduce their attack survivability. This could be exploited by those seeking to damage these systems. □

1. Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. From molecular to modular cell biology. *Nature* **402**, 47–52 (1999).
2. Claffy, K., Monk, T. E. *et al.* Internet tomography. *Nature Web Matters* [online] (7 Jan. 99) ⟨http://helix.nature.com/webmatters/tomog/tomog.html⟩ (1999).
3. Albert, R., Jeong, H. & Barabási, A.-L. Diameter of the World-Wide Web. *Nature* **401**, 130–131 (1999).
4. Kumar, R., Raghavan, P., Rajalopagan, S. & Tomkins, A. in *Proc. 9th ACM Symp. on Principles of Database Systems* 1–10 (Association for Computing Machinery, New York, 2000).
5. Huberman, B. A. & Adamic, L. A. Growth dynamics of the World-Wide Web. *Nature* **401**, 131 (1999).
6. Faloutsos, M., Faloutsos, P. & Faloutsos, C. On power-law relationships of the internet topology, ACM SIGCOMM '99. *Comput. Commun. Rev.* **29**, 251–263 (1999).
7. Wasserman, S. & Faust, K. *Social Network Analysis* (Cambridge Univ. Press, Cambridge, 1994).
8. Jeong, H., Tombor, B., Albert, R., Oltvai, Z. & Barabási, A.-L. The large-scale organization of metabolic networks. *Nature* (in the press).
9. Erdös, P. & Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17–60 (1960).
10. Bollobás, B. *Random Graphs* (Academic, London, 1985).
11. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).
12. Zegura, E. W., Calvert, K. L. & Donahoo, M. J. A quantitative comparison of graph-based models for internet topology. *IEEE/ACM Trans. Network.* **5**, 770–787 (1997).
13. Williams, R. J. & Martinez, N. D. Simple rules yield complex food webs. *Nature* **404**, 180–183 (2000).
14. Maritan, A., Colaiori, F., Flammini, A., Cieplak, M. & Banavar, J. Universality classes of optimal channel networks. *Science* **272**, 984–986 (1996).
15. Banavar, J. R., Maritan, A. & Rinaldo, A. Size and form in efficient transportation networks. *Nature* **399**, 130–132 (1999).
16. Barthélémy, M. & Amaral, L. A. N. Small-world networks: evidence for a crossover picture. *Phys. Rev. Lett.* **82**, 3180–3183 (1999).
17. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–511 (1999).
18. Barabási, A.-L., Albert, R. & Jeong, H. Mean-field theory for scale-free random networks. *Physica A* **272**, 173–187 (1999).
19. Redner, S. How popular is your paper? An empirical study of the citation distribution. *Euro. Phys. J. B* **4**, 131–134 (1998).
20. Lawrence, S. & Giles, C. L. Accessibility of information on the web. *Nature* **400**, 107–109 (1999).
21. Milgram, S. The small-world problem. *Psychol. Today* **2**, 60–67 (1967).
22. Bunde, A. & Havlin, S. (eds) *Fractals and Disordered Systems* (Springer, New York, 1996).
23. Paxson, V. End-to-end routing behavior in the internet. *IEEE/ACM Trans. Network.* **5**, 601–618 (1997).
24. Adamic, L. A. The small world web. *Lect. Notes Comput. Sci.* **1696**, 443–452 (1999).

....

Controlled surface charging as a depth-profiling probe for mesoscopic layers

Ilanit Doron-Mor*, **Anat Hatzor***, **Alexander Vaskevich***,
Tamar van der Boom-Moav†, **Abraham Shanzer**†, **Israel Rubinstein***
& Hagai Cohen‡

* *Department of Materials and Interfaces,* †*Department of Organic Chemistry,*
‡*Department of Chemical Services, Weizmann Institute of Science,*
Rehovot 76100, Israel

....

Probing the structure of material layers just a few nanometres thick requires analytical techniques with high depth sensitivity. X-ray photoelectron spectroscopy[1] (XPS) provides one such method, but obtaining vertically resolved structural information from the raw data is not straightforward. There are several XPS depth-profiling methods, including ion etching[2], angle-resolved XPS (ref. 2) and Tougaard's approach[3], but all suffer various limitations[2–5]. Here we report a simple, non-destructive XPS depth-profiling method that yields accurate depth information with nanometre resolution. We demonstrate the technique using self-assembled multilayers on gold surfaces; the former contain 'marker' monolayers that have been inserted at predetermined depths. A controllable potential gradient is established vertically through the sample by charging the surface of the dielectric overlayer with an electron flood gun. The local potential is probed by measuring XPS line shifts, which correlate directly with the vertical position of atoms. We term the method 'controlled surface charging', and expect it to be generally applicable to a large variety of mesoscopic heterostructures.

Charging is commonly considered an experimental obstacle to accurate determination of binding energies in XPS measurements of poorly conducting surfaces[6]. To compensate the extra positive charge (a natural consequence of photoelectron emission) and to stabilize the energy scale on a reasonably correct value, an electron flood gun is routinely used, creating a generally uniform potential across the studied volume. This, however, is often impossible with systems comprising components that differ in electrical conductivity[7–9]. In such cases, chemical information may be smeared due to XPS line shifts that follow local potential variations. On the other hand, this very effect can be used to gain structural information[10–12]. We have shown recently that surface charging can be used to analyse self-assembled monolayers on heterogeneous substrates, providing lateral resolution on a scale given by the substrate structural variations—that is, much smaller than the probe size[13]. Here we show that controlled surface charging (CSC) can be applied to obtain nanometre-scale depth resolution in laterally homogeneous structures.

Well-defined correlation between the local potential and the vertical (depth) axis can be achieved in some systems; for example, those consisting of an insulating overlayer on a metallic (grounded) substrate, where the overlayer is laterally uniform. We use such a system in the work reported here. In the present application the flood electron gun is not used for charge neutralization, but rather is operated at a considerably higher flux of electrons; this creates a dynamic balance between charge generation and discharge rates, which is controlled by variation of the flood-gun parameters (that is, electron flux and kinetic energy). Evidently, the extra negative charge is accumulated on the film surface, and a dynamic steady state is reached where the charge leakage to the grounded substrate is balanced by the net incoming flux. With the systems studied here, essentially no space charge is developed within the dielectric film. The resultant vertical field is therefore constant, such that the local potential, correlated directly with XPS line shifts, varies linearly with the depth scale (z), unfolding the spectrum with respect to the vertical scale. The steady-state situation, with a surface charge density σ and an overlayer dielectric constant ϵ, is thus modelled by a parallel-plate capacitor, where the local potential ϕ between the plates is given by equation (1) (σ is determined by the overlayer resistivity and the flood-gun operating conditions).

$$\phi(z) = 4\pi\sigma z/\epsilon \qquad (1)$$

The studied substrates comprised a 100-nm {111} textured gold layer, evaporated on a highly polished doped Si(111) wafer[14]. Metal−organic coordinated multilayers were constructed layer-by-layer on the gold surface as previously described[14]. The base layer was a disulphide-bishydroxamate molecule (compound **1** in Fig. 1d)[15]. Two different bifunctional ligand repeat units were used, a tetrahydroxamate (**2**, Fig. 1d)[14] and a diphosphonate (**3**, Fig. 1d)[16–18], connected by three different tetravalent metal ions: Zr^{4+}, Ce^{4+} or Hf^{4+}.

XPS measurements were performed on a Kratos Analytical AXIS-HS Instrument, using a monochromatized Al (Kα) source ($h\nu = 1{,}486.6$ eV) at a relatively low power[13]. The flood gun was used under varying conditions, of which four are presented here (see Fig. 2 legend). Accurate determination of line shifts was achieved by graphically shifting lines until a statistically optimal match was obtained. This procedure, correlating the full line shape (100−200 data points), allowed excellent accuracy ($\sim$0.03 eV) in the determination of energy shifts, much beyond the experimental energy resolution (that is, source and bare line widths). In cases where the external charging induced line-shape changes, larger uncertainties were considered. The shifts were translated to potential differences relative to the gold by subtracting the corresponding gold line shifts. Special effort was made to determine time periods characteristic of the stabilization of local potentials, which were generally longer for the lower flood-gun flux. This was of particular importance in view of beam-induced damage[5], which was observed after an hour or more and found to slightly distort the CSC behaviour. The latter will be discussed elsewhere.

For a general demonstration of CSC, two sets of multilayers with varying thicknesses (2 to 10 layers) were constructed (Fig. 1a), using the tetrahydroxamate ligand **2**. Each set was assembled with a different binding ion, that is, Zr(IV) or Ce(IV), respectively. The spectral response to the flood-gun flux is shown in Fig. 2a. The various surface elements manifest energy shifts and line-shape changes, both associated with the development of vertical potential gradients: the Au 4f line shifts by a very small amount, attributed to the finite (sample-dependent) conductivity of the silicon substrate. All other overlayer elements exhibit much larger shifts, with differences correlated with their spatial (depth) distribution. The energy shifts and line shapes of elements not shown in Fig. 2 (N, O) are fully consistent with the model.

In these two sets of samples, all the overlayer elements (except sulphur) are distributed along the vertical scale. This complicates the derivation of local potentials in general, and specifically the overall potential difference, V_0, developed across the entire overlayer. (V_0 here is relative to the 'gun-off' situation.) Yet an approximate analysis can be performed by curve-fitting the metal ion lines to sets of discrete signals, corresponding to ion interlayers at