

Growing optimal scale-free networks via likelihood

Michael Small,* Yingying Li, Thomas Stemler, and Kevin Judd

School of Mathematics and Statistics, University of Western Australia, Crawley, WA, Australia, 6009

(Received 1 April 2014; revised manuscript received 11 November 2014; published 7 April 2015)

Preferential attachment, by which new nodes attach to existing nodes with probability proportional to the existing nodes' degree, has become the standard growth model for scale-free networks, where the asymptotic probability of a node having degree k is proportional to $k^{-\gamma}$. However, the motivation for this model is entirely *ad hoc*. We use exact likelihood arguments and show that the optimal way to build a scale-free network is to attach most new links to nodes of low degree. Curiously, this leads to a scale-free network with a single dominant hub: a starlike structure we call a *superstar network*. Asymptotically, the optimal strategy is to attach each new node to *one of the nodes* of degree k with probability proportional to $\frac{1}{N+\zeta(\gamma)(k+1)^\gamma}$ (in a N node network): a stronger bias toward high degree nodes than exhibited by standard preferential attachment. Our algorithm generates optimally scale-free networks (the superstar networks) as well as randomly sampling the space of all scale-free networks with a given degree exponent γ . We generate viable realization with finite N for $1 \ll \gamma < 2$ as well as $\gamma > 2$. We observe an apparently discontinuous transition at $\gamma \approx 2$ between so-called superstar networks and more treelike realizations. Gradually increasing γ further leads to reemergence of a superstar hub. To quantify these structural features, we derive a new analytic expression for the expected degree exponent of a pure preferential attachment process and introduce alternative measures of network entropy. Our approach is generic and can also be applied to an arbitrary degree distribution.

DOI: [10.1103/PhysRevE.91.042801](https://doi.org/10.1103/PhysRevE.91.042801)

PACS number(s): 89.75.Hc, 89.75.Da, 89.75.Fb

I. INTRODUCTION

Complex networks appear to be virtually ubiquitous [1], and, moreover, there is currently a growing industry in the study of networks with a power-law distribution of node degree. This recent activity can be traced back to Barabási and Albert's seminal work [2] in which they described the preferential attachment growth model. They showed that if one grows a network by adding nodes in such a way that new nodes preferentially attached to high-degree nodes, then the resulting network will be scale-free; that is, the result of such a process is a connected network with the probability of a vertex having degree k being proportional to $k^{-\gamma}$.

The surprising power of the *Barabási and Albert preferential attachment growth model* (hereafter, BA) is twofold. First, the scale-free distribution of node degree is implicit and arises naturally. Second, BA generates networks which seem to explain much, but not all, of the great variety of scale-free networks in nature. However, BA does not generate random representative realizations from the set of all scale-free graphs [3,4]. In Ref. [5] we proposed a Monte Carlo Markov Chain framework which does just this, thereby showing that many of the properties attributed to a scale-free degree distribution are actually dependent on the particular generative model.

However, the algorithm in Ref. [5] is not a growth model. In this paper we ask: What is the optimal (in terms of likelihood) way to arrive at a scale-free network via a growth process? The surprising result is that BA is not the right answer. The optimal (via a greedy likelihood maximization process) way of growing a scale-free network is to most often connect to low-degree nodes. This leads to a starlike scale-free network with a single dominant hub: what we call a *superstar network*. Careful

examination of the relative prevalence of low- and high-degree nodes reveals that, when cast in a manner analogous to BA, our algorithm chooses nodes with degree k with probability proportional to $\frac{k^\gamma}{N^2}$ (for $k \ll N$). This is a *stronger bias towards higher degree nodes* than posited in Ref. [2] and one that explicitly incorporates the target node degree. Nonetheless, both our algorithm and BA rely on knowledge of the global degree sequence of the growing network, yet neither method needs global connectivity information.

Our result does not diminish the remarkable observation that BA will naturally lead to a scale-free network. Conversely, our algorithm incorporates the scale exponent γ explicitly and hence allows one to grow a scale-free network of arbitrary exponent; indeed, the algorithm can easily be modified to grow networks of an arbitrary degree distribution. Our algorithm provides a mechanism for the generation of scale-free (often nonsmall world [6]) networks with $\gamma < 2$. Such networks have been widely observed [6–8], but largely viewed as inconsistent with the current generative models and hence rather pathological. Typical networks produced by this algorithm are depicted in Fig. 1. Contrary to the titular claim of [9], we observe both sparse and dense scale-free networks. Indeed, just as is claimed in Ref. [9], we see that for $\gamma \ll 2$, the so-called graphical fraction becomes small; viable networks become scarcer. However, for γ closer to two, finite dense networks are fairly common [5]. The structural change in the networks of Fig. 1 at $\gamma \approx 2$ reflects the effect of increasing the number of high-degree nodes as highlighted by Ref. [9].

Section II introduces our primary algorithm, an analytic derivation of the optimal (maximum likelihood) model for growth of a scale-free network and our main results. In Sec. III we address some minor issues concerning the maximum degree of the resultant network; we extend the results of Sec. II to build truncated power-law networks by introducing a node-degree cutoff. The cutoff acts to explicitly forbid

*michael.small@uwa.edu.au

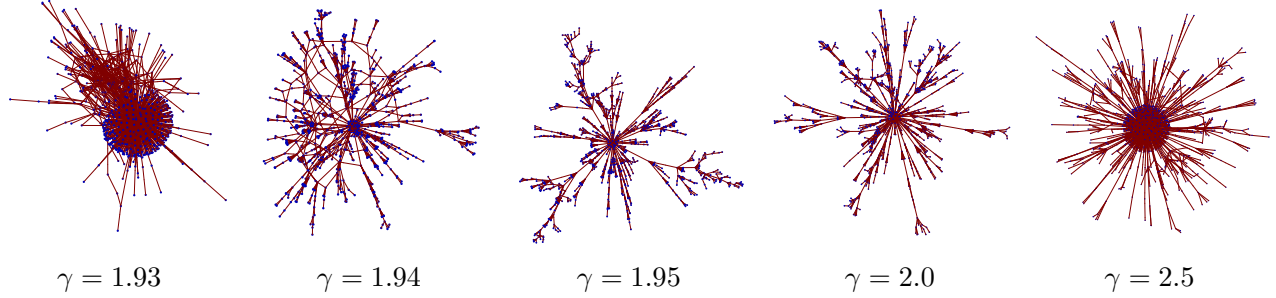


FIG. 1. (Color online) Representative realizations of the optimal scale-free network generation scheme with $\gamma = 1.93, 1.94, 1.95, 2.0,$ and 2.5 ($N = 10^3$). For $\gamma \ll 2$ the algorithm is unable to find viable networks; however, for $1 \ll \gamma < 2$ viable networks do exist (certainly for finite N it is easy to see why this is the case), and such *dense* networks are, in fact, quite plentiful [5]. As γ approaches 2.0 there is a sudden phase transition to large treelike networks with a proliferation of cross-links. As γ increases further, the cross-links gradually become scarcer and the networks we obtain evolve from highly treelike back to hub-centric superstar networks.

networks from growing too large. Nonetheless, we see that the cutoff alone is neither enough to explain nor suppress the emergence of superstar networks. Section IV concludes.

Following the main paper are two short and technical appendices. In Appendix A we address some technical issues concerning the dependence of scale-free exponent γ on the number of added links m (this extends the well-known asymptotic result: in the tail of the distribution $\gamma \rightarrow 3$ independent of m as $N \rightarrow \infty$). Finally, Appendix B addresses some technical computational issues concerning growth of maximum likelihood networks for $m > 1$.

The next section presents our primary results.

II. GREEDY MAXIMUM LIKELIHOOD GROWTH

Let G_N be a network of N nodes. Let n_k denote the number of nodes in G_N with k links. Hence $N = \sum_k n_k$ and $\mathbf{n}(G_N) = [n_1, n_2, \dots, n_k, \dots, n_N]$ is the histogram of degree distribution. A scale-free network is usually *defined* to be a graph for which the probability of a node having degree k is given by

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma)}. \quad (1)$$

The denominator $\zeta(\gamma)$ is the Riemann zeta function and provides the necessary normalization. However, this is a very restricted definition of scale-free, and it will not be sufficient for our purposes. This equality can only hold asymptotically, and for real networks deviation from this definition is to be expected. For BA, one must also account for the minimum degree m , such that $p_k = 0$ for $k \leq m$, and hence (1) is replaced by a shifted power law. In what follows we will provide a rigorous probabilistic definition of scale-free, based on the likelihood of an observed degree histogram conforming to the ideal distribution p_k .

Starting from some seed network G_s of s nodes we wish to add nodes so that

$$\lim_{N \rightarrow \infty} \frac{n_k}{N} = p_k. \quad (2)$$

We do this with a series of moves, adding both nodes and edges, in such a way that we produce networks that are increasingly likely to satisfy Eq. (2). The likelihood of G_N conforming to

degree distribution p_k is given by the multinomial distribution

$$P(G_N) = N! \prod_{k=1}^N \frac{p_k^{n_k}}{n_k!}. \quad (3)$$

There is a small discrepancy between n_k and p_k . As defined above $p_k > 0$ for all k , including $k \geq N$. One way to circumvent this would be to use a truncated power law instead of (1). However, the introduction of such truncation incurs an additional parameter which we choose to ignore at present. In Sec. III we examine the effect of a cutoff parameter more closely. We eventually conclude [see Eq. (9)] that our attachment rule naturally induces a more effective size dependent cutoff of its own.

Our criteria (3) depends on only node degree; we assume that all connected networks with a given degree sequence are equally probable [5, 10]. One could also condition (3) on other desirable network properties: assortativity or embeddability, for example. We seek a sequence of moves yielding a sequence of networks with increasing $P(G)$. Each move can either add a new node with one edge, or add another edge from the last added node to the rest of the network. We only propose moves which are modifications to the last node added and its connections; we add nodes one at a time and with a variable number of links.

The first move we consider is the addition of one node of degree one. By adding a new node with one link, we move from G_N to G_{N+1} , and the marginal payoff is

$$\begin{aligned} Q_{\text{node}}(k) &= \frac{P(G_{N+1})}{(N+1)\tilde{p}P(G_N)} = \frac{(N+1)! \prod_{k=1}^N \frac{p_k^{n_k}}{n_k!}}{(N+1)\tilde{p}N! \prod_{k=1}^N \frac{p_k^{n_k}}{n_k!}} \\ &= \begin{cases} \frac{p_2}{n_2+1} \frac{1}{\tilde{p}} & k=1 \\ \frac{p_1}{n_1+1} \frac{n_k}{p_k} \frac{p_{k+1}}{n_{k+1}+1} \frac{1}{\tilde{p}} & k>1 \end{cases}. \end{aligned} \quad (4)$$

The additional term $(N+1)\tilde{p}$ accounts for the additional node. To provide a fair comparison between the numerator and denominator we explicitly include the extra node: combinatorially, it could be any of the $(N+1)$ nodes, yet as it is not connected to existing nodes its corresponding probability is chosen to be the complement of all other possibilities $\tilde{p} = 1 - \sum_{k=1}^N p_k$. Both simulation and dimensional arguments indicate that this is the correct choice. We find that choice of \tilde{p} which do

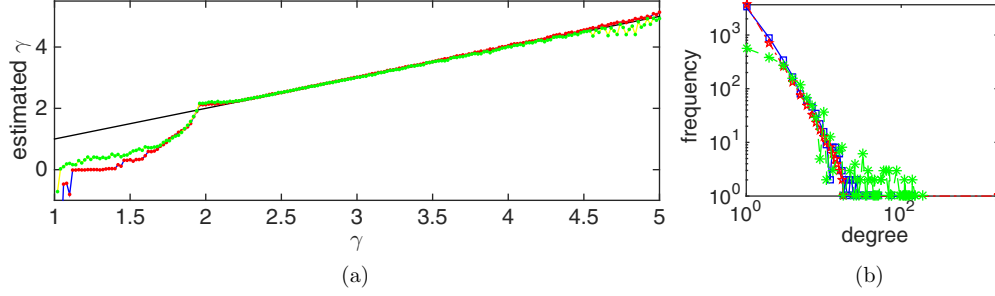


FIG. 2. (Color online) (a) As we vary $\gamma \in [1.5, 5]$ we generate optimal and random realization of our scale-free network growth algorithm. We plot γ against $\hat{\gamma}$ (the maximum likelihood estimate of the scale-free exponent from the sample degree distribution) for both optimal (red) and partially random (green: slightly larger variance at extrema) realizations with $N = 10^4$ nodes. The identity line is also shown. (b) Typical degree distribution for BA (blue, squares), optimal attachment (red, stars), and semirandom ($q = 0.5$, green asterisks). The partially random scheme (with $q = 0.5$ as described in the text) did not significantly affect the final degree histogram, and yet this allowed much more variation in the resultant networks; for $q < 1$ we see a transition from superstar networks towards the usual results of BA.

not scale with N in this way does not produce viable networks. Denote by n'_k the terms in the new histogram $\mathbf{n}(G_{N+1})$. The only difference between the histograms $\mathbf{n}(G_{N+1})$ and $\mathbf{n}(G_N)$ is that G_{N+1} has an additional node with one link (the last node added), and that link connects to a node formerly of degree k . Hence, $n'_1 = n_1 + 1$; $n'_k = n_k - 1$; $n'_{k+1} = n_{k+1} + 1$; and $n'_i = n_i$ (for $i \notin \{1, k, k+1\}$).

The second move is the addition of another edge from the last added node to the rest of the network. Suppose that the last node added to the network already has j links (initially, $j = 1$). A slightly more complicated counting argument yields

$$Q_{\text{edge-}j}(k) = \frac{P(\tilde{G}_{N,(j+1)})}{P(G_N)} = \begin{cases} \frac{n_{j-1}}{p_{j-1}} \frac{p_{j+1}}{n_{j+1}+1} & k = j-1 \\ \frac{p_{j+1}^2}{(n_{j+1}+1)(n_{j+1}+2)} \frac{n_j(n_j-1)}{p_j^2} & k = j \\ \frac{n_j}{p_j} \frac{p_{j+2}}{n_{j+2}+1} & k = j+1 \\ \frac{n_j}{p_j} \frac{p_{j+1}}{n_{j+1}+1} \frac{n_k}{p_k} \frac{p_{k+1}}{n_{k+1}+1} & |k-j| > 1 \end{cases} \quad (5)$$

The two cases $|k-j|=1$ are needed to avoid spurious terms when an edge is added between nodes of similar degree.

Equations (4)–(5) are sufficient to define two network growth algorithms: one a random process, and one a greedy optimization. We start with a seed network G_s and then at each step we compute $Q_{\text{node}}(k)$ and $Q_{\text{edge-}j}(k)$ as given above. From among these $2N$ marginal payoffs we select the maximum $Q_*(k) > 1$ and adopt the specified move, either adding a link or a new node. This yields a greedy process which we will call the *optimal* scale-free network algorithm (for a given γ and G_s). Alternatively, we may seek random realizations by treating each $Q_*(k) > 0$ as being proportional to the probability of accepting that particular move and by selecting among the possibilities randomly. This yields a *random* scale-free network growth algorithm (for given γ).

Like the BA algorithm, the optimal approach requires knowledge of the degree sequence of the network. Since neither this algorithm nor BA requires true *global* knowledge (of clustering, for example), it is easy to imagine generalization requiring only local information, utilizing the information of a random walker on the network, for example. Scenarios in-

volving random walkers, with local information and likelihood maximizing attachment rules, will be considered elsewhere.

In Fig. 1 we show representative realizations of the optimal scale-free network for γ between 1.93 and 2.5. Figure 1 depicts the strong dependence of network structure on γ and an explosive transition first to treelike graphs (at $\gamma \approx 1.94$) and then a gradual drift back toward superstar networks. Note that, while γ is the degree distribution exponent prescribed for the algorithm, it is not necessarily the actual exponent of the degree distribution of a particular realization of that algorithm. For comparison, in Fig. 2 we plot exactly this quantity. What we observe is an excellent agreement between the target (prescribed) γ and the actual realizations $\hat{\gamma}$, for $\gamma > 2$. For $\gamma < 2$, we observe a nonlinear relationship (as the degree histogram becomes difficult to satisfy), with $\hat{\gamma}$ systematically smaller than the prescribed value. Of course, once $\gamma < 2$ the mean of the asymptotic degree distribution is no longer finite (for $N \rightarrow \infty$). Nonetheless, such networks do exist (for $N < \infty$), and our algorithm continues to grow them. As the asymptotic results no longer converge, the comparison between γ and $\hat{\gamma}$ is much weaker.

We now consider the middle ground between our optimal and random algorithms. Define a parameter q , such that at each time step there is a probability q of performing an attachment move deterministically, and probability $1-q$ of making a random assignment. The deviation between the values generated by the optimal algorithm ($q = 1$) and by a somewhat random counterpart ($q = 0.5$) is small. We stress that both algorithms control only the degree of the connected nodes; their *location* is always random (one is free to choose from among all nodes of equal degree). The optimal algorithm has a deterministic impact on the histogram, while for the random algorithm both the link placement and the changes in the histogram are random.

Next we ask how best to frame these algorithms in terms comparable to BA: we seek to distill from (4) and (5) an analogous rule. Asymptotically, we may suppose (assuming that the algorithm works) that $n_k \rightarrow Np_k$ and hence $n_k \approx \frac{N}{\zeta(\gamma)} k^{-\gamma}$. Substituting this and (1) into (4) and (5) gives

$$Q_{\text{node}} \rightarrow \begin{cases} \frac{N+1}{N+\zeta(\gamma)2^\gamma} & k = 1 \\ \frac{N+1}{N+\zeta(\gamma)} \frac{N}{N+\zeta(\gamma)(k+1)^\gamma} & k > 1 \end{cases}, \quad (6)$$

and

$$Q_{\text{edge}-j} \rightarrow \begin{cases} \frac{N}{N+\zeta(\gamma)(j+1)^\gamma} & k = j-1 \\ \frac{N(N-\zeta(\gamma)j^\gamma)}{[N+\zeta(\gamma)(j+1)^\gamma][N+2\zeta(\gamma)(j+1)^\gamma]} & k = j \\ \frac{N}{N+\zeta(\gamma)(j+2)^\gamma} & k = j+1 \\ \frac{N^2}{[N+\zeta(\gamma)(j+1)^\gamma][N+\zeta(\gamma)(k+1)^\gamma]} & |k-j| > 1 \end{cases} \quad (7)$$

Hence, we see that the maximal likelihood scale-free network is obtained by observing that the probability of attaching a node to a link of degree k is proportional to

$$\text{Prob}(\text{a degree-}k \text{ node}) \propto \frac{1}{N + \zeta(\gamma)(k+1)^\gamma}. \quad (8)$$

However, before comparing (8) to BA we must note that (8) is the probability of attaching to *any one* of the nodes of degree k ; asymptotically we expect there to be

$$n_k \rightarrow Np_k \approx \frac{N}{\zeta(\gamma)} k^{-\gamma}$$

such nodes. BA [2] says that the probability of attaching to a node, if that node has degree k , is proportional to k . Hence, our maximum likelihood algorithm approach says that the

optimal strategy is to link to a node of degree k with probability proportional to

$$\text{Prob}(\text{node-}i|\text{degree node-}i = k) \propto \frac{1}{N} \frac{k^\gamma}{\frac{N}{\zeta(\gamma)} + (k+1)^\gamma}. \quad (9)$$

For $k \ll N$ this is proportional to $\frac{k^\gamma}{N^2}$, and as $k \rightarrow N$ this probability is approximately $\left(\frac{k}{k+1}\right)^\gamma \rightarrow 1$, reducing the likelihood of very high-degree nodes and acting as an implicit degree cutoff. While most links are to low-degree nodes, we find that high-degree nodes are more likely to receive links. The likelihood is significantly stronger than BA, proportional to an increasing power law, and dependent on the desired exponent γ . The combination $\frac{k^\gamma}{N^2}$ reveals an interesting connection between the parameters γ and N and the probability p_k . Other authors (Refs. [11,12], for example) have proposed models that are explicitly dependent on the network size N . What we do is show that the (likelihood) optimal growth model induces a natural scale dependent cutoff of its own, as node degrees become large the probability of attaching to any of the large hubs becomes (asymptotically) equal. Beyond some cutoff all hubs are equally attractive.

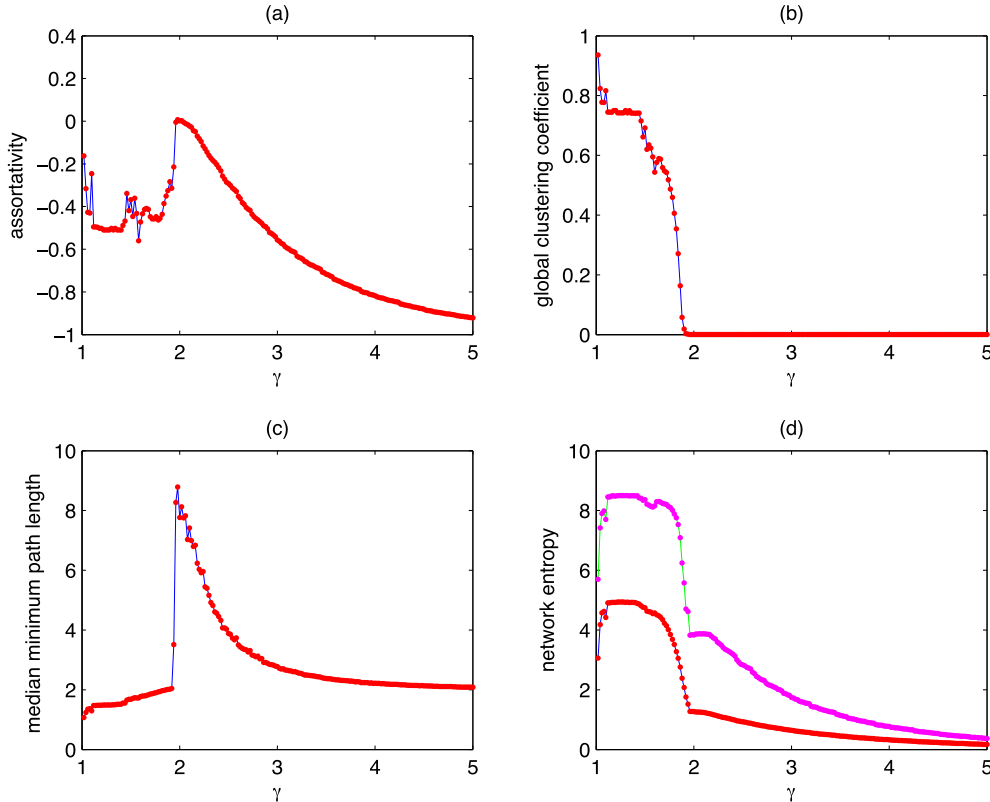


FIG. 3. (Color online) For $\gamma \in (1, 5]$ we compute optimal realizations of our scale-free network generation algorithms ($N = 10^4$). We show (a) assortativity (linear Pearson correlation coefficient), (b) global clustering coefficient, (c) shortest path length, and (d) two measures of network entropy. Note that the hublike nature of networks with large γ is evident from the shortest path length, while the global clustering coefficient drops to zero at $\gamma \approx 2$. Interestingly, assortativity remains negative, peaking with a value of 0 at $\gamma \approx 2$ and then declines rapidly. The shortest path length has clear evidence of the undersize networks for $\gamma \ll 2$, an abrupt transition to treelike networks near 2, and then a gradual decay to a single dominant hub as γ increases further. The two network entropy statistics compute the entropy of the degree sequence [13,14] (lower, red/blue line) and the entropy of the link-degree coincidences (upper, magenta/green line). Both entropy measures show an abrupt, nondifferentiable transition at $\gamma \approx 2$. Each data point reported in these figures is the corresponding statistic values estimated from a single network (with γ increasing between simulations in steps of size 0.01); variance can be inferred from the smoothness of the plotted curves.

The difference between (4)–(5) and (6)–(7) is that the former incorporates deviation from the target degree distribution p_k as a sequence of ratios of the form $\frac{n_k}{p_k}$. The latter assigns links proportional to (9). Unfortunately, an attachment algorithm based on these *asymptotic attachment* rules does not perform well. Simulation has shown that for small N the asymptotic approximation is usually poor, and there appears to be little hope of appropriate convergence. We could be tempted to define a third implementation of our growth algorithm: randomly select nodes according to the distributions (6)–(7) and iterate. Such an algorithm generates what we refer to as *asymptotically* scale-free networks. However, systematic results from this algorithm are not reported here, and performance is generally very poor. For small seed networks, the initial configuration is far from the asymptotic ideal and convergence is poor. For large seed networks, the result depends on the choice of initial network.

In Fig. 3 we compute the usually quoted properties of scale-free networks, for realizations of our first two algorithms. We observe a systematic dependence of these properties on the exponent γ . In particular, there is a sudden, apparently not differentiable, transition in the vicinity of $\gamma \approx 2$ as the network structure rearranges. This transition is indicative of the underlying structural change in the network for γ bigger than two, the onset of the superstar hub structure depicted in Fig. 1. We note that for $\gamma > 2$ the typical maximal node degree scales as $N^{1/(\gamma-1)}$ [15]. For $\gamma < 2$ this implies faster than linear growth and results in the (low-degrees exponent) emergence of a superstar structure in our model.

More generally, we find the range of behaviours demonstrated in Fig. 3 is far wider than what one would observe with straightforward BA. In addition to network topological measures, we also report two measures of network entropy. First, following Bianconi [13,14] (and our own independent and *ad hoc* treatment [16]), we compute the entropy of the events defined by the sample degree histogram; that is, $-\sum \frac{n_k}{N} \log \frac{n_k}{N}$. However, this quantity does not take into account the structure between nodes of the network (i.e., how the high and low-degree nodes are distributed within that topology). Hence, we also compute what we call the *network link entropy*. If $e_{i,j}$ is the sample probability of an edge joining nodes of degree i and j then this version of entropy is computed as $-\sum_{i,j} e_{i,j} \log e_{i,j}$. Figure 3(d) illustrates the result of both computations.

Figure 4 demonstrates that, even if γ is restricted to values which one obtains from BA, the schemes we propose here exhibit a much wider variation in network structures. The assortativity of our algorithm is stronger (more negative), and as consequence, the mean path length is lower. The BA algorithm illustrated here has minimum degree $m = 1$, and hence one can compute (see Appendix A) that the expected value of γ is approximately 2.471, in excellent agreement with Fig. 4(c). In comparison, the optimal algorithm adds more links per node and achieves a significantly higher value of γ .

Using likelihood arguments we have illustrated a new network growth model that generates structures distinct from those observed elsewhere. As a consequence of these likelihood methods we have observed the emergence of a natural node degree cutoff in Eq. (9). Previously other authors have

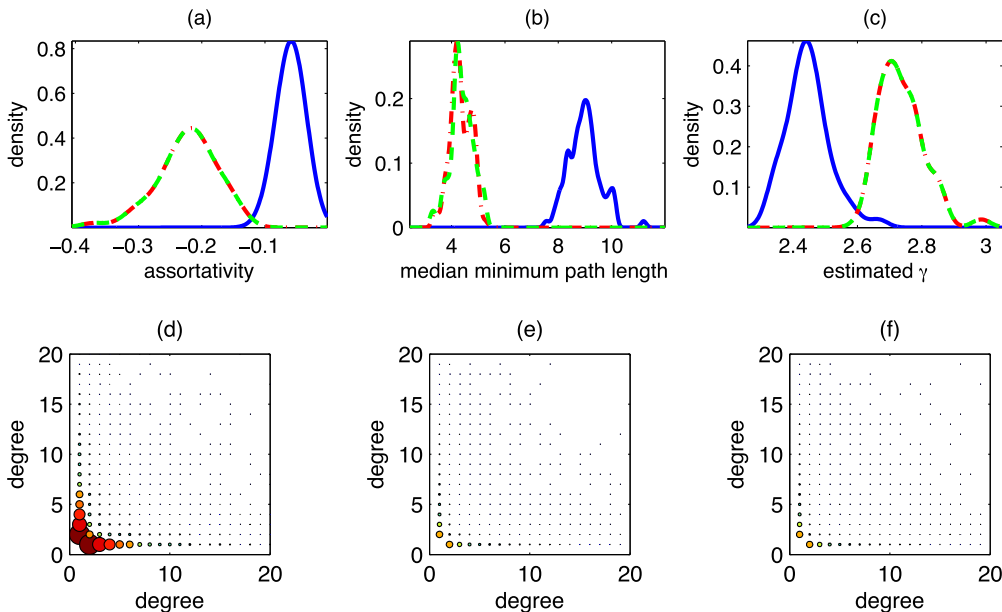


FIG. 4. (Color online) We generate BA networks (blue solid lines), estimate the scale exponent γ , and then generate networks according to our first two schemes with this estimated value $\hat{\gamma}$ and $N = 10^4$ nodes (red dot-dashed and green dashed lines). Displayed here are the usual network properties estimated from the resultant networks and depicted as histograms (generated from 100 network realizations via a Gaussian kernel smoothing algorithm): (a) assortativity, (b) shortest path length, and (c) estimated exponent $\hat{\gamma}$ (adaptive binning). Remarkably the distributions reported for the optimal scheme (red dot-dashed) is almost identical to the results of the randomization scheme ($q = 0.5$ and green dashed lines). As assortativity is a linear measure, it is not particularly good at describing the detail of degree-degree correlation. Panels (d)–(f) illustrate scatter plots (circle size and colour proportional to likelihood/number) of actual degree-degree structure for representative networks: (d) BA, (e) optimal, and (f) random.

proposed other node degree cutoffs (Ref. [11] for example). In the next section, we present a technical comparison of these truncation methods to the growth process proposed in this paper.

III. TRUNCATED POWER LAWS, AN UPPER BOUND ON NODE DEGREE, AND A “NATURAL” CUTOFF FUNCTION

We have seen that Eq. (9) acts as an implicit cutoff and limits the growth of very high-degree nodes. While it is pleasing to observe the manifestation of this cutoff directly in our maximum likelihood model, there are several models that impose explicit (but so-called “natural”) cutoffs in the power-law degree distribution of evolving scale-free networks [15]. It is perhaps useful to ask whether the addition of an explicit mechanistic cutoff to our optimal scheme will either significantly alter our results or provide an explanation for the superstar effect and deviation from the BA model. In this section we show that the answer to both propositions is “no,” except when imposing extreme restrictions on the node degree.

To illustrate these results we explore the effect of imposing an arbitrary maximum degree $C_{\max} < N$ on a scale-free network network following the optimal growth procedure described in Sec. II. The growth process of Sec. II remains the primary contribution of this paper.

The optimal method proposed in the previous section tends to generate scale-free networks with a single big hub, which means the degree distribution in log-log scale has one particular extremum in the far tail of the distribution far away from the power law. Because this datum corresponds to an extremely low-probability event it does not significantly affect the likelihood evaluation and is hence a natural solution to the greedy optimization process. Conversely we note that the BA method produces networks with significant deviation among the low-degree nodes, deviation which is routinely explicitly ignored when estimating the exponent of such networks (see Appendix A). Nonetheless, our networks produce a small deviation from the ideal scale-free property in the resulting degree histogram, and it is natural to wonder whether the superstar networks we observe are only an artifact of this single node in the far tail of the distribution. To test this we repeat the optimal growth process described in Sec. II after modifying (1) to include an explicit cutoff function. We now consider the truncated power-law distribution

$$p_k = \begin{cases} Ck^{-\gamma}, & k \leq C_{\max} \\ Ck^{-\alpha\gamma}, & \text{otherwise} \end{cases} \quad (10)$$

where C is the normalization constant and α can be any number greater than one when $\gamma > 2$, and C_{\max} is the cutoff value. We call this the *truncated* maximum likelihood method; as this approach is parameterized by the constant C_{\max} , we will equivalently refer to this as the C_{\max} method.

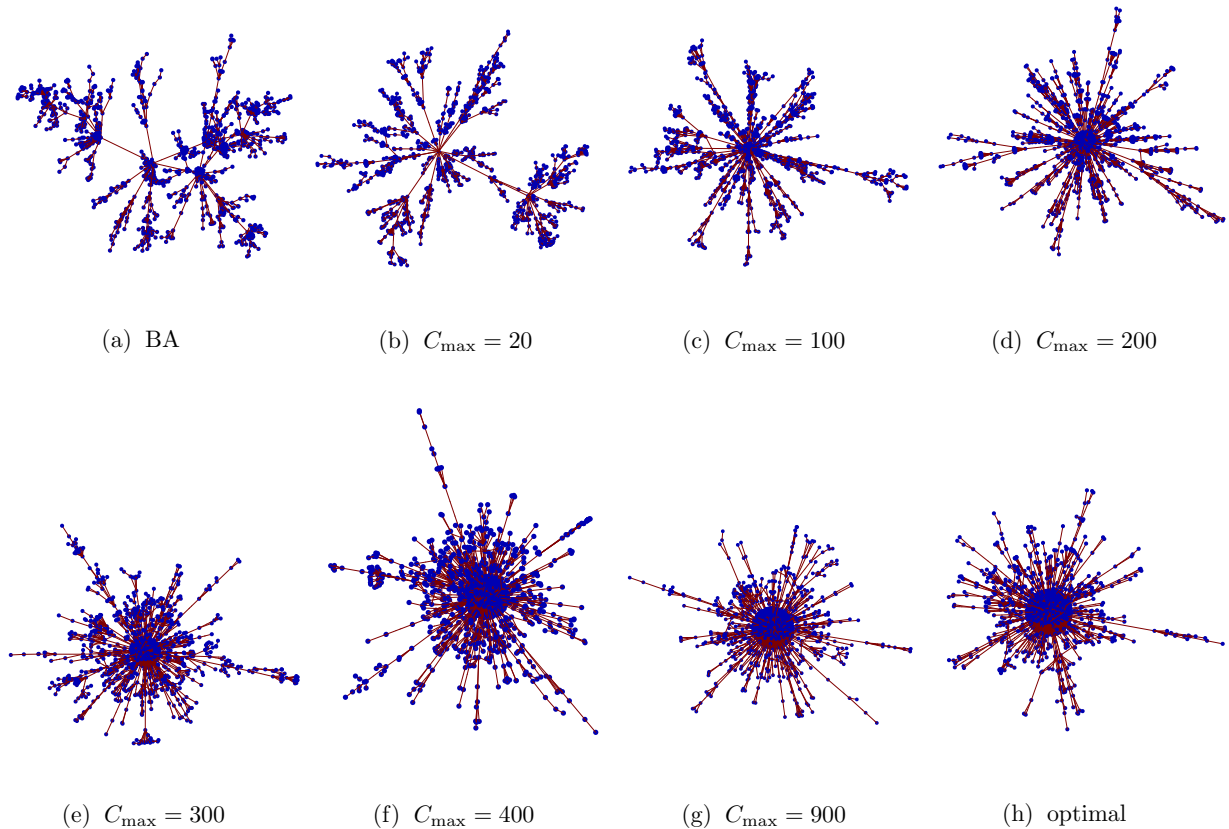


FIG. 5. (Color online) Representative realizations of the BA method, optimal method (Sec. II), and an imposed C_{\max} with $C_{\max} = 20, 100, 200, 300, 400, 900$ ($N = 10^3$) (Sec. III). When C_{\max} is quite small, i.e., $C_{\max} \leq 100$, networks look like the BA network because of a scattered structure. For $C_{\max} > 100$, networks tend to have less hubs and gradually evolve to concentrated superstar networks, equivalent to the optimal method.

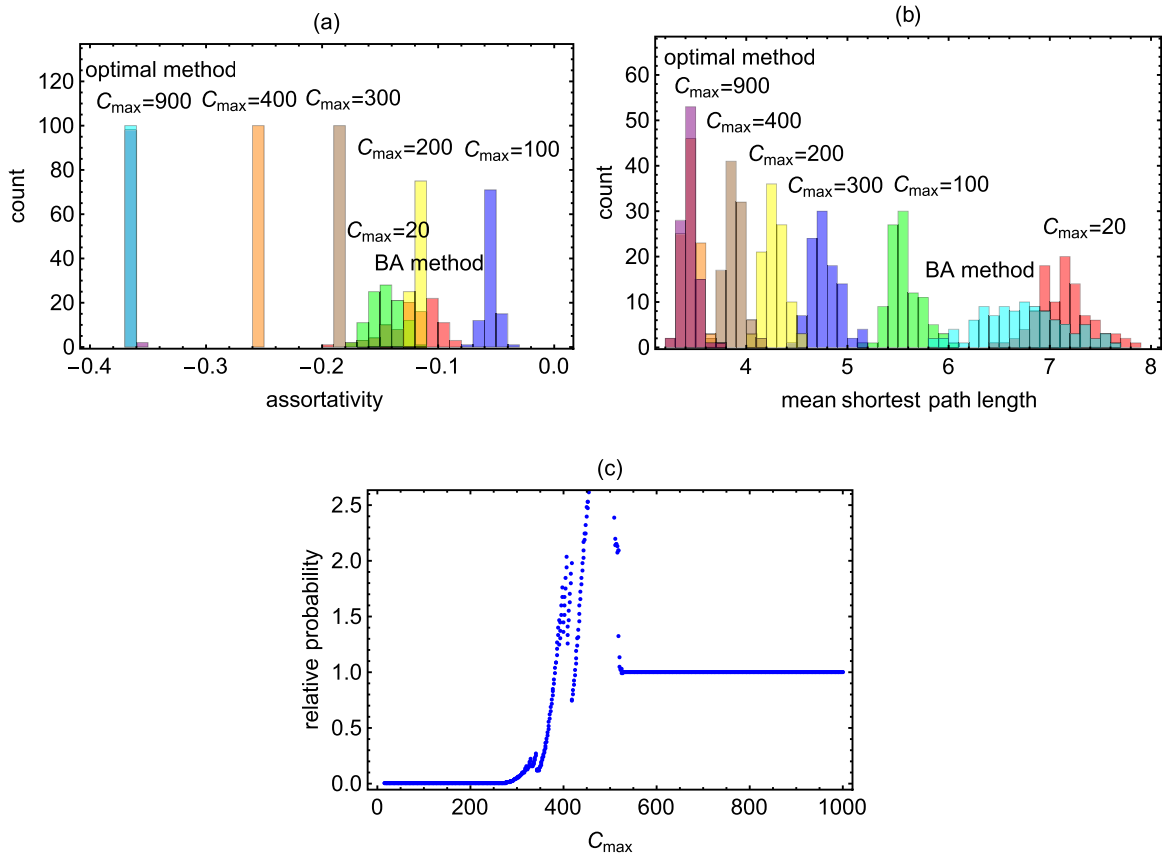


FIG. 6. (Color online) In (a) and (b), we apply the optimal method, the BA method and the truncated C_{\max} method with different C_{\max} to generate scale-free networks. For each method or each value of C_{\max} , we generate 100 networks, and then display the histogram of common properties of scale-free networks: assortativity and shortest path length. The number on the top of the bars indicates the C_{\max} value. From panels (a) and (b) we can see that when the C_{\max} value is large, the networks we have very similar properties to the optimal maximum likelihood networks. This makes sense because if C_{\max} is large, the cutoff function has so little power that the optimal method and the C_{\max} method are almost the same. Then, when we decrease C_{\max} , the properties of networks move towards those of BA networks. Notably, when C_{\max} is small, the properties of C_{\max} networks are quite similar to those of BA networks. All of this indicates that the C_{\max} method can actually link the BA method and optimal method. Panel (c) illustrates relative probability (relative to optimal maximum likelihood network) of C_{\max} networks, with C_{\max} ranging from 15 to N . Surprisingly, when C_{\max} is neither very big nor very small, the relative probability will be higher than 1, which means the probability of truncated C_{\max} network will be bigger than optimal networks. This suggests that the truncated approach provides a useful middle ground to generate networks “between” BA and the optimal growth networks.

When C_{\max} is large, the effect of the additional cutoff term in the power-law distribution function is small, and the truncated method behaves similarly to the greedy optimal method described above: the network results in a single dominant hub and a potential superstar network. However, as we decrease C_{\max} , the cutoff function has more influence, and we observe that this gradually splits the superhub into multiple smaller (but still large) hubs. Figure 5 demonstrates this process. That is, the effect of the introduction of a maximum degree C_{\max} results in the largest hub being reduced to a rich club of multiple large (but nonetheless, *smaller*) degree hub nodes, and these nodes are interconnected. One can view this rich club of high-degree nodes as a virtual superhub: replacing these nodes with a single node of degree greater than C_{\max} results in networks exactly equivalent to the previous section. The superstar networks and their hub nodes are not an artifact.

Figure 5 provides representative realizations of the BA method, the optimal method, and the C_{\max} method (for C_{\max} between 20 and N). In Fig. 5 we can observe the structure of

networks with different C_{\max} . It is easy to notice that when we decrease C_{\max} , the network evolves from a superstar structure to a scattered structure. In Fig. 6 properties of C_{\max} networks move from those of optimal networks to BA networks when we decrease C_{\max} from N to 20, further supporting the qualitative observations of Fig. 5. Another interesting observation is that when C_{\max} is rather small, networks can be very similar to BA networks, both in the structural sense illustrated in Fig. 5 and in the properties shown in Fig. 6. The C_{\max} method provides a link between the BA and optimal method via the parameter C_{\max} .

To identify whether a given network is more similar to the BA or optimal model, the easiest and most direct way is to count the number (and size) of hubs. Hubs have frequently been shown to have a disproportionate importance in the overall network structure [17–19], heavily influencing properties including assortativity, transitivity, rich clubs, stability, and centrality. Studying the location and connectivity of hub nodes offers significant insight into the overall structure and

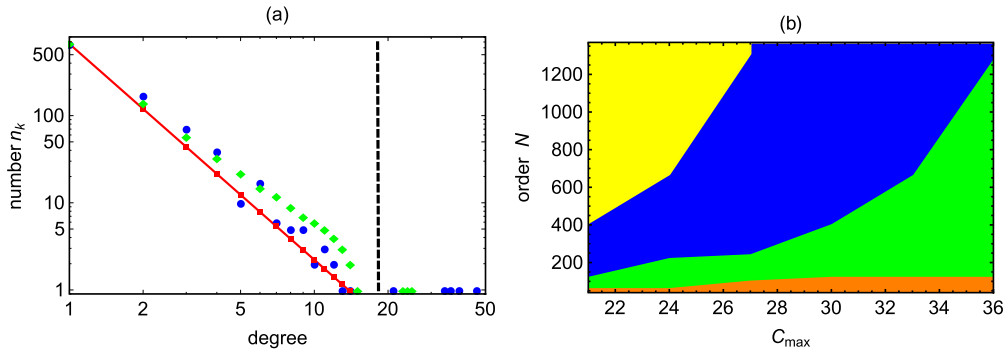


FIG. 7. (Color online) Panel (a) illustrates our constructive definition of hubs. Starting with a fixed $\gamma = 2.47$ (corresponding to the estimated γ in comparable BA networks), we construct a straight line in the degree distribution on the log-log scale (the solid red line) $\log(n_k) = \log(n_1) - \gamma \log(k)$ where k is degree and n_k is the number of nodes with degree k . The intersection of this line and x axis occurs at $\log(n_1)/\gamma$. As there is some variability at the end of the degree distribution, we add θ to exclude the tail. In panel (a), the green diamonds show the degree distribution of C_{\max} network with $C_{\max} = 25$, and blue circles show the BA network. Here we use $\theta = 3$ and draw the minimum degree of hubs as the black line. From (a) we can see our definition successfully distinguishes the hubs in the networks. In (b), for $C_{\max} \in [21, 36]$ and $\gamma = 2.47$ (the estimated γ of BA method when minimum degree is 1), we compute realizations of C_{\max} method and the number of hubs in the representative networks. The orange, green, blue, and yellow areas (from bottom to top in the plot) indicate the number of hubs will be 1, 2, 3, 4 with particular C_{\max} and N in the representative areas. Note that if we fix C_{\max} , when we increase the size of networks, the number of hubs will increase; if we fix N and increase C_{\max} , the number of hubs will decrease.

properties of a network [19]. If there is only one hub (as with the superstar networks), the network is highly concentrated; if there are many hubs, the network is more scattered.

Now we wish to quantify the proliferation and placement of hubs, and we provide the following working definition for that purpose. Naturally, the only judge of whether a node is a hub is the degree. We define the minimum degree of hubs as

$$\log(n_1)/\gamma + \theta,$$

and any node with an equal or higher degree is counted as a hub. Here γ is the degree exponent of the true asymptotic distribution, not that estimated from the data. The parameter θ can be any value more than 1; with different θ the exact number of hubs may vary, but the overall tendency remains the same. This is not the only definition of hubs, but alternative definitions yield the same tendency.

When we change the C_{\max} value, we see that the number of hubs also changes. Figure 7 shows the tendency of the number of hubs. Decreasing C_{\max} or increasing the size of the network both lead to an increase in the number of hubs. This can be explained as follows. When the network grows, it is natural to generate more centers because as the population N grows, nodes tend to (perhaps) gather into different groups and gather with different nodes instead of being all together. If C_{\max} decreases, the cutoff function has more influence and so forces the richer nodes in the network to have fewer connections. However, as those richer nodes have fewer connections, their former neighbors need to connect elsewhere, and those new connections increase the degree of other (relatively) low-degree nodes, increasing the prevalence of hubs.

We have now provided a thorough examination of the effect of natural and explicitly imposed degree cutoffs on the growth of hubs. We find that the absence of cutoffs alone is not enough to explain the superstar networks we generate with the algorithm introduced in Sec. II. In the next section we provide a brief conclusion.

IV. CONCLUSION

Although we have derived four separate algorithms, we examine only three of them: the asymptotic scheme gives poor results if the seed network G_s is small. Nonetheless, this algorithm does provide insight into the asymptotic behavior of the other methods. Each of the three algorithms we present here provides a technique to obtain random realizations of networks consistent with a particular growth process. It has been argued that growing a network inherently biases the random sampling of the wider space of all networks consistent with a given degree distribution [3]. While in Ref. [5] we address the issue of random realizations from the space of all networks defined by a particular degree distribution, in this paper we propose a more narrowly defined growth algorithm. We demonstrate that BA does not provide the most *likely* algorithm for network growth consistent with a particular power-law degree distribution. In a sense, this extends the arguments of Ref. [3]; not only is random growth biasing one's selection from the space of all networks, BA is a biased selection [20,21] from the space of randomly grown scale-free networks. Our algorithm provides a new approach to growing scale-free networks with an arbitrary degree exponent; moreover, these networks exhibit a range of structural features beyond what one would expect from BA.

Of particular interest are the superstar networks that emerge from the optimal algorithm. These networks possess a unique structure not previously explored via standard growth processes. Recent work on explosive synchronization in star networks [22] demonstrates the importance of understanding this particular class of networks. Here we see that superstar networks emerge via a natural and optimal growth process. This also provides a natural mechanism for the very large diameter and small exponent (i.e., $\gamma < 2$) scale-free network observed in the real-world transmission of avian influenza [6], a large number of superstar hubs distributed geographically. Similar small exponent superstar networks have also been

observed for networks of musical preference [7] and sexual promiscuity [8]. One possible explanation for the origin of this superstar structure is offered by the strongly linked hub nodes in a rich club [19]; as we observe, the imposition of a strict node-degree cutoff leads to the formation of a cluster of high-degree nodes acting as a virtual superstar. Nonetheless, the algorithm we present provides a simple mechanism for generation of networks such as these. Conversely, the single dominant “superstar” can easily be forbidden by truncating (1) with a harsher upper bound dependent on N . Doing so produces networks with a distributed cluster of hubs.

Mathematica and MATLAB implementations of the algorithms described in this paper are available from the first author.

ACKNOWLEDGMENTS

M.S. is funded by the Australian Research Council via a Future Fellowship (FT110100896) and Discovery Project (DP140100203). Y.L. is supported by the UWA-USTC research training programme.

APPENDIX A: EXPECTED DEGREE

We provide an analytic expression for the expected degree exponent γ of a preferential scale-free network. The standard result [23] holds that, for BA growth, $\gamma \rightarrow 3$ independent of m . However, this result is an asymptotic one concerning the tail of the distribution. A more useful statistic for what we are doing here is to estimate γ from the entire distribution (1). Doing so yields quite a different answer and we achieve excellent agreement between theory and computation. Moreover, this is a far more useful statistical measure for finite networks than the asymptotic result. After introducing our results we will briefly discuss the reasons behind the deviations from the results in Ref. [23] in a little more detail.

We perform a preferential attachment growth process to generate a scale-free network. At each stage we add a new node with m new links. Let k denote the degree of a node, and N the number of nodes in the network. The degree distribution is assumed to converge to a power law (for $k \geq m$) of the form $k^{-\gamma}$, and we obtain an exact implicit relationship for γ , m , and N . We verify this with numerical calculations over several orders of magnitude. Although this expression is exact, it provides only an implicit expression for $\gamma(m)$. Nonetheless, we provide a reasonable guess as to the form of this curve and perform curve fitting to estimate the parameters of that curve, demonstrating excellent agreement between numerical fit, theory, and simulation.

Preferential attachment [2] is the archetypal growth mechanism for scale-free networks. Asymptotically, under certain circumstances, such networks produce a degree distribution which converges asymptotically to a power law with exponent 3. But this is not true in general, and it is not true for arbitrary finite networks generated along the way. In this appendix we derive straightforward analytic results for the expected exponent γ of a scale-free network with power-law degree distribution $p(k) \propto k^{-\gamma}$.

We assume that the network is grown with a BA process as described in Ref. [2]. With each new node we add m links, and the growth process is terminated when the network has N

nodes. We make the approximation that the degree distribution of this finite network follows a shifted power law [24] with some exponent γ .

Hence, a BA network with minimum degree m will add exactly m new links for each new node. The expected degree

$$E(k) = 2m \tag{A1}$$

(since each link has two ends and contributed to the degree of two nodes). Conversely, the probability that a node has degree k is given by

$$P(k|\gamma, d) = \begin{cases} 0 & k < m \\ \frac{k^{-\gamma}}{K(\gamma)} & k \geq m \end{cases}$$

where the normalization factor $K(\gamma)$ is inconvenient. However,

$$\zeta(\gamma) = \left(\sum_{k=1}^{m-1} + \sum_{k=m}^{\infty} \right) k^{-\gamma} = \sum_{k=1}^{m-1} k^{-\gamma} + K(\gamma),$$

and hence it is easily computable.

The expected degree is

$$E(k) = \sum_{k=1}^{\infty} k P(k|\gamma) = \frac{\sum_{k=m}^{\infty} k^{1-\gamma}}{\zeta(\gamma) - \sum_{k=1}^{m-1} k^{-\gamma}}. \tag{A2}$$

Equating (A1) and (A2), we have that the asymptotic value of γ satisfies

$$\zeta(\gamma) = \sum_{k=1}^{m-1} k^{-\gamma} + \frac{1}{2m} \sum_{k=m}^{\infty} k^{1-\gamma}. \tag{A3}$$

Replacing the left-hand side of (A3) with the corresponding infinite sum and canceling identical terms we obtain

$$\sum_{k=m}^{\infty} (2m - k) k^{-\gamma} = 0. \tag{A4}$$

Solving (A4) allows us to determine the expected value of γ for the BA algorithm with a particular choice of minimum degree m . In particular, for $m = 1$ we recover $2\zeta(\gamma) = E(k)$.

In Fig. 8 we illustrate the agreement between sample preferential attachment networks of various sizes and the prediction of (A4). The curve appears to be asymptotic to $\gamma = 3$, and so we fit a function of the form $\hat{\gamma}(m) = 3 - (m + \alpha)^{-\beta}$ to the solution of the series (A4). We find that

$$\gamma(m) \approx 3 - \frac{1}{(m + 0.925)^{0.9932}}.$$

These results are required to explain expected degree distributions observed in Sec. II (Fig. 4), and in that case also show excellent agreement.

As noted above, this is not the same answer as that provided in the review of Albert and Barabási [23]. The discrepancy arises from the methods used to estimate the exponent γ . Essentially, the standard maximum likelihood approach described by Newman [25] imposes a minimum degree x_{\min} and estimates γ using a maximum likelihood expression:

$$\gamma = 1 + n \Big/ \sum_{x_i > x_{\min}} \frac{x_i}{x_{\min}}.$$

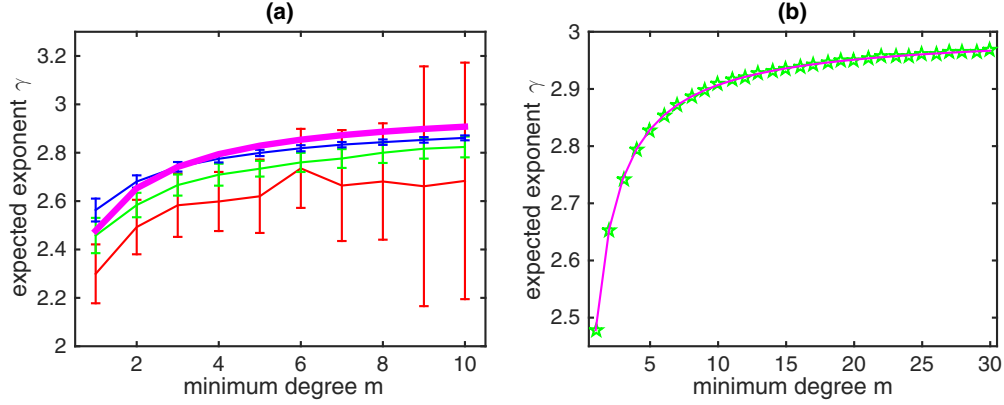


FIG. 8. (Color online) Left panel: Expected values of γ as a function of m [Eq. (A2)] (heavy magenta line) and estimated values of γ from 30 independent realizations of BA networks of size N (mean \pm standard deviation). We take $m \in [1, 10]$ and $N = 10^3$ (red), 10^4 (green), 10^5 (blue). Right panel: γ as a function of m computed via the solution of (A4) (stars) and estimated from a function fit of the form $\hat{\gamma}(m) = 3 - (m + \alpha)^{-\beta}$. The best fit (obtained from a fit on $m \in [1, 10]$) is then extrapolated over the domain. Parameter values are $\alpha = 0.9205$ and $\beta = 0.9932$.

This expression becomes independent of x_{\min} only once $x_{\min} > \min_{i=1, \dots, N} x_i$ and invokes Bayes' rule with a uniform prior on γ . Consequently, for x_{\min} sufficiently large one observes that for preferential attachment $\gamma \rightarrow 3$ independent of m . Instead, we choose to maximise likelihood directly over the finite degree histogram.

The review paper [23] derives this asymptotic degree distribution in three different ways:

(1) *Continuum theory*: Critically, the continuum approach assumes that k_i is a continuous real variable when it is in fact a discrete random process. Moreover, the results hold only asymptotically. The continuity assumption holds only in the tail of the distribution, avoiding the systematic bias away from the power law for low-degree nodes.

(2) *Master equation*: This approach actually obtains a slightly different expression for $P(k)$, which is dependent on m and scales as the inverse of a cubic polynomial. Again, this is an approximation that is valid in the tail but not the head of the distribution.

(3) *Rate equation*: Similar to the previous approach, one will obtain an equivalent expression under the same assumption.

Of these three approaches it is only the continuum approach which faithfully yields the claimed result that $P(k) \propto k^{-3}$, and this is true only in the tail (where the assumption the k_i is a continuous real variable is valid). Albert and Barabási acknowledge this in Sec. VII C of their review: “these methods (the master- and rate- equation approaches), not using a continuum assumption, appear more suitable for obtaining exact results in more challenging network models.”

To estimate γ in the case of pure preferential attachment (the BA process) there is no reason to insist on any choice other than $x_{\min} = m$; the entire distribution should be scale-free and to do otherwise unnecessarily favors the presumed asymptotic behavior of the tail of a finite graph. We stress that we employ our estimate of γ as nothing more than a descriptive statistic. For this purpose, it makes no sense to seek the asymptotic value which is independent of the structure we are trying to quantify. However, as the graphs we generate conform to the power law over their entire range, the characterization

we produce here is also the correct exponent to describe that distribution.

APPENDIX B: GENERALIZED MAXIMUM LIKELIHOOD GROWTH

In the main text we focus on growth by adding a single new node or a single edge at each time step. However, the BA model of preferential attachment adds each new node and m edges simultaneously. The resulting degree distribution is not (1) and has the additional constraint that $p_k = 0$ for $k < m$ where $m > 1$. In particular $p_1 = 0$, and there is no chance of encountering a node with degree 1 (or any degree less than m). One possible computational expedient to overcome this problem is to replace $p_k = 0$ with $p_k = \epsilon > 0$ for $k < m$. In this appendix we provide the exact likelihood expression: extension of (5) and (4) for the case where one adds $m > 1$ links simultaneously.

Let $\mathcal{D}_{m,N}$ denote the sequence of degrees of m nodes chosen from among the nodes of the existing graph of N nodes. The degrees of these m nodes in N are denoted as below, and q_k here denotes the number of nodes in N having degree k :

$$\mathcal{D}_{m,N} = \{\underbrace{k_1, k_1, \dots, k_1}_{q_{k_1}}, \underbrace{k_2, \dots, k_2}_{q_{k_2}}, \dots, \underbrace{k_s, \dots, k_s}_{q_{k_s}}\}, \quad (\text{B1})$$

where $k_1 < k_2 < \dots < k_s$ and $\sum_{i=1}^s q_{k_i} = m$. From the main text, we know that \mathcal{Q}_{node} can be described as

$$\mathcal{Q}_{node} = \frac{P(G_{N+1})}{(N+1)\tilde{p}P(G_N)}.$$

However, the expression for connecting simultaneously to m nodes is far less straightforward than the case (4) for $m = 1$. In order to provide a more tractable formula, we need to change our notation. Let $l_1 = k_1$. If $k_2 = k_1 + 1$, then $l_1 + 1 = k_2$; if

not, let $l_2 = k_2$. Hence we can rewrite $\mathcal{D}_{m,N}$ as

$$\mathcal{D}_{m,N} = \{l_1, l_1 + 1, \dots, l_1 + b_1, l_2, l_2 + 1, \dots, l_2 + b_2, \dots, l_t, l_t + 1, \dots, l_t + b_t, \}, \quad (\text{B2})$$

where $k_1 = l_1, \dots, k_{b_1+1} = l_1 + b_1$ and $k_{(\sum_{i=1}^{t-1} b_i)+t-1} = l_t, \dots, k_m = l_t + b_t$. The sequence on the right-hand side of (B2) is a complete list of available node degrees (not counting multiplicities) parameterized under the l_i and b_j .

Using this new notation, the same counting argument as described in the main text for $m = 1$ now yields

$$Q_{\text{node}-m}(\mathcal{D}_{m,N}) = \begin{cases} \frac{p_m}{n_m+1} \frac{1}{\bar{p}} \prod_{i=1}^t \frac{p_{l_i+b_i+1}^{q_{l_i+b_i+1}}}{p_{l_i}^{q_{l_i}}} \frac{n_{l_i+b_i+1}!}{(n_{l_i+b_i+1}+q_{l_i+b_i})!} \frac{n_{l_i}!}{(n_{l_i}-q_{l_i})!} \times \dots & l_1 > m \\ \prod_{j=1}^{b_i} p_{l_i+j}^{q_{l_i+j-1}-q_{l_i+j}} \frac{n_{l_i+j}!}{(n_{l_i+j}-q_{l_i+j}+q_{l_i+j-1})!} & \\ \frac{p_m}{(n_m+1-q_m)} \frac{1}{\bar{p}} \prod_{i=1}^t \frac{p_{l_i+b_i+1}^{q_{l_i+b_i+1}}}{p_{l_i}^{q_{l_i}}} \frac{n_{l_i+b_i+1}!}{(n_{l_i+b_i+1}+q_{l_i+b_i})!} \frac{n_{l_i}!}{(n_{l_i}-q_{l_i})!} \times \dots & l_1 = m \\ \prod_{j=1}^{b_i} p_{l_i+j}^{q_{l_i+j-1}-q_{l_i+j}} \frac{n_{l_i+j}!}{(n_{l_i+j}-q_{l_i+j}+q_{l_i+j-1})!} & \end{cases}. \quad (\text{B3})$$

There are several things we need to notice. First, when $m = 1$, formula (B3) is the same as the formula (4) given in the main text. Second, when we add a new edge, $Q_{\text{edge}-j}$ is identical to (5). Finally, when a new node added, it connects to the

m nodes at the same time. The following is not allowed: if $n_k = 0$, and $n_{k-1} > 0$, the new node is connected to a node with degree $k - 1$ and then connected to the same node which now has degree k . So if $n_k = 0$, then $k \notin \mathcal{D}_{m,N}$.

-
- [1] M. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).
- [2] A. Barabási and R. Albert, *Science* **286**, 509 (1999).
- [3] D. S. Callaway, J. E. Hopcroft, J. M. Kleinberg, M. E. J. Newman, and S. H. Strogatz, *Phys. Rev. E* **64**, 041902 (2001).
- [4] A. Bekessy, P. Bekessy, and J. Komlos, *Studia scientiarum mathematicarum Hungarica* **7**, 343 (1972).
- [5] K. Judd, M. Small, and T. Stemler, *Europhys. Lett.* **103**, 58004 (2013).
- [6] M. Small, D. M. Walker, and C. K. Tse, *Phys. Rev. Lett.* **99**, 188702 (2007).
- [7] R. Lambiotte and M. Ausloos, *Phys. Rev. E* **72**, 066107 (2005).
- [8] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Åberg, *Nature (London)* **411**, 907 (2001).
- [9] C. I. Del Genio, T. Gross, and K. E. Bassler, *Phys. Rev. Lett.* **107**, 178701 (2011).
- [10] L. Zhang, M. Small, and K. Judd, [arXiv:1309.0961v2](https://arxiv.org/abs/1309.0961v2).
- [11] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, *Phys. Rev. Lett.* **85**, 4633 (2000).
- [12] S. N. Dorogovtsev and J. F. F. Mendes, *Europhys. Lett.* **50**, 1 (2000).
- [13] G. Bianconi, *Europhys. Lett.* **81**, 28005 (2008).
- [14] G. Bianconi, *Phys. Rev. E* **79**, 036114 (2009).
- [15] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, *Eur. Phys. J.* **B 38**, 205 (2004).
- [16] M. Small, in *IEEE International Symposium on Circuits and Systems Proceedings* (IEEE, Piscataway, 2013), pp. 2509–2512.
- [17] H. Zhang, J. Zhang, C. Zhou, M. Small, and B.-H. Wang, *New J. Phys.* **12**, 023015 (2010).
- [18] X.-K. Xu, J. Zhang, and M. Small, *Phys. Rev. E* **82**, 046117 (2010).
- [19] V. Colizza, A. Flammini, M. Serrano, and A. Vespignani, *Nature Phys.* **2**, 110 (2006).
- [20] M. Small, X. Xu, J. Zhou, J. Zhang, J. Sun, and J.-a. Lu, *Phys. Rev. E* **77**, 066112 (2008).
- [21] M. Catanzaro, M. Boguná, and R. Pastor-Satorras, *Phys. Rev. E* **71**, 027103 (2005).
- [22] Y. Zou, T. Pereira, M. Small, Z. Liu, and J. Kurths, *Phys. Rev. Lett.* **112**, 114102 (2014).
- [23] R. Albert and A.-L. Barabási, *Rev. Modern Phys.* **74**, 47 (2002).
- [24] That is, exactly a power law for degree $k \geq m$ where $m \geq 1$ is the number of edges added with each new node.
- [25] M. Newman, *Contemp. Phys.* **46**, 323 (2005).