What exactly are the properties of scale-free and other networks?

Kevin Judd, Michael Small and Thomas Stemler

School of Mathematics and Statistics, The University of Western Australia - Perth, Australia

received 8 May 2013; accepted in final form 6 September 2013
published online 30 September 2013

PACS 89.75.Fb – Structures and organization in complex systems
PACS 89.75.Da – Systems obeying scaling laws
PACS 05.10.Ln – Monte Carlo methods

Abstract – The concept of scale-free networks has been widely applied across natural and physical sciences. Many claims are made about the properties of these networks, even though the concept of scale-free is often vaguely defined. We present tools and procedures to analyse the statistical properties of networks defined by arbitrary degree distributions and other constraints. Doing so reveals the highly likely properties, and some unrecognised richness, of scale-free networks, and casts doubt on some previously claimed properties being due to a scale-free characteristic.

Introduction. – Scale-free networks are loosely defined to be networks where the number of connections per node has a power-law distribution [1]. Such a definition is problematic because firstly the specification is vague, and secondly determining whether a histogram has a power-law distribution is notoriously difficult [2–5]. Algorithms such as preferential attachment [1], duplication models [6,7], and configuration models [8], can all generate putative scale-free networks, but it is not clear that the networks are representative of typical scale-free networks [9,10]. Preferential attachment has become a de facto standard, and is often treated as synonymous with scale-free.

Scale-free clearly refers to a statistical property of networks. Imagine a process that generates random graphs with specified properties, like being scale-free with power-law \( \gamma \), having a specified number of nodes \( N \), and minimum vertex degree \( d \). Graph properties like the number of nodes, or the minimum vertex degree, can be prescribed and verified precisely, but a characteristic like being scale-free with power-law \( \gamma \) is statistical in nature and can only be prescribed probabilistically and verified on average. Moreover, it is not clear cut whether a graph is representative of a power-law \( \gamma \); there is just a probability the graph is generated for a specified \( \gamma \). The graphs of most interest are the typical graphs, which have a high probability of generation for a given \( \gamma \).

This letter describes tools and procedures to explore random graphs that are typical for specified properties. In particular, we examine scale-free graphs with prescribed size \( N \), power-law \( \gamma > 1 \), and minimum vertex degree \( d \).

With tools like ours not being previously available, an investigation of this type has not been possible: the ability to precisely control \( \gamma \) is particularly important. What we discover is an unrecognised richness of scale-free graphs as \( d \) and \( \gamma \) vary; fig. 1 illustrates two examples of typical graphs that are unlike those accessible by preferential attachment. We find that low-degree nodes can play a more significant role than previously realised; some properties previously claimed to be due to scale-free characteristics are more likely due to effects of the minimum vertex degree \( d \). We see clear structural changes with \( \gamma \): when \( d = 1 \) there are three critical power-law values (\( \gamma_0 \approx 2.47876 \), \( \gamma_1 \approx 2.18482 \), and \( \gamma_2 = 2 \)) as graphs transition from dense to tree-like forms. Our results require some elementary
calculations and a meticulous computational algorithm, which together demand clear motivation and description to be understood and implemented.

Characteristics defined by degree histograms. – Our aim is to define processes to generate random graphs with specified properties; in particular graphs of a fixed size and prescribed distributions of vertex degrees.

Let \( \mathcal{G}_N \) be the set of connected graphs with \( N \) nodes and at most one edge between any pair of nodes. Since \( \mathcal{G}_N \) is a finite set, then a process that randomly generates graphs with fixed characteristics is equivalent to assigning a probability mass \( \Pr(G) \) to each graph \( G \in \mathcal{G}_N \).

For \( G \in \mathcal{G}_N \) let \( n_k \) be the number of nodes of degree \( k \), so that \( \mathbf{n}(G) = (n_1, \ldots, n_{N-1}) \) is the histogram of the degrees of the nodes of \( G \). Let \( \mathcal{P}(\mathbf{n}) \subseteq \mathcal{G}_N \) be the set of graphs with histogram \( \mathbf{n} \). From a statistical point of view, if a characteristic depends only on \( \mathbf{n} \), then any graph \( G \in \mathcal{P}(\mathbf{n}) \) will serve as a representative of this characteristic. Hence, write the probability mass \( \Pr(G) = \Pr(G) \cdot \mathcal{P}(\mathbf{n}) \), and stipulate that \( \Pr(G) \) is the same for all graphs in \( \mathcal{P}(\mathbf{n}) \). Then, graphs within \( \mathcal{P}(\mathbf{n}) \) are equally likely, and are treated as equal representatives of the properties of \( \mathcal{P}(\mathbf{n}) \). (The actual value of \( \Pr(G) \) for \( G \in \mathcal{P}(\mathbf{n}) \) is not important to the investigation, and is in practice rarely possible to compute. In the following \( \Pr(G) \) is a normalisation that depends implicitly on other factors, like the sizes of the sets \( \mathcal{P}(\mathbf{n}) \), which are extremely difficult to compute.)

Imagine then a process that randomly selects graphs from \( \mathcal{G}_N \) such that there is a probability \( p_k \) of a node having degree \( k \) for each \( k = 1, \ldots, N-1 \). For graphs generated this way \( \mathbf{n}(G) \) has a multinomial distribution,

\[
\Pr(\mathbf{n}(G)) = N! \prod_{k=1}^{N-1} \frac{p_k^{n_k}}{n_k!}
\]

(1)

The methodology that follows can be applied to arbitrary degree distributions, and used to explore any statistical property that is determined by a degree distribution alone.

A natural choice to obtain scale-free graphs for \( \gamma > 1 \) is a (shifted) zeta distribution, where \( p_k = 0 \) for \( k < d \), and \( p_k = \Pr(k|\gamma, d) = (k-d+1)^{-\gamma}/\zeta(\gamma) \),

(2)

for \( k \geq d \), where \( \zeta(\gamma) = \sum_{k=1}^{\infty} k^{-\gamma} \).

Our specification of \( \Pr(G) \) above encapsulates the characteristics of size \( N \), scale-free power-law \( \gamma \), and minimum degree \( d \). In the next section we develop a process to randomly select graphs according to \( \Pr(G) \). Moreover, the following method can provide an unbiased sample of all graphs with a prescribed distribution \( \Pr(G) \).

It is relevant to note that configuration models [8] are not an example of the required process for three reasons: 1) they do not prescribe a multinomial distribution of \( \mathbf{n} \); 2) although capable of generating graphs with specified \( \mathbf{n} \) they do not restrict to \( \mathcal{G}_N \), as graphs can have self-loops and multiple edges, or be disconnected; 3) they do not ensure graphs within \( \mathcal{P}(\mathbf{n}) \) are equally likely. Modification to achieve requirement 3) is non-trivial.

Generating random graphs. – The question now is how to randomly generate scale-free graphs as we define. In this section we provide the necessary foundation to understand our results and algorithm, which we describe in detail later. A Markov-chain Monte Carlo (MCMC) approach [12,13] can be used to sample \( \Pr(G) \). Starting with an arbitrary initial graph \( G \in \mathcal{G}_N \) one systematically proposes random simple modifications of the graph \( G \) to produce a different graph \( G' \in \mathcal{G}_N \). The quantity

\[
Q(G'|G) = \frac{\Pr(G')}{\Pr(G)} = \prod_{k=1}^{N-1} \frac{p_k^{n_k}}{n_k!} \frac{n_k!}{p_k^{s_k}}
\]

(3)

measures the relative likelihood of the graphs. An MCMC approach accepts \( G' \) as a replacement for \( G \) if \( Q \geq 1 \) and if \( Q < 1 \), then \( G' \) is accepted with probability \( Q \). Asymptotically, an unbiased sample of random graphs with the specified distribution \( \Pr(G) \) are obtained.

The most basic modifications are adding and deleting a single edge. For example, if \( G' \) is obtained from \( G \) by adding an edge between a node of degree \( k \) and another node of degree \( l \), where \( |k-l| > 1 \), then \( n_k \) and \( n_l \) decrease by one, and \( n_{k+1} \) and \( n_{l+1} \) increase by one. Using (3) it can be easily derived that adding an edge \( (s = 1) \), or deleting an edge \( (s = -1) \), between a node of degree \( k \) and another node of degree \( l \), results in

\[
Q = \begin{cases} 
\frac{p_k p_{k+1}}{p_k p_{k+1}} & |k-l| > 1, \\
\frac{n_k}{p_k} & k = l + s, \\
\frac{n_l}{p_l} & k = l - s, \\
\frac{n_k (n_k - 1)}{p_k^2} & k = l, n_k \geq 2.
\end{cases}
\]

(4)

For scale-free graphs substitute (2) into (4), and note that the zeta functions cancel out entirely.

Adding and deleting edges is transitive in \( \mathcal{G}_N \), but another useful modification that accelerates convergence is to disconnect one or more edges from a node, and reconnect these edges to other nodes. Consider the simplest case where a giver node has degree \( k \), that \( m < k \) edges are disconnected and these are all reconnected to a receiver node of degree \( l \); we will call such modifications gifting. The degrees of the giver and receiver nodes decrease and increase respectively by \( m \); it follows that since necessarily \( n_k, n_l \geq 1 \), then

\[
Q = \begin{cases} 
\frac{n_k p_l}{p_k (n_k - 1)} & k = l \neq 0, m, 2m, \\
\frac{n_k (n_k - 1)}{p_k (n_k + 1)} & k = l, n_k \geq 2, \\
1 & k = l + m,
\end{cases}
\]

(5)

Substitution of (2) gives \( Q \) for scale-free graphs.
What exactly are the properties of scale-free and other networks?

Care is required making any modification to a graph because it can result in graphs that are not in $G_N$, or fail other required constraints, like minimum vertex degree. Given the assumed properties of $G_N$, an edge can only be added if the nodes are not already connected, and deletion of an existing edge is only allowed if the resulting graph is connected. Gifting can also result in disconnected graphs. Ensuring the graph is connected is an involved process, and somewhat tangential to our narrative, so to avoid disrupting our focus we will deal with this detail later when describing the algorithm used in our calculations.

Properties of scale-free graphs. – We have introduced some basic methods of modifying graphs in $G_N$ to obtain different graphs in $G_N$. These can be used in an MCMC approach, but an MCMC approach can be slow to converge, especially if the arbitrary initial graph is far from being scale-free. An alternative approach is to seek highly likely graphs by maximising $Pr(G)$. This can be achieved by considering various modifications of $G$ and selecting the modification that has the largest $Q$. Our algorithm to implement gradient ascent of $Q$, which is used in our calculations, is described and discussed later.

So then, we come to answer the question in the title. Figure 2 shows highly likely scale-free graphs obtained for $d = 1$ using the gradient ascent algorithm stated later; these graphs are typical representatives. We find there is a significant subset of scale-free graphs that are trees, especially for larger $\gamma$ values. For intermediate $\gamma$ scale-free graphs are similar to scale-free trees with some extra edges linking nodes, which break the tree structure by creating loops. For smaller $\gamma$ values the proportion of edges creating loops increases.

This tree-like nature deserves closer attention, because it has been noted elsewhere [14]. If $G \in G_N$, then a spanning tree $S(G)$ of $G$ has exactly $N - 1$ edges, so the number of edges $E(G)$ of $G$ in excess of $E(S(G)) = N - 1$ is an indication of the amount of cross-linking. For a graph with vertex-degree probabilities $p_k$, the expected number of edges is the sum over $k$ of $kp_k$. For the distribution (2) with $d = 1$ this sum evaluates to $\zeta(\gamma - 1)/\zeta(\gamma)$. Hence, the expected excess $X_s(G) = E(G) - E(S(G))$ for scale-free graphs in $G_N$ with $\gamma > 2$, $d = 1$, is

$$E[X_s(G)] = N \times \left( \frac{1}{2} \frac{\zeta(\gamma - 1)}{\zeta(\gamma)} - 1 \right) + 1. \quad (6)$$

The quantity in brackets is decreasing, diverging to positive infinity as $\gamma$ approaches 2 from above, negative for $\gamma > 70 \approx 2.47876$, and equal to one at $\gamma = \gamma_1 \approx 2.18482$. Excesses are stated for the graphs in fig. 2.

The expected excess (6) implies that for $\gamma > 70$ scale-free graphs are always expected to be trees. In fact, there should be a deficit of edges, which forces scale-free graphs toward the extreme of one node of very large degree and many nodes of degree one, like fig. 2(d). For $\gamma_1 < \gamma < 70$ scale-free graphs are expected to have less cross-links than links in a spanning tree, and so are tree-like. For $\gamma < \gamma_1$ there are expected to be more cross-links than spanning-tree links, so the graphs are not expected to be tree-like, and are frequently observed in real-world networks [14]. For $\gamma > 2$ the excess is expected to be a bounded multiple of the number of nodes $N$. When $\gamma < 2$ the expected number of links grows without bound as $N$ increases until nearly fully connected graphs are produced. These far from tree-like graphs have been termed dense [15], but, contrary to the claim of the cited work, there are many such dense graphs for any $N$ and $1 < \gamma < 2$; our supplied algorithm finds them easily, because it searches within the set of graphs $G_N$, rather than by histogram.

We conclude that random scale-free graphs with $d = 1$ and $\gamma_1 < \gamma < 70$ are best characterized as having an underlying, more-or-less scale-free, tree, and are expected to

---

Table 1: Typical behaviour of gradient ascent of $Q$ for a graph of $N = 300$ nodes. Columns are: $E_0$ and $E_\ast$, initial and final number of edges; $\log P$, final log-likelihood; number of additions, deletions, and gifting moves; total moves and predicted lower and upper bounds on number of moves. Recall that the excess edges over a tree is $X_s = E_\ast - N + 1$.

| $\gamma$ | $E_0$ | $E_\ast$ | $\log P$ | Add | Del | Gift | Total | Lower | Upper |
|----------|-------|----------|----------|-----|-----|------|-------|-------|-------|
| $\gamma = 1.5$ | 300 452 | -87.41 | 183 | 1 | 61 | 245 | 43 | 558 |
| $\gamma = 2.0$ | 500 504 | -86.97 | 36 | 32 | 82 | 150 | 129 | 970 |
| $\gamma = 2.5$ | 700 626 | -87.75 | 20 | 94 | 117 | 231 | 164 | 1526 |

Fig. 2: Random scale-free graphs with $N = 300$, $d = 1$, $\gamma$ as stated, using the algorithm described in text. $X_s$ is the number of edges in excess of a spanning tree. Graph layout by neato [11].
be scale-free for \( \gamma > \gamma_0 \), with nodes of very large degree expected for larger \( \gamma \). These characteristic structures are well illustrated in fig. 2. Furthermore, the characteristics are confirmed by observing the relative frequency of the different modifications during the likelihood ascent. Table 1 compiles a summary of the typical behaviour of likelihood ascent. For \( \gamma = 1.5 \), the non-tree-like case with many cross-links, the likelihood ascent is mainly adding and gifting edges. For \( \gamma = 2.5 \), the tree-like case with few cross-links, the likelihood ascent is mainly deleting and gifting edges.

Generating large graphs. — We have introduced means to obtain increasingly scale-free graphs by modification of an initial graph. Given our conclusion that scale-free graphs for \( \gamma > \gamma_1 \) are essentially tree-like, this suggests an alternative constructive approach for building random scale-free graphs by adding individual nodes with one link.

If a graph \( G \in \mathcal{G}_N \) is modified by adding a node with one link to an existing node of degree \( k \), then a new graph \( G' \in \mathcal{G}_{N+1} \) is created with

\[
Q = \begin{cases} 
\binom{N+1}{k+1} \frac{p_k - 1}{p_k}, & k > 1, \\
\frac{p_k - 1}{p_k}, & k = 1,
\end{cases}
\]

which follows from eq. (1) similar to (3). Hence, eq. (7) provides an optimal preferential attachment rule, however, attachment rules alone need not result in highly likely graphs [9]. Locally optimum graphs can be obtained only if after one or more attachments the link modifications previously described are used.

Figure 3 shows the growth of excess links \( X_s(G) \), for \( G \in \mathcal{G}_N \) constructed in this way with eq. (7) used as the attachment rule. The staircase shape results from occasional avalanches of link modifications after a period of mainly node-link additions with few modifications. There appears to be a self-organised criticality.

For \( 1 < \gamma < 2 \) (as illustrated in fig. 3) the excess should grow without bound until the graph is almost fully connected, but fig. 3 clearly shows that the number of nodes needs to be many orders of magnitude larger before this effect is apparent; only \( \gamma = 1.5 \) has more cross-links than spanning-tree links for \( N > 600 \) nodes. While \( \gamma < 2 \) is often considered as a pathological case, our calculations to obtain fig. 3 illustrate that such networks are easily created; indeed networks with \( \gamma < 2 \) have been observed in real-world data [16]. For \( 2 < \gamma < \gamma_0 \) the growth of \( X_s(G) \) should become asymptotically linear in \( N \), that is, a constant proportion of cross-links. Once again, graphs many orders of magnitude larger are required before this effect is apparent.

In fig. 4 two frequently applied measures of network properties (clustering and assortativity [17]) are used to compare a few small real networks that have been claimed to be scale free, networks generated by standard Barabási-Albert preferential attachment with \( d = 1, 2 \), and random unbiased samples of scale-free networks conforming to (2) with \( d = 1 \), generated using (7) and the following algorithm. We observe the real networks and random networks have a wider range of values, indicating that preferential attachment creates a biased subset of the potential scale-free networks. Elsewhere we present extensive...
computations that reveal the nature of the bias and the causes of the extreme values of some real networks [18].

Algorithm. – In this, and the next, section we state and then describe the computational algorithm we use to support our results. This algorithm sequentially modifies the links of a given graph to obtain another graph with the same number of nodes but higher likelihood of being scale free. Links are modified by the three operations of adding, deleting, and gifting. The algorithm requires only the number of nodes \( n \) and the target degree distribution, which in the case of (2) is defined by the single power-law parameter \( \gamma \).

Steps 3) and 4) treat the current graph \( G \) as a representative of an equivalence class of graphs having node-degree histogram \( \mathbf{n}(G) \); these two steps determine which equivalence classes are accessible from \( G \) using the allowed modifications, and which achieve the largest increase in \( Q \).

Steps 5) and 6) determine if there is a specific valid modification of \( G \) that can make the potentially best transition to another equivalence class.

The following notation is used: \( G \in \mathcal{G}_N \) is the current graph with nodes numbered 1 to \( N \); \( A \) is the adjacency matrix of \( G \); \( A_{ij} = 1 \) if node \( i \) is linked to \( j \), zero otherwise; \( a_i \) is column \( i \) of \( A \); \( I_k \) is the set of nodes with degree \( k \), by definition \( n_k = |I_k| \) and \( i \in I_k \) if \( \|a_i\|_1 = k \); \( N_i \) = \{ \( j : A_{ij} = 1 \) \}, the nodes linked to node \( i \); \( C(i,j,\mathcal{L}) = 1 \) if there exists a path between node \( i \) and \( j \) after the links in \( \mathcal{L} \) have been deleted, and zero otherwise.

1) Choose an arbitrary initial graph \( G \in \mathcal{G}_N \).

2) Compute the node-degree histogram \( \mathbf{n}(G) \) and log-likelihood \( \log \Pr(\mathbf{n}(G)) \).

3) Choose a degree \( k \) such that \( n_k > 0 \); chosen uniformly amongst non-zero \( n_k \).

4) Compute log \( Q_{klm} \) for all potentially valid \( k, l, m \), that is all possible changes involving a node of degree \( k \) and another of degree \( l \), where \( m \) identifies the type of change: (\( m = A \)) addition of a link, eq. (4); (\( m = D \)) deletion of a link, eq. (4); (\( m \in \mathbb{Z}^+ \)) gift \( m \) links from a node of degree \( k \) to a node of degree \( l \), eq. (5).

5) Check the list of log \( Q_{klm} > 0 \) in descending order of magnitude for the first valid change, if none, return to step 3). Validity is tested as follows. Find candidate nodes \( i \in I_k, j \in I_l \), for the change as follows:

\( (A) \quad A_{ij} = 0; \)
\( (D) \quad A_{ij} = 1 \) and \( C(i,j,\{(i,j)\}) = 1; \)
\( (m) \quad a_i^T a_j - A_{ij} \leq n_k - m. \)

If there is more than one pair of \( (i,j) \) then choose uniformly randomly between them. This is achieved by choosing uniformly random permutations [19] of the elements of \( I_k \) and \( I_l \), then running over \( i \) in permutation order with an embedded loop over \( j \) in permutation order until the first valid pair is found.

6) Make the valid change to \( G \) and return to step 2). For addition \( (c = 1) \) and deletion \( (c = 0) \), set \( A_{ij} = A_{ji} = c \). Step 5)(\( m \)) ensures that \( m \) nodes linked to node \( i \) can be moved to node \( j \); however, some choices of the \( m \) nodes can result in disconnected graphs. Proceed as follows:

a) If \( A_{ij} = 0 \) and \( a_i^T a_j = 0 \), then choose uniformly at random \( b \) from \( N_i \) where \( C(j,h,\{(i,g) : g \in N_i\}) = 1 \). Otherwise, set \( h = j \).

b) Choose \( m \) nodes uniformly randomly from \( \{ g \in N_i : g \notin N_j, g \neq h \} \). Relink these nodes by setting \( A_{ig} = A_{gi} = 0 \) and \( A_{gh} = A_{hg} = 1 \).

The algorithm employs the test \( C(i,j,\mathcal{L}) \) for a graph \( G \) as to whether node \( i \) is connected to node \( j \) when the links in \( \mathcal{L} \) are removed. The purpose of this test is to ensure that a graph remains connected after some change to the graph. A graph \( G \) is connected if the matrix \( C_p = I + A + A^2 + \ldots + A^p \) has no zero elements for \( p = N - 1 \). A pair of nodes \( i \) and \( j \) is connected by a path if there exists \( p < N \) such that \( C_{pq} \neq 0 \). If \( a_i \) and \( a_j \) are columns of \( A \), respectively, \( w_p+1 = A w_p + w_p \), and \( w_1 = a_j \), then \( C_{pq} \approx a_j^T w_p \), which is a relatively efficient computation for each \( p \). Hence, \( C(i,j,\mathcal{L}) \) requires computing \( C_{pq} \) where \( A \) is modified so that \( A_{uv} = A_{vu} = 0 \) for each \( (u,v) \in \mathcal{L} \).

The algorithm should terminate at a local optimum of \( \log \Pr(\mathbf{n}) \) when either \( \log Q_{klm} \) is non-positive for all \( klm \), or there are no valid changes with \( \log Q_{klm} > 0 \) for any \( klm \). However, step 3) makes a random choice of \( k \), so as stated the algorithm will not terminate and requires an additional stopping criterion. For example, if more than a prescribed number of trials of \( k \) result in no valid changes, then test each \( k \) with \( n_k > 0 \) sequentially, if none results in valid changes then a local optimum of \( \log \Pr(\mathbf{n}) \) has been reached.

Algorithm performance. – We provide here some comments on the performance of the algorithm. The histogram \( \hat{\mathbf{n}} \) that maximizes \( \Pr(\mathbf{n}) \) for the multinomial distribution (1), has \( \hat{n}_k \approx N p_k \). Substituting into (3) and using Stirling’s approximation of a factorial, obtains

\[
Q(\hat{G}|G) \approx \prod_{k=1}^{N-1} \left( \frac{n_k}{\hat{n}_k} \right)^{n_k} \frac{1}{n_k^\frac{1}{2}}.
\]

This approximation assumes \( n_k \neq \hat{n}_k \) only if \( n_k \gg 1 \). The expressions (4), (5), and (7) for \( Q \) were the result of small changes to graphs, but by (8) it can be seen that modest deviations from \( n_k \approx \hat{n}_k \) can result in significant change of probability mass. Hence, most graphs in \( \mathcal{G}_N \) do not display the scale-free property. On the other hand, there are a lot of scale-free graphs [10,20,21].

Of importance to gradient ascent of \( Q \) is an estimate of the number of moves required to reach a high-likelihood scale-free graph \( \hat{G} \) with \( \mathbf{n}(\hat{G}) = \hat{\mathbf{n}} \) from an initial graph \( G \).
A lower bound on the number of moves required is around \[ \frac{1}{2} \sum_{k=1}^{N-1} |n_k - \tilde{n}_k| \]; here each move corrects the degrees of a pair of nodes. The worst case upper bound is around \[ \frac{1}{2} \sum_{k=1}^{N-1} k|n_k - \tilde{n}_k| \], where a move is required to relocate each edge end. The lower bound should be tight, the upper bound is expected to be typically a significant over-estimate. These upper and lower bounds are included in Table 1; it can be seen that in all cases the algorithm is closer to the lower bound than the upper bound.

Table 1 also shows that when the initial graph has minimal edges, then near optimal graphs are obtained, but when the initial graph has much too many edges, then the algorithm often gets caught in good, but sub-optimal, graphs. This suggests building a graph up is more effective than reducing a graph down. In the latter case some initial deletions can leave the graph in a configuration that cannot be easily corrected; note the large number of additions and deletions in these cases in Table 1, which results from frequent readjustment from earlier deletions. The optimal starting condition appears to have as many edges as the optimal graph, which can be predicted in advance; this minimises additions and deletions. Another disadvantage of initial graphs with too many edges, not revealed by Table 1, is that they require many more evaluations of potential \( Q \) values, meaning individual moves take longer.

Conclusion. – The algorithm we have presented provides for the unbiased sampling of random graphs of finite size with a prescribed degree distribution. Using this algorithm we find that many properties of scale-free graphs that have previously been supposed to be generic are actually due to additional constraints. In particular, we conclude that scale-free graphs, with \( d = 1 \), are not particularly robust to deletions for \( \gamma > \gamma_1 \); being based on an underlying tree, any robustness derives from the cross-linking, which is more prevalent for smaller \( \gamma \) values, but expected to be entirely absent for \( \gamma > \gamma_0 \). When a graph is a tree, deletion of any non-leaf node will result in destruction of the giant component. Provided one chooses non-leaf nodes, the process will proceed rapidly with either targeted or random node selection. Hence, for such graphs the robustness will be due to the presence of a large number of leaf nodes, rather than being due to the hubs. This means that our understanding of some qualities, like robustness and clustering, needs to be examined more closely. Previously, such properties have been claimed to be due to the scale-free property of graphs. Our evidence suggests that additional, often implicit, assumptions and constraints, such as, a minimum degree of nodes \( d > 1 \) are equally important. Some earlier work has hinted at similar conclusions. Chung and Lu [22] studied the configuration model and showed that a transition point arose beyond which the network loses its giant component. In turn, that transition point depends critically on minimum degree \( d \).

This strong effect of \( d \) highlights a potential mis-conception. Usually when power-laws arise (phase transitions, Hirsch exponent, fractal dimension, extreme events, self-organized criticality) the asymptotics of the distribution dominates the interesting physics. Here, it seems, the properties of the most numerous nodes are important, too, that is, the shape of the left side of the degree histogram, rather than the tail on the right. These effects can be explored using the methods we have presented. It should be possible to make further characterizations of random graphs under different probability constraints on the degrees of nodes. The methodology presented here can be adapted to cases of \( d > 1 \); Fig. 1(b) was obtained by applying the described algorithm with a shifted zeta distribution (2) with \( d = 2 \). The three basic graph modifications of addition, deletion and gifting, were sufficient for \( d = 1,2 \), but it is likely that different modifications will be required for efficient algorithms for \( d > 2 \).

REFERENCES

[1] Barabási A. and Albert R., Science, 286 (1999) 509.
[2] Clauset A., Shalizi C. and Newman M., SIAM Rev., 51 (2009) 661.
[3] Goldstein M., Morris S. and Yen G., Eur. Phys. J. B, 41 (2004) 255.
[4] Stumpf M. and Porter M., Science, 335 (2012) 665.
[5] Judd K., Physica D, 56 (1992) 216.
[6] Pastor-Satorras R., Smith E. and Solé R., J. Theor. Biol., 222 (2003) 199.
[7] Chung F., Lu L., Dewey T. and Galas D., J. Comput. Biol., 10 (2003) 677.
[8] Molloy M. and Reed B., Random Struct. Algorithms, 6 (1995) 161.
[9] Callaway D., Hopcroft J., Klenberg J., Newman M. and Strogatz S., Phys. Rev. E, 64 (2001) 041902.
[10] Catanzaro M., Boguñá M. and Pastor-Satorras R., Phys. Rev. E, 71 (2005) 027103.
[11] North S., Technical report. www.graphviz.org (2004).
[12] Fitzgerald R., Numerical Bayesian Methods Applied to Signal Processing (Springer) 1997.
[13] Gamerman D., Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference (Chapman & Hall) 1997.
[14] Kim D., Noh J. and Jeong H., Phys. Rev. E, 70 (2004) 046126.
[15] Del Genio C., Gross T. and Bassler K., Phys. Rev. Lett., 107 (2011) 178701.
[16] Small M., Walker D. and Tse C., Phys. Rev. Lett., 99 (2007) 188702.
[17] Newman M., Networks: An Introduction (Oxford University Press) 2010.
[18] Zhang L., Small M. and Judd K., Exactly scale-free scale-free networks, arXiv:1309.0961 (2013).
[19] Knuth D. E., Fundamental Algorithms, 2nd edition, Vol. 1 (Addison-Wesley, Reading, Mass.) 1973.
[20] Bekassy A., Bekessy P. and Komlos J., Stud. Sci. Math. Hung., 7 (1972) 343.
[21] Van der Hofstad R. and Spencer J., Eur. J. Comb., 27 (2006) 1294.
[22] Chung F. and Lu L., Ann. Comb., 6 (2002) 125.