

## Structural constraints in complex networks

S Zhou<sup>1</sup> and R J Mondragón<sup>2</sup>

<sup>1</sup> Department of Computer Science, University College London, Ross Building  
Adastral Park, Ipswich, IP5 3RE, UK

<sup>2</sup> Department of Electronic Engineering, Queen Mary, University of London  
Mile End Road, London, E1 4NS, UK  
E-mail: s.zhou@adastral.ucl.ac.uk

*New Journal of Physics* **9** (2007) 173

Received 10 February 2007

Published 28 June 2007

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/9/6/173

**Abstract.** We present a link rewiring mechanism to produce surrogates of a network where both the degree distribution and the rich-club connectivity are preserved. We consider three real networks, the autonomous system (AS)-Internet, protein interaction and scientific collaboration. We show that for a given degree distribution, the rich-club connectivity is sensitive to the degree-degree correlation, and on the other hand the degree-degree correlation is constrained by the rich-club connectivity. In particular, in the case of the Internet, the assortative coefficient is always negative and a minor change in its value can reverse the network's rich-club structure completely; while fixing the degree distribution and the rich-club connectivity restricts the assortative coefficient to such a narrow range, that a reasonable model of the Internet can be produced by considering mainly the degree distribution and the rich-club connectivity. We also comment on the suitability of using the maximal random network as a null model to assess the rich-club connectivity in real networks.

## Contents

<b>1. Introduction</b>	<b>2</b>
<b>2. Link rewiring algorithms</b>	<b>3</b>
2.1. Maximal cases I: preserving $P(k)$ . . . . .	3
2.2. Maximal cases II: preserving both $P(k)$ and $\phi(k)$ . . . . .	4
<b>3. Results</b>	<b>5</b>
<b>4. Discussion</b>	<b>8</b>
<b>5. Conclusions</b>	<b>9</b>
<b>Acknowledgments</b>	<b>9</b>
<b>References</b>	<b>10</b>

## 1. Introduction

In graph theory the degree  $k$  is defined as the number of links a node has. The distribution of degree  $P(k)$  provides a global view of a network's structure and is one of the most studied topological properties. Many complex networks are scale-free because they exhibit a power-law degree distribution, i.e.  $P(k) \sim k^{-\gamma}$ ,  $\gamma > 1$  [1]–[8]. A more complete description of a network's structure is obtained from the joint degree distribution  $P(k, k')$  [9]–[11], which is the probability that a randomly selected link connects a node of degree  $k$  with a node of degree  $k'$ . The degree distribution can be obtained from the joint degree distribution:  $P(k) = (\bar{k}/k) \sum_{k'} P(k, k')$ , where  $\bar{k}$  is the average degree.

The joint degree distribution characterizes the degree–degree correlation [12, 13] between two nodes connected by a link. But in practice  $P(k, k')$  can be difficult to measure, in particular for a finite-size and scale-free network [14]. Nevertheless the degree–degree correlation can be inferred from the average degree of the nearest neighbours of  $k$ -degree nodes [10, 15, 16], which is a projection of the joint degree distribution given by

$$k_{\text{nn}}(k) = \frac{\bar{k} \sum_{k'} k' P(k, k')}{k P(k)}. \quad (1)$$

If the nearest-neighbours average degree  $k_{\text{nn}}$  is an increasing function of  $k$  then the network is assortative, where nodes tend to attach to alike nodes, i.e. high-degree nodes to high-degree nodes and low-degree nodes to low-degree nodes. If  $k_{\text{nn}}(k)$  is a decreasing function of  $k$  then the network is disassortative, where high-degree nodes tend to connect with low-degree nodes. A network's degree–degree correlation, or mixing pattern, can also be summarized by a single scalar called the assortativity coefficient  $\alpha$ ,  $-1 \leq \alpha \leq 1$  [12]

$$\alpha = \frac{L^{-1} \sum_{j>i} k_i k_j a_{ij} - \left[ L^{-1} \sum_{j>i} 1/2(k_i + k_j) a_{ij} \right]^2}{L^{-1} \sum_{j>i} 1/2(k_i^2 + k_j^2) a_{ij} - \left[ L^{-1} \sum_{j>i} 1/2(k_i + k_j) a_{ij} \right]^2}, \quad (2)$$

where  $L$  is the total number of links,  $k_i$ ,  $k_j$  are the degrees of nodes  $i$  and  $j$ , and  $a_{ij}$  is an element of the network's adjacency matrix, where  $a_{ij} = 1$  if nodes  $i$  and  $j$  are connected by a

link otherwise  $a_{ij} = 0$  [17]. For an uncorrelated network  $\alpha = 0$ , for an assortative network  $\alpha > 0$  and for a disassortative network  $\alpha < 0$ .

In some scale-free networks the best connected nodes, rich nodes, tend to be very well connected between themselves. A rich-club is the set of nodes  $\mathcal{R}_{>k}$  with degrees larger than a given degree  $k$ . The connectivity between members of the rich-club is measure by the rich-club connectivity [18], which is defined as the ratio of the number of links  $E_{>k}$  shared by the nodes in the set  $\mathcal{R}_{>k}$  to the maximum possible number of links that the rich nodes can share

$$\phi(k) = \frac{E_{>k}}{|\mathcal{R}_{>k}| \cdot (|\mathcal{R}_{>k}| - 1)/2} = \frac{1}{|\mathcal{R}_{>k}| \cdot (|\mathcal{R}_{>k}| - 1)} \sum_{i,j \in \mathcal{R}_{>k}} a_{ij}, \quad (3)$$

where  $|\mathcal{R}_{>k}|$  is the number of nodes in the set  $\mathcal{R}_{>k}$  [17, 19]. The rich-club connectivity as a function of the degree is a global property of a network. The rich-club connectivity is a different projection of the joint degree distribution [19]

$$\phi(k) = \frac{N\bar{k} \sum_{k'=k+1}^{k_{\max}} \sum_{k''=k+1}^{k_{\max}} P(k', k'')}{\left[ N \sum_{k'=k+1}^{k_{\max}} P(k') \right] \cdot \left[ N \sum_{k'=k+1}^{k_{\max}} P(k') - 1 \right]}, \quad (4)$$

where  $N$  is the total number of nodes and  $k_{\max}$  is the maximum degree in the network. The rich-club connectivity and the degree–degree correlation are not trivially related.

Our motivation here is twofold. Firstly, to determine whether the description of a network using  $P(k)$  and  $\phi(k)$  produces a reasonable model of a real network. We consider three real networks, the autonomous system (AS)-Internet, protein interaction and scientific collaboration. Our approach is, from a real network, to create surrogate networks with the same  $P(k)$ , or even the same  $\phi(k)$ , as the original network, and then compare properties of the surrogates with the original network. Secondly, we are interested in the properties of the surrogates, in particular the maximal random case of a network, as it has been used as a ‘null model’ to assess network properties.

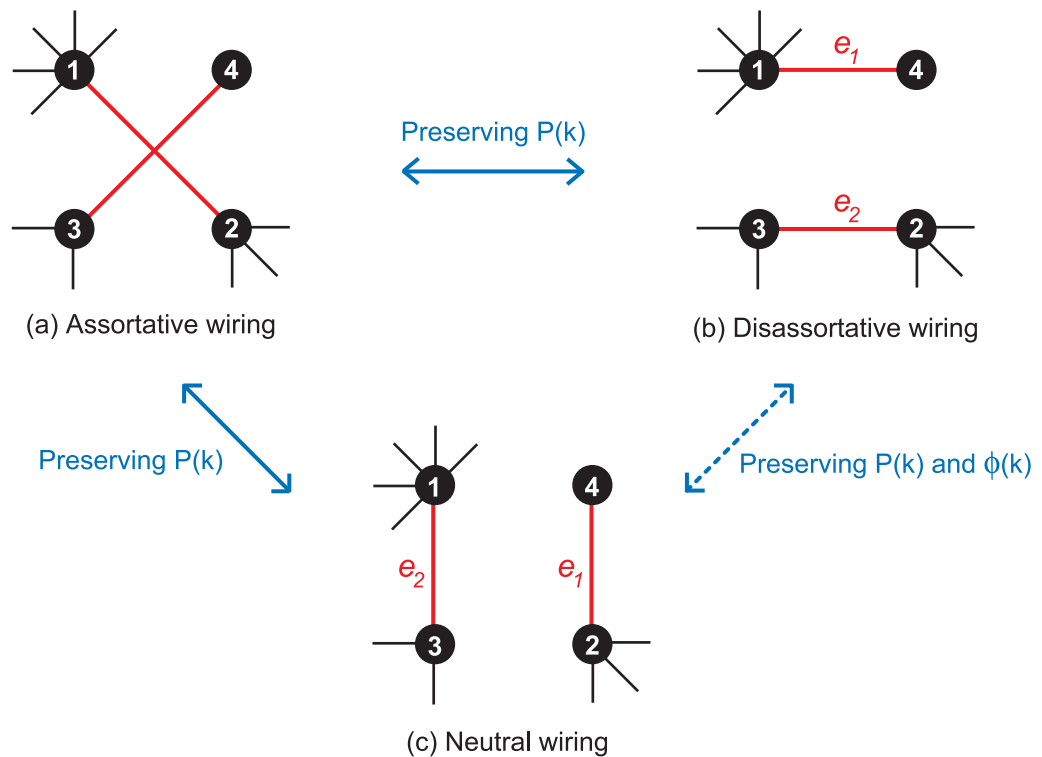
## 2. Link rewiring algorithms

We create surrogate networks by using the link rewiring algorithms [20, 21].

### 2.1. Maximal cases I: preserving $P(k)$

The broad degree distribution  $P(k)$  is an important characteristic for complex networks and it should be preserved by any link rewiring process [22]. Figure 1 shows that any four nodes with degrees  $k_1 \geq k_2 \geq k_3 \geq k_4$  can be connected by two links in three possible wiring patterns. One can see that reconnecting a pair of links from one wiring pattern to another preserves the degree of individual nodes and therefore preserves the degree distribution  $P(k)$ . It is possible to obtain different kinds of surrogate networks by rewiring links in the following ways.

- **Maximal random case I:** randomly choose a pair of links and swap two of their end nodes. This is equivalent to reconnecting the four end nodes using a wiring pattern chosen at random. The process is repeated a sufficiently large number of times.



**Figure 1.** The four end nodes of a pair of links can be reconnected in three wiring patterns: (a) assortative wiring, where one link connects the two nodes with larger degrees and the other link connects the two nodes with smaller degrees; (b) disassortative wiring, where one link connects the node with the highest degree with the node with the lowest degree and the other link connects the two remaining nodes and (c) neutral wiring.

- **Maximal assortative case I:** reconnect a pair of links in the assortative wiring pattern (see figure 1(a)) and repeat the process until all link pairs are assortatively wired.
- **Maximal disassortative case I:** similarly, reconnect all pairs of links using the disassortative wiring pattern (see figure 1(b)).

For all link rewiring algorithms in this paper, a pair of links can be rewired only if the resulting graph remains as a single connected component.

## 2.2. Maximal cases II: preserving both $P(k)$ and $\phi(k)$

It is possible to modify the link rewiring process such that the rich-club connectivity is preserved as well. For a given degree  $k$  the rich-club connectivity  $\phi(k)$  depends on the number of links  $E_{>k}$  shared by the nodes belonging to the set  $\mathcal{R}_{>k}$ . Any rewiring between nodes belonging to  $\mathcal{R}_{>k}$ , or between nodes outside  $\mathcal{R}_{>k}$ , will not change  $E_{>k}$  hence  $\phi(k)$  will remain the same. As shown in figure 1,  $E_{>k_1}$ ,  $E_{>k_2}$ ,  $E_{>k_3}$  and  $E_{>k_4}$  in disassortative wiring (figure 1(b)) and neutral wiring (figure 1(c)) are the same, because the link  $e_1$  only and always belongs to  $E_{>k_4}$ , and the other link  $e_2$  only and always belongs to  $E_{>k_3}$  and  $E_{>k_4}$ . This means that when reconnecting a pair of

links between disassortative wiring and neutral wiring,  $\phi(k)$  remains unchanged for all degrees. This allows us to obtain a different set of maximal cases for a network while preserving both the networks  $P(k)$  and  $\phi(k)$ .

- **Maximal random case II:** if a chosen pair of links are assortatively wired, they are discarded and a new pair of links is selected; otherwise the four end nodes are reconnected using either disassortative wiring or neutral wiring at random.
- **Maximal assortative case II:** if a pair of links are not assortatively wired, the four end nodes are reconnected using neutral rewiring, which will produce a more assortative mixing than using the disassortative wiring. The process is repeated for all pairs of links.
- **Maximal disassortative case II:** if a pair of links are not assortatively wired, the four end nodes are reconnected using disassortative wiring. The process is repeated for all pairs of links.

### 3. Results

Table 1 describes the data sets and some of their topological properties. Figure 2(a) shows that the three networks have a power-law decay in  $P(k)$ . The degree distribution of the Internet is well approximated by  $P(k) \sim k^{-\gamma}$ ,  $\gamma \simeq 2.24$  [23], and it exhibits a fat tail where the maximum degree,  $k_{\max} = 2070$ , is larger than the power-law natural cut-off degree  $k_{\text{cut}} = 1573$ . The degree distributions of the protein interaction and the scientific collaboration deviate from a strict power-law and have short tails. Figure 2(b) shows that the scientific collaboration exhibits assortative mixing behaviour, which is common in social networks. The Internet and protein interaction exhibit disassortative mixing behaviour, which is typical for technological and biological networks. The mixing behaviours are also confirmed by evaluating the assortative coefficient of the networks (see  $\alpha$  in table 1). Figure 2(c) shows that the three data sets exhibit different rich-club structures. Rich nodes in the disassortative Internet are significantly more tightly interconnected with each other than in the assortative scientific collaboration. Only the Internet contains a rich-club clique where the top 16 richest nodes are fully connected with each other (see  $n_{\text{clique}}$  in table 1). One can see that an assortative network does not always exhibit a strong rich-club structure, neither does a disassortative network always lack a rich-club structure. Indeed high-degree nodes have very large numbers of links and only a few of them are enough to provide connectivity to other high-degree nodes, whose number is anyway small [5].

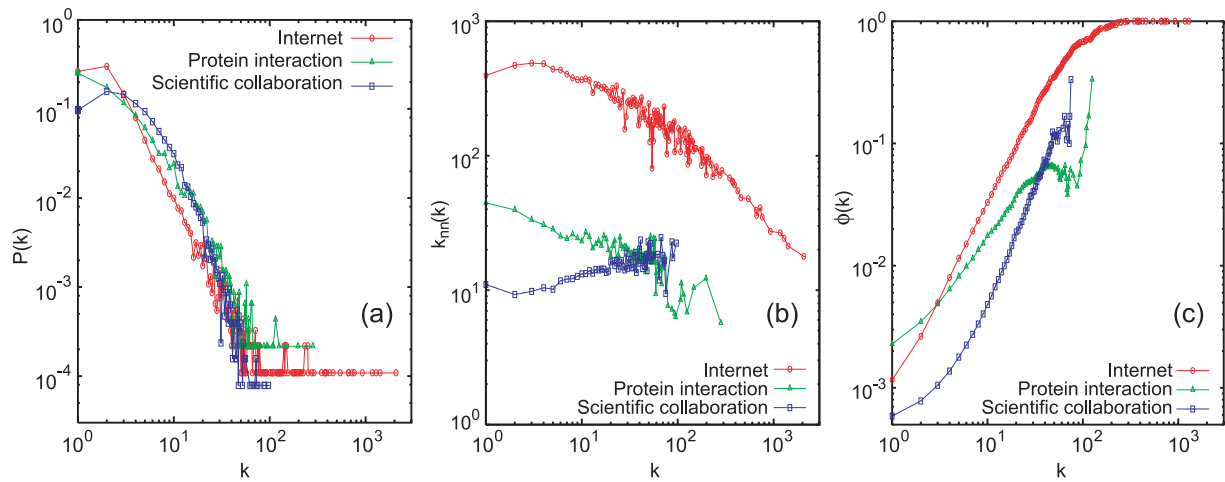
A relevant metric of a network is the average shortest path length  $\ell$  between all nodes. As shown in table 1 the average shortest path length in the Internet is significantly smaller than the average shortest path length expected in a random graph with the same network size. The Internet is so small [32] because it exhibits both a strong rich-club structure and strong disassortative mixing behaviour. While members of the rich-club are tightly interconnected with each other and they collectively function as a ‘super’ traffic hub for the network, the disassortative mixing ensures that the majority of the network nodes, peripheral low-degree nodes, are always near the rich-club core. Thus a typical shortest path between two peripheral nodes consists of three hops, the first hop is from the source node to a member of the rich-club, the second hop is between two club members and the final hop is to the destination node. One can see that a combination of the degree–degree correlation and the rich-club connectivity can also explain the distribution of the hierarchical path [33] and the short cycles [20] in a network.

**Table 1.** Three real networks considered are: (a) the Internet network at the AS level [5, 15], [23]–[27] from data collected by CAIDA [28], in which nodes represent Internet service providers and links connections among these; (b) the protein interaction network [6, 29] of the yeast *Saccharomyces cerevisiae* (<http://dip.doe-mbi.ucla.edu/>); and (c) the scientific collaboration network [30, 31], in which nodes represent scientists and a connection exists if they co-authored at least one paper in the archive. The three networks contain multiple components. In this paper, we study the giant component of the networks. We show the following properties: the number of nodes  $N$  and links  $L$  in the giant component, the average degree  $\bar{k} = 2L/N$ , the power-law exponent  $\gamma$  by fitting  $P(k)$  with  $k^{-\gamma}$  for degrees between 6 (the average degree) and 40, the maximum degree  $k_{\max}$ , the power-law natural cut-off degree  $k_{\text{cut}} = N^{1/(\gamma-1)}$  [9], the assortative coefficient  $\alpha$ , the rich-club connectivity  $\phi(k_{>40})$ , the rich-club exponent  $\theta$  obtained by fitting  $\phi(k)$  with  $k^\theta$  for degrees between 6 and 40, the size of rich-club clique  $n_{\text{clique}}$ , the average shortest path length  $\ell$ , and the average shortest path length expected in a random graph  $\ell^* = \ln N / \ln \bar{k}$  [9].

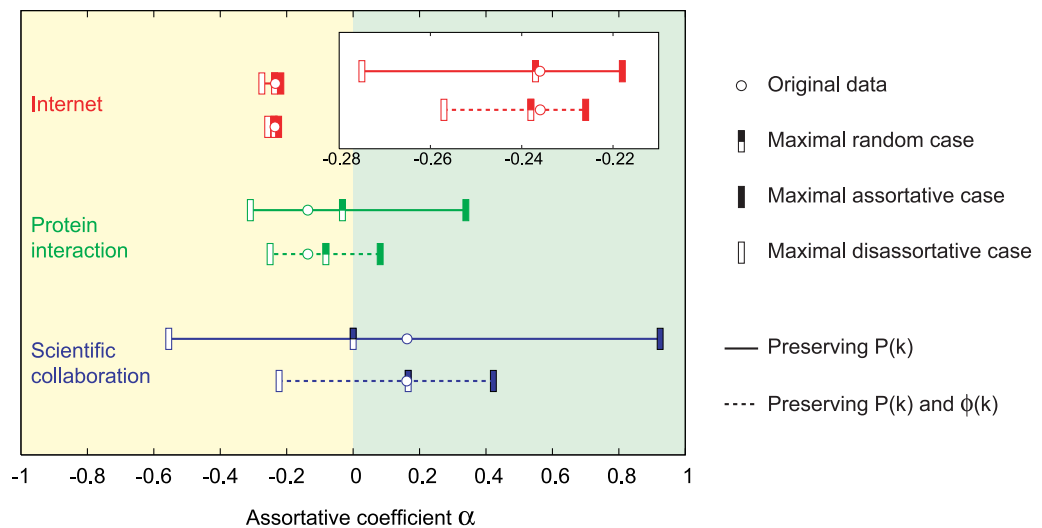
	Internet	Protein interaction	Scientific collaboration
Number of nodes $N$	9200	4626	12722
Number of links $L$	28957	14801	39967
Average degree $\bar{k}$	6.3	6.4	6.3
Power-law exponent $\gamma$	2.24	2.14	2.90
Maximum degree $k_{\max}$	2070	282	97
Power-law cut-off degree $k_{\text{cut}}$	1573	1641	145
Assortative coefficient $\alpha$	−0.236	−0.137	0.161
Rich-club connectivity $\phi(k_{>40})$	26.8%	6.4%	7.1%
Rich-club exponent $\theta$	1.52	0.97	1.94
Rich-club clique $n_{\text{clique}}$	16	0	0
Average shortest path length $\ell$	3.1	4.2	6.8
Expected in a random graph $\ell^*$	5.0	4.5	5.1

Figure 3 shows the range of the assortative coefficient  $\alpha$  obtained by the link rewiring algorithms preserving the degree distribution (case I) against that preserving both the degree distribution and the rich-club connectivity (case II). The maximal random case of a real network is averaged over 40 surrogate networks, each of which is obtained by repeating the appropriate link rewiring process for  $1000 \times L$  times, where  $L$  is the total number of links in the network.

For case I preserving  $P(k)$ , the maximal random rewiring of the protein interaction and the scientific collaboration almost decorrelates the networks, and the assortative and disassortative rewiring can produce surrogate networks in the range from assortative to disassortative. This is in contrast to the Internet, where the maximal random case is almost as disassortative as the original data. In fact all the surrogate networks produced by rewiring the Internet are disassortative, the assortative coefficient is always negative and its value is restricted to a very small range. This behaviour of the Internet is due to the restriction of having a finite network that has a power-law decay in its degree distribution and that the maximum degree is larger than the natural cut-off degree [9, 14].



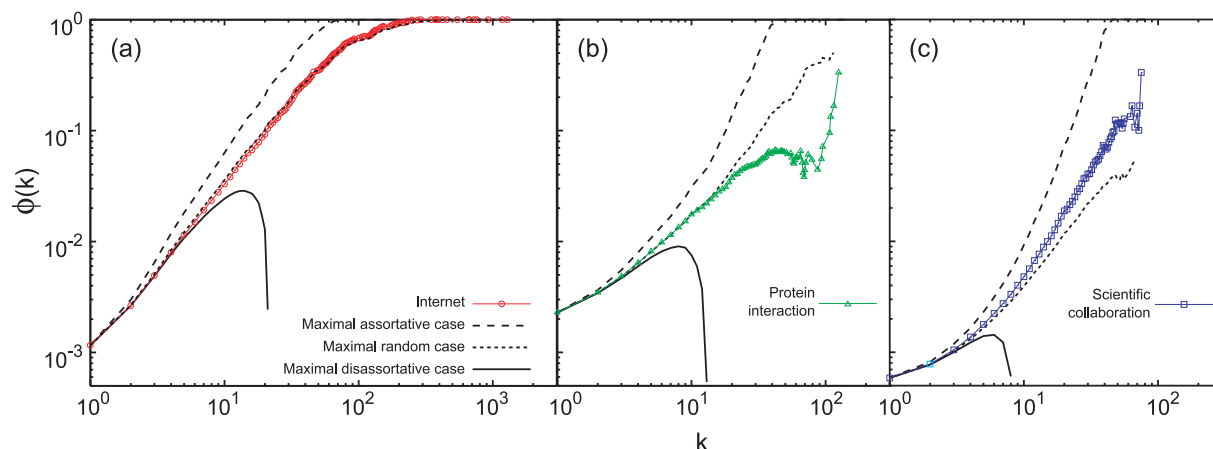
**Figure 2.** Topological properties of the Internet, the protein interaction and the scientific collaboration networks: (a) the degree distribution,  $P(k)$ ; (b) the nearest-neighbours average degree of  $k$ -degree nodes,  $k_{nn}(k)$ ; and (c) the rich-club connectivity as a function of degree,  $\phi(k)$ .



**Figure 3.** Range of the assortative coefficient  $\alpha$  of the three networks under study obtained by the link rewiring algorithms preserving  $P(k)$  (case I) comparing with that preserving both  $P(k)$  and  $\phi(k)$  (case II). The inset shows the enlargement for the Internet. The standard deviation of  $\alpha$  for a maximal rewiring case is smaller than the symbol representing it.

For case II preserving both  $P(k)$  and  $\phi(k)$ , the range of  $\alpha$  is narrower than case I when only fixing  $P(k)$ . This result confirms the analytical analysis by Krioukov and Krapivsky [34] that the rich-club connectivity constrains a network's degree–degree correlation. In the case of the Internet, the assortative coefficient is restricted to an even smaller range. This observation suggests that a reasonable model of a real network can be produced by modelling the degree





**Figure 4.** Rich-club connectivity  $\phi(k)$  of (a) the Internet, (b) the protein interaction, and (c) the scientific collaboration, comparing with the three maximal cases I obtained by preserving  $P(k)$ .

distribution and the rich-club connectivity, e.g. the positive-feedback preference (PFP) model [27, 35, 36] for the Internet.

Figure 4 shows the rich-club connectivity of the three networks each compared with their three maximal cases I obtained by preserving  $P(k)$ . The rich-club connectivity changes dramatically due to the rewiring. For all the maximal assortative networks there is a notable increase of  $\phi(k)$  throughout all degrees and all contain a fully connected rich-club clique which consists of nodes with degrees larger than 78, 48 and 46 for the Internet, the protein interaction and the scientific collaboration respectively. For all the maximal disassortative networks there is a complete collapse of the rich-club structure such that there is no single link shared between nodes with degrees larger than 23. This suggests that networks with the same degree distribution can have very different rich-club connectivity. In other words the degree distribution does not constrain the rich-club connectivity. The rich-club connectivity is sensitive to the change of a network's degree–degree correlation. For the Internet, a minor change in the assortative coefficient within the narrow range of  $\alpha \in (-0.218, -0.275)$  could reverse the rich-club structure completely. This highlights the importance of measuring the rich-club connectivity when evaluating a network model.

#### 4. Discussion

The maximal random network obtained by preserving  $P(k)$  has been used to discern whether the existence of an interaction between two proteins is due to chance or not [6]. To do this, the probability that the two nodes share a link in the protein interaction network is compared against the probability that the same two nodes will share a link in the maximal random network. The maximal random network is used as a null model because in this case it is almost a decorrelated network (see figure 3).

Recently the maximal random network has also been used as a null model to detect the origin of the rich-club connectivity in real networks [19]. The argument is that if the rich-club connectivity of the original network is the same as that of the maximal random network then the rich-club connectivity was created by chance, otherwise there was an ‘organizational principle’



responsible for the existence (or the lack) of the rich-club structure. In the case of the Internet, the original network<sup>3</sup> and the maximal random network have similar rich-club connectivity (see figure 4(a)), then the conclusion in [19, 38] was that ‘*hubs in the Internet . . . are not tightly interconnected*’ and ‘*the Internet does not have an oligarchic structure whereas, for example, scientific collaborations do*’. However, as shown in figure 2(c), the Internet does contain a well connected rich-club core and we do not need more statistical analysis to support this observation.

To understand the problem of using the maximal random network of the Internet as a null model, one needs to realize that the maximal random network in this case is not an uncorrelated network. On the contrary it is a strongly correlated network and is almost as disassortative as the Internet. Rich nodes in both the original network and the maximal random network are tightly interconnected, and the similarity between the rich-club connectivity of the two networks does not imply that the Internet lacks a rich-club structure.

Notice that the maximal random network for the Internet with  $P(k)$  fixed is more disassortative than the original network, and the maximal random network with  $P(k)$  and  $\phi(k)$  both fixed is even more disassortative (see inset in figure 3). This suggests that the rich-club structure depends strongly on the nature of the degree–degree correlation and it was not formed by chance. This strong dependence on the tail of the degree distribution ( $k_{\max}$ ) and the degree–degree correlation has also been noted in the estimates of large cliques that appear in random scale-free networks [39]. A more detailed analysis of the null model of the rich-club connectivity will be published elsewhere.

## 5. Conclusions

The rich-club connectivity and the degree–degree correlation describe the global structure of a network from different perspectives. We show that for a given degree distribution, the rich-club connectivity is sensitive to the degree–degree correlation, and on the other hand the degree–degree correlation is constrained by the rich-club connectivity. In particular for the case of the Internet, the assortative coefficient is always negative and a minor change in its value can reverse the network’s rich-club structure completely; if both the degree distribution and the rich-club connectivity are fixed, the assortative coefficient is restricted to such a narrow range that a reasonable model of the Internet can be produced by considering mainly the degree distribution and the rich-club connectivity.

We also clarify some misinterpretations that have appeared in the literature which use the maximal random case as a null model to assess the rich-club connectivity in real networks. We remark that some care is needed to avoid reaching misleading conclusions, in particular when studying the Internet.

## Acknowledgments

We thank A Vespignani and V Colizza for sharing the data sets of the protein interaction and scientific collaboration networks. We also thank CAIDA for providing the data set of the Internet AS graph. SZ is partially funded by the UK Nuffield Foundation under grant no. NAL/01125/G, and RJM is partially funded by the UK EPSRC under grant no. EP/C520246/1.

<sup>3</sup> Reference [19] used the Internet data collected by the Route Views project [37], which exhibits a similar rich-club structure to the the Internet data collected by CAIDA used in this paper.

## References

- [1] Barabási A and Albert R 1999 Emergence of scaling in random networks *Science* **286** 509
- [2] Albert R and Barabási A L 2002 Statistical mechanics of complex networks *Rev. Mod. Phys.* **74** 47–97
- [3] Newman M E J 2003 The structure and function of complex networks *SIAM Rev.* **45** 167–256
- [4] Dorogovtsev S N and Mendes J F F 2003 *Evolution of Networks—From Biological Nets to the Internet and WWW* (Oxford: Oxford University Press)
- [5] Pastor-Satorras R and Vespignani A 2004 *Evolution and Structure of the Internet—A Statistical Physics Approach* (Cambridge: Cambridge University Press)
- [6] Maslov S and Sneppen K 2002 Specificity and stability in topology of protein networks *Science* **296** 910–3
- [7] Wasserman S and Faust K 1994 *Social Network Analysis* (Cambridge: Cambridge University Press)
- [8] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D-U 2006 Complex networks: structure and dynamics *Phys. Rep.* **424** 175–308
- [9] Dorogovtsev S N and Mendes J F F 2002 Evolution of networks *Adv. Phys.* **51** 1079
- [10] Dorogovtsev S N 2003 Networks with given correlations *Preprint cond-mat/0308336*
- [11] Mahadevan P, Krioukov D, Fall K and Vahdat A 2006 Systematic topology analysis and generation using degree correlations *Proc. of SIGCOMM'06* (New York: ACM) pp 135–46
- [12] Newman M E J 2002 Assortative mixing in networks *Phys. Rev. Lett.* **89** 208701
- [13] Newman M E J 2003 Mixing patterns in networks *Phys. Rev. E* **67** 026126
- [14] Boguñá M, Pastor-Satorras R and Vespignani A 2004 Cut-offs and finite size effects in scale-free networks *Eur. Phys. J. B* **38** 205–10
- [15] Pastor-Satorras R, Vázquez A and Vespignani A 2001 Dynamical and correlation properties of the Internet *Phys. Rev. Lett.* **87** 258701
- [16] Vázquez A, Boguñá M, Moreno Y, Pastor-Satorras R and Vespignani A 2003 Topology and correlations in structured scale-free networks *Phys. Rev. E* **67** 046111
- [17] Costa L da F, Rodrigues F A, Traverso G and Villas Boas P R 2005 Characterization of complex networks: a survey of measurements *Preprint cond-mat/0505185*
- [18] Zhou S and Mondragón R J 2004 The rich-club phenomenon in the Internet topology *IEEE Commun. Lett.* **8** 180–2
- [19] Colizza V, Flammini A, Serrano M A and Vespignani A 2006 Detecting rich-club ordering in complex networks *Nat. Phys.* **2** 110–15
- [20] Maslov S, Sneppenb K and Zaliznyaka A 2004 Detection of topological patterns in complex networks: correlation profile of the Internet *Physica A* **333** 529–40
- [21] Xulvi-Brunet R and Sokolov I M 2004 Reshuffling scale-free networks: from random to assortative *Phys. Rev. E* **70** 066102
- [22] Newman M E J, Strogatz S H and Watts D J 2001 Random graphs with arbitrary degree distributions and their applications *Phys. Rev. E* **64** 026118
- [23] Faloutsos M, Faloutsos P and Faloutsos C 1999 On power-law relationships of the Internet topology *Comput. Commun. Rev.* **29** 251–62
- [24] Vázquez A, Pastor-Satorras R and Vespignani A 2002 Large-scale topological and dynamical properties of Internet *Phys. Rev. E* **65** 066130
- [25] Chen Q, Chang H, Govindan R, Jamin S, Shenker S J and Willinger W 2002 The origin of power laws in Internet topologies (Revisited) *Proc. of INFOCOM 2002* (Washington DC: IEEE Computer Society) pp 608–17
- [26] Mahadevan P, Krioukov D, Fomenkov M, Huffaker B, Dimitropoulos X, Claffy K and Vahdat A 2006 The Internet AS-level topology: three data sources and one definitive metric *Comput. Commun. Rev.* **36** 17–26
- [27] Zhou S and Mondragón R J 2004 Accurately modelling the Internet topology *Phys. Rev. E* **70** 066108
- [28] The Cooperative Association For Internet Data Analysis <http://www.caida.org/>

- [29] Colizza V, Flammini A, Maritan A and Vespignani A 2005 Characterization and modeling of protein–protein interaction networks *Physica A* **352** 1–27
- [30] Newman M E J 2001 Scientific collaboration networks. I. Network construction and fundamental results *Phys. Rev. E* **64** 016131
- [31] Newman M E J 2001 Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality *Phys. Rev. E* **64** 016132
- [32] Cohen R and Havlin S 2003 Scale-free networks are ultrasmall *Phys. Rev. Lett.* **90** 058701
- [33] Trusina A, Maslov S, Minnhagen P and Sneppen K 2004 Hierarchy measures in complex networks *Phys. Rev. Lett.* **92** 17
- [34] Krioukov D and Krapivsky P 2006 Power laws as a pre-asymptotic regime of the PFP model online at [http://www.caida.org/publications/presentations/2006/isma0605\\_dima/](http://www.caida.org/publications/presentations/2006/isma0605_dima/)
- [35] Zhou S 2006 Understanding the evolution dynamics of Internet topology *Phys. Rev. E* **74** 016124
- [36] Zhou S, Qiang Zhang G and Qing Zhang G 2005 The Chinese Internet AS-level topology *Preprint cs.NI/0511101*
- [37] Route Views Project, University of Oregon, Eugene <http://www.routeviews.org/>
- [38] Amaral L A N and Guimerà R 2006 Lies, damned lies and statistics *Nat. Phys.* **2** 75–6
- [39] Bianconi G and Marsili M 2006 Emergence of large cliques in random scale-free networks *Europhys. Lett.* **74** 740