

## Socioeconomic networks with long-range interactions

Rui Carvalho\*

Centre for Advanced Spatial Analysis, University College London, 1-19 Torrington Place, London WC1E 6BT, United Kingdom

Giulia Iori†

Department of Economics, School of Social Science, City University, Northampton Square, London EC1V 0HB, United Kingdom

(Received 4 June 2007; revised manuscript received 11 March 2008; published 22 July 2008)

We study a modified version of a model previously proposed by Jackson and Wolinsky to account for communication of information and allocation of goods in socioeconomic networks. In the model, the utility function of each node is given by a weighted sum of contributions from all accessible nodes. The weights, parametrized by the variable  $\delta$ , decrease with distance. We introduce a growth mechanism where new nodes attach to the existing network preferentially by utility. By increasing  $\delta$ , the network structure evolves from a power-law to an exponential degree distribution, passing through a regime characterized by shorter average path length, lower degree assortativity, and higher central point dominance. In the second part of the paper we compare different network structures in terms of the average utility received by each node. We show that power-law networks provide higher average utility than Poisson random networks. This provides a possible justification for the ubiquitousness of scale-free networks in the real world.

DOI: [10.1103/PhysRevE.78.016110](https://doi.org/10.1103/PhysRevE.78.016110)

PACS number(s): 89.65.Gh, 89.65.-s, 89.75.-k, 87.23.Ge

### I. INTRODUCTION

The study of socioeconomic networks is a burgeoning field in the physics and economics literature, with major progress having been attained over the last decade [1–6]. Individuals and firms interact through networks to share information and resources, exchange goods and credit, make new friendships or partnerships, etc. The structure of the network through which interactions take place may thus have an important effect on the success of the individual or the productivity of the firm [1]. Furthermore, the network of interactions among socioeconomic agents plays an important role for the stability and efficiency of socioeconomic systems [7]. Theories about how interaction networks form are thus essential for a deeper understanding of the development and organization of society as a whole.

The economics literature focuses mainly on equilibrium networks and the network formation mechanisms are based on utility maximization and costs minimization. The aim of most economic papers is to identify, among the set of equilibrium networks, the geometry that optimizes efficiency<sup>1</sup> in the sense of social benefit. Likewise, economists are interested in the stability<sup>2</sup> of equilibrium networks under link deletion, addition, or rewiring [1,2]. A shortcoming of these models is that the equilibrium networks are often too simple in their geometry (stars, complete networks, interlinked stars, etc.), typically as a consequence of the symmetries that need

to be assumed in the payoff functions in order to make the models analytically tractable [8]. The physics literature, instead, has mainly focused on the characterization of the structure of real networks and proposed dynamic models, mostly based on probabilistic rules, capable of reproducing the observed geometrical structures (Poisson, stretched exponential, and scale-free networks) [9–11].

In this paper we try to combine the physics and economic approaches, by introducing a stochastic network formation mechanism inspired by economists' utility maximization models, which naturally extends the well known physicists' preferential attachment rule [12]. One of the most interesting models of socioeconomic network formation was introduced by Jackson and Wolinsky in 1996 [1]. In their model, the formation and evolution of links is driven by a utility maximization mechanism. The model is based on the assumption that agents may derive benefit not only from the nodes to which they are directly connected (their nearest neighbors), but also from the ones they are connected to indirectly (possibly via long paths). Less distant connections are more valuable than more distant ones, but connections to the nearest neighbors are costly. The utility of node  $i$  is defined as

$$u_i = w_{ii} + \sum_{j \neq i} w_{ij} \delta^{d_{ij}} - \sum_{j \in \mathcal{V}(i)} c_{ij}, \quad (1)$$

where the contribution to the utility of  $i$  from  $j$  may depend on the weight  $w_{ij}$  of the edge between  $i$  and  $j$  (or, alternatively, on the fitness of node  $j$ );  $0 \leq \delta < 1$  captures the idea that the utility gain from indirect connections decreases with distance;  $d_{ij}$  is the number of links in the shortest path between  $i$  and  $j$  ( $d_{ij} = \infty$  if there is no path between  $i$  and  $j$ );  $\mathcal{V}(i)$  is the set of nearest neighbors of  $i$ ; and  $c_{ij}$  are the (node specific) costs to establish a directed connection between  $i$

\*rui@ruicarvalho.org

†g.iori@city.ac.uk

<sup>1</sup>A network  $g$  is efficient with respect to an aggregate utility measure  $u$  if  $u(g) \geq u(g') \forall g' \in \mathcal{G}$  [1].

<sup>2</sup>A network is pairwise stable when no node would benefit from severing an existing link, and no two nodes would benefit from forming a new link [1]. Pairwise stability is stronger than Nash equilibrium and is aimed at sequential updating.

and  $j$ .<sup>3</sup> Costs can also be differentiated in costs of initially creating or maintaining an edge [2].

The papers by Jackson and Wolinsky [1], as well as the one by Bala and Goyal [2], are mainly concerned with stability and efficiency of the network resulting from different dynamic updating rules. In particular, Jackson and Wolinsky study pairwise stability when agents can update only a link at a time (either deleting or creating it), while Bala and Goyal allow agents to rearrange all their connections at once. The updating is deterministic in both models, and a new configuration is accepted only if it increases the utility of the agent. These two papers show that the star network is both efficient and stable for a wide range of the parameters when  $\delta=1$ . Nonetheless, a multiplicity of network architectures exist in [2] for  $0 < \delta < 1$  which could be a strict Nash equilibrium, and to which the system may converge depending of the initial conditions. Feri [13] has shown that for sufficiently large networks the star network is stochastically stable for almost all the range of parameters, even for  $0 < \delta < 1$ .

Here we focus on the connections model of Jackson and Wolinsky, i.e., the case  $w_{ij}=1$ ,  $w_{ii}=0$ , and  $c_{ij}=c$ . In this case, the utility can be written as

$$u_i = \sum_{l=1}^{j_{\max}^{(i)}} \sum_{\{k|d_{ik}=l\}} \delta^l - \sum_{j \in \mathcal{V}(i)} c = \sum_{l=1}^{j_{\max}^{(i)}} \delta^l z_l^{(i)} - cz_1^{(i)}, \quad (2)$$

where the sum in  $l$  is over all shortest paths of length  $l$  from node  $i$ , the sum in  $k$  is over all nodes whose shortest path from  $i$  is  $d_{ik}=l$ ,  $j_{\max}^{(i)}$  is the path length of the node the furthest away from node  $i$ , and  $z_l^{(i)}$  is the number of  $l$ th-nearest neighbors of node  $i$ . The utility of a node is expressed in Eq. (2) as a weighted sum of the number of nodes accessible from  $i$  on outward “layers” of increasing distance from  $i$ . Thus, we start at node  $i$  and multiply  $\delta$  by the number of nodes that are joined by an edge to  $i$ —this being the first layer. We then add  $\delta^2$  times the number of nodes that are joined by an edge to a node in the first layer—this is the second layer. We continue in this way until no new nodes are found. Hence, expression (2) incorporates implicitly the well known breath-first search algorithm [14]. In this paper, we will consider only networks with zero costs. Therefore, Eq. (2) becomes

$$u_i = \sum_{l=1}^{j_{\max}^{(i)}} \delta^l z_l^{(i)}. \quad (3)$$

In the first part of the paper we focus on a specific network growth mechanism and examine the resulting network topology. If each new node attached deterministically to the existing node with maximal utility, the resulting network would be a star. The randomness generated by a probabilistic attaching rule can be interpreted as costs and barriers to gather information, or bounded rationality, all of which limit the ability to establish links in an optimal way, thus possibly generating more realistic geometries than the star network. It

is thus worthwhile to ask which network topologies are to be found when new nodes arrive steadily and create links with existing nodes in a probabilistic way, proportionally to the utility of existing nodes. In this way, we build on the preferential attachment growth rule of Barabási and Albert [9,12] which can be recovered from Eq. (3) when  $l_{\max}^{(i)}=1$ . Furthermore, preferential attachment is, arguably, the most extensively studied mechanism of network formation and one that has revealed insights into properties observed in real networks. Therefore, it is important to understand the robustness of the specific rule of linear preferential attachment by node degree, which is one of the aims of this paper.

Often, the specific network growth mechanisms are unknown and only the topology of the equilibrium network can be extracted from data. One obvious question is then how the observed equilibrium networks rank in terms of their efficiency, e.g., Erdős-Rényi versus scale-free networks. We address this question in the second part of the paper and derive analytical results by using the generating function approach [15]. We show that power-law networks are more efficient than Poisson random network when individual utility is defined by Eq. (3), thus providing a possible explanation for why scale-free networks are so ubiquitous.

## II. GROWING NETWORKS

In the classic Barabási and Albert model [12], a network is grown by adding, at every time step, a new node that attaches to  $m$  existing nodes with a probability proportional to their degree,  $\Pi(k_i)=k_i/\sum_{j=1}^N k_j$ . At time  $t$ , the resulting network has size  $N_t=m_0+t$ , where  $m_0 \geq m$  is the size of the (fully connected) network at time  $t=0$ . Preferential attachment models were in fact introduced in the literature already by Yule [16] and Simon [17] and are a generalization of the Polya urn scheme.

The preferential attachment mechanism generates a scale-free probability density of incoming links that leads to the stationary result  $p(k)=2m^2/k^\gamma$ , with  $\gamma=3$  independently of  $m$ . The model is also characterized by a clustering coefficient larger than the one found for the Erdős-Rényi networks (for  $m > 1$ ) and no clear assortative or disassortative behavior [9].

Several models have been proposed lately to investigate extensions of the preferential attachment mechanism through edge removal and rewiring, inheritance, redirection or copying; node competition, aging and capacity constraints; and accelerated growth of networks to name just a few (see [9,10,18,19] for reviews). Of particular relevance to our approach are fitness models [20–23], where the probability of attaching to a node is proportional to the node fitness,

$$\Pi(k_i) \sim \frac{f_i k_i}{N \sum_{j=1} f_j k_j}. \quad (4)$$

Here we extend the preferential attachment rule by introducing a growing mechanism inspired by the work of Jackson and Wolinsky [1]. Our contribution is to propose preferential attachment by node utility. Thus, the probability that a new node  $j$  will be connected to an existing node  $i$  depends on the utility of  $i$ , such that

<sup>3</sup>In [1] costs are assumed to be equally, or cooperatively, shared by  $i$  and  $j$ , but extensions to the noncooperative case have also been explored.

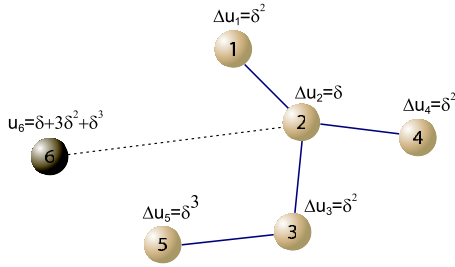


FIG. 1. (Color online) Schematic layout of network growth when  $m=m_0=1$ . The addition of a new node, 6, implies an increase of the utility of nodes 1–5 which is simply  $\delta^l$ , where  $l$  is the path length from node 6. The simplicity of this updating mechanism allowed simulations to be run with  $N=10^5$  when  $m=1$ .

$$\Pi_i = \frac{u_i}{\sum_{k=1}^N u_k} \quad (5)$$

where the utility of node  $i$ ,  $u_i$ , is given by Eq. (3). Attachment happens with uniform distribution for  $\delta=1$  generating an exponential distribution of node degree.<sup>4</sup> While the model is not defined for  $\delta=0$ , the preferential attachment rule (5) is invariant up to multiplicative factors in Eq. (3), so for  $\delta \neq 0$  the qualitative behavior of the model remains unchanged if we define utility as

$$u'_i = \frac{u_i}{\delta} = k_i + \sum_{l=2}^{l_{\max}^{(i)}} \delta^{l-1} z_l^{(i)}, \quad (6)$$

where  $k_i$  is the degree of node  $i$ . Thus, as  $\delta \rightarrow 0$  our model converges to the Barabási-Albert model and generates a scale-free network.

Our model has resemblances with the fitness models discussed above. However, there is a fundamental discrepancy: we regard utility as a time-dependent measure of node fitness, whereas existing models assume that node fitness does not change with time.

At each time step, a new node  $j$  joins the network and the utility of existing nodes changes. When  $m=1$ , the utility increment to an existing node  $i$  at distance  $l$  from  $j$  is given by  $\Delta u_i = \delta^l$  and therefore, at each time step, the computation of utility for the network can be completed in  $O(N)$  time. Figure 1 is a diagram of a possible network configuration with  $m=m_0=1$  after  $t=5$  time steps, showing the change in utility of existing nodes,  $\Delta u_i = \delta^l$ . When  $m > 1$ , the increment in the utility of node  $i$  depends on the existing network geometry and  $\Delta u_i \geq \delta^l$ . Therefore, when  $m > 1$ , we need to recompute the utility of all existing nodes at every time step, and the computation runs in  $O(N^2)$  time as it involves running a *breadth-first search* algorithm from every node. This is the reason that we ran simulations for  $N=10^5$  when  $m=1$ , but only up to  $N=5 \times 10^3$  when  $m > 1$ .

Existing nodes  $i$  at a higher distance than a certain  $l_{\max}^{(j)}$  from new node  $j$  receive a contribution  $\Delta u_i = \delta^{l_{\max}^{(j,i)}}$

$< 10^{-(\text{precision})}$ , which is less than the number of significant digits that the computer can store (typically the precision is 32 in double precision), and do not need to have their utility updated in the simulations. This maximal distance  $l_{\max}$  is defined as

$$10^{-(\text{precision})} > \delta^{l_{\max}} \Leftrightarrow l_{\max} > -\frac{(\text{precision})}{\log_{10} \delta}. \quad (7)$$

Our implementation of the algorithm updates the utility of all nodes accessible from the new node  $j$  up to distance  $l_{\max} = -32/\log_{10} \delta$ . The code was implemented in C++ and ran on a Condor framework (high-throughput computing) [24] for several values of  $\delta$ . Ensemble averages were taken over 30 runs.

Expressions (3) and (6) predict the existence of two distinct regimes: a scale-free regime as  $\delta \rightarrow 0$  ( $\delta \neq 0$ ) and a random growth regime for which the degree distribution is exponential at  $\delta=1$ . We are interested in exploring how the network evolves from one limiting regime to the other as we increase  $\delta$ .

Figures 2(a)–2(c) show the distribution of degree for  $m=1, 2$ , and 5. We also plot the corresponding distribution for the Barabási-Albert (BA) model (solid curves shifted vertically). For very small  $\delta \sim 0.01$ , preferential attachment by degree is indistinguishable from preferential attachment by utility and the probability distributions of both quantities decay as  $p(x) \sim x^{-\gamma}$  with  $\gamma=3$ . The power-law decay in the BA model is known to be a peculiarity of the linear preferential attachment mechanism and is destroyed by small perturbations like, for example, a nonlinear attachment rule  $\Pi(k_i) \sim k_i^\alpha$  [12]. Here we also observe a departure from the scale-free regime as we increase  $\delta$ . Furthermore, in the Barabási-Albert model the degree distribution decays as a power law with exponent  $\gamma=3$  independently of  $m$ . In our model, increasing  $m$  has the effect of homogenizing the utility of the nodes (the distance between pairs of nodes decreases with increasing connectivity in the networks). Consequently, deviations from the power-law decay are observed at lower and lower values of  $\delta$  as we increase  $m$ .

Betweenness centrality is plotted in Figs. 2(g)–2(i) as  $m$  is varied. Recent results have shown that the distribution of loads (or betweenness) scales with a power law [25,26]  $p(g) \sim g^{-\alpha}$  where  $\alpha=2$  for a tree (and hence for  $m=1$ ). This justifies the collapse of the curves of the distribution of betweenness in Fig. 2(g). As can be observed in Figs. 2(h) and 2(i), the distribution of betweenness deviates from the power-law behavior as  $m$  is increased.

For intermediate values of  $\delta$  and  $m=1$  a number of interesting features appear. First of all, we observe that the utility distribution becomes steplike [Fig. 2(d)] suggesting the presence of subsets of nodes that share similar utilities. This phenomenon can also be inferred from the network layouts in Fig. 3 (for networks of  $10^3$  nodes with  $m=1$ ), which are produced using the Kamada-Kawai spring layout [27]. Essentially, the Kamada-Kawai layout assigns stronger springs to vertices that are closer in the graph-theoretic sense (i.e., by following edges) and therefore places them closer together. In the case  $m=1$  (a tree), nodes close to the hubs on the

<sup>4</sup>This is model A in [12].

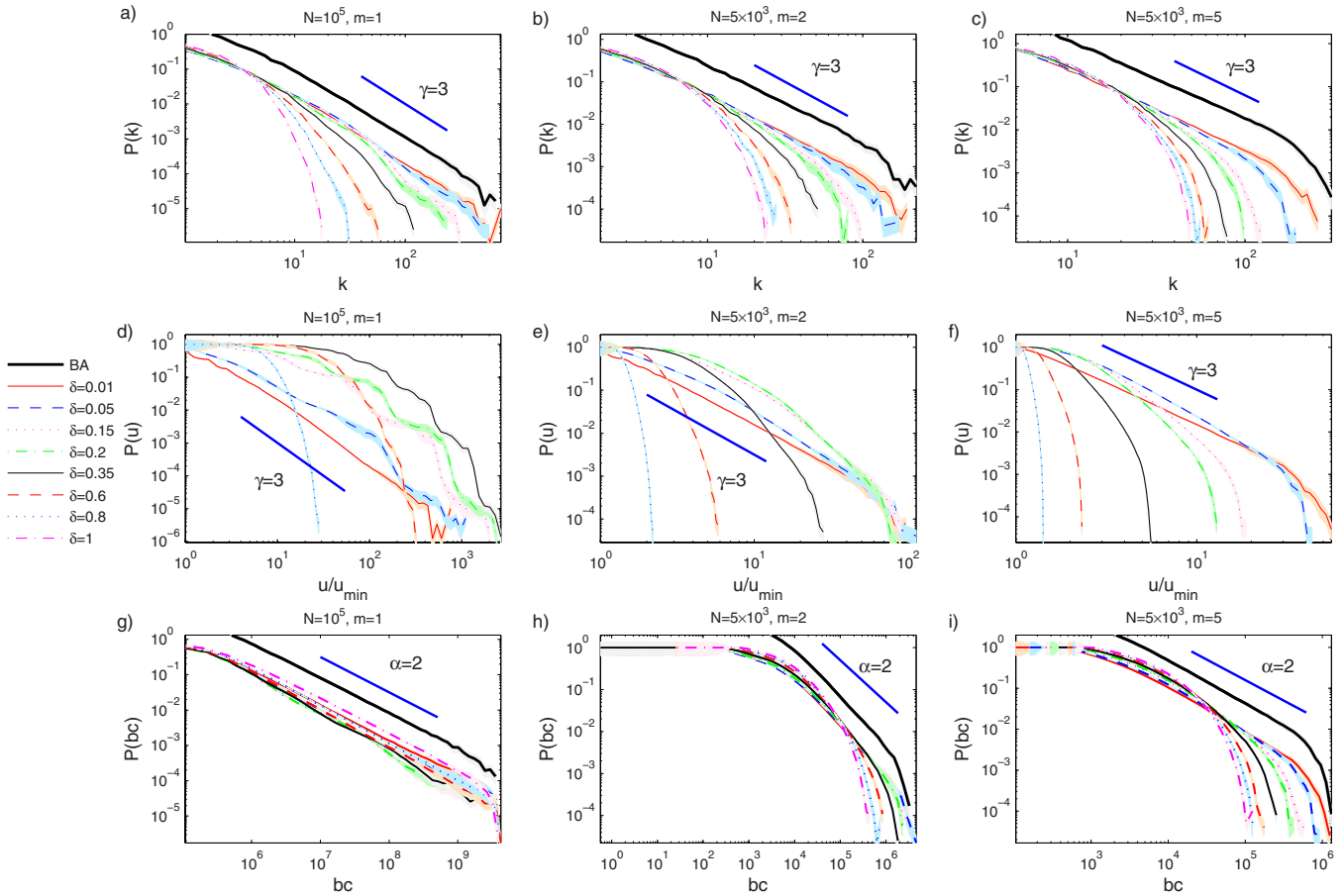


FIG. 2. (Color online) Cumulative distribution function of degree (a)–(c), utility (d)–(f), and betweenness centrality (g)–(i) for several values of  $\delta \in ]0, 1]$ , and  $m=1, 2$ , and  $5$ . We also plot the corresponding distribution of degree and betweenness for the BA model (curves were shifted vertically). Simulations were averaged over 30 runs in networks with  $N=10^5$  ( $m=1$ ) or  $N=5 \times 10^3$  ( $m=2$  and  $5$ ). Colored bands around the curves are 95% confidence intervals.

graph layout will also be close in graph terms and therefore we can interpret the layout heuristically: these nodes have similar utility. When  $\delta=0.01$  (close to the BA scale-free regime), the layout shows a few utility hubs (the dark vertices) surrounded by clouds of nodes that disperse as we move further away from the hubs; for  $\delta=0.2$  denser clouds of nodes cluster around a smaller number of hubs, and can still be observed farther away from the hub; for higher  $\delta$  the clouds start dispersing; eventually for  $\delta=1$  all nodes have the same utility.

The rearrangement of the network as  $\delta$  increases from zero gives rise to non monotonous behavior of the average path length  $\bar{l}$ , degree assortativity  $r_{\text{degree}}$  [28], and central point dominance  $C'_B$  [29]. Both  $\bar{l}$  and  $r_{\text{degree}}$  show a minimum for  $\delta \sim 0.15$  ( $N=10^5$  and  $m=1$ ) and  $C'_B$  has a maximum around the same value. This behavior can be observed in Figs. 4(a)–4(c). The average path length  $\bar{l}(\delta)$  is measured relative to the path length of a random growing network [i.e., relative to  $\bar{l}(1)$ ]. The scale-free regime is characterized by a shorter path length than the random growth regime. Here we observe an even further contraction of the network for values of  $\delta$  up to  $\sim 0.5$ . Note that the average path length of a star network is  $l^*=2$  for large  $N$ . When normalizing with  $l(1)$

$=20.16$  we get a value of  $l^* \sim 0.1$ . While this value is still much smaller than  $\bar{l}(0.15)$ , the network contraction seems to indicate a move toward a more starlike configuration. To further investigate this point, we compute the central point dominance measure introduced by Freeman [29] and plot it in Fig. 4(b). This measure is defined as the average difference in node centrality (measured by node total betweenness) between the most central point and all the others. The central point dominance takes a value between zero (for a graph where all points have the same centrality) and one (for the wheel or star graph). The maximum of  $C'_B$  for  $\delta \sim 0.15$  in Fig. 4(b) confirms that the network is becoming more starlike around these values of  $\delta$ .

Next, we plot the assortativity of the network in Fig. 4(c). We implement as measure of assortativity the degree assortativity  $r_{\text{degree}}$  [28] which takes values from  $-1$  to  $+1$ : negative values for disassortative networks, 0 if the networks are neither disassortative or assortative, and  $+1$  for fully assortative networks. This value approaches zero for large  $N$  in the BA model [30] and is negative for a star. Our model shows a lower assortative mixing than the BA model for values of  $\delta$  up to  $\sim 0.5$ . The decrease in  $r_{\text{degree}}$  is also consistent with our hypothesis that the network is becoming more star like at  $\delta \sim 0.15$ . While the network goes through this rearrangement

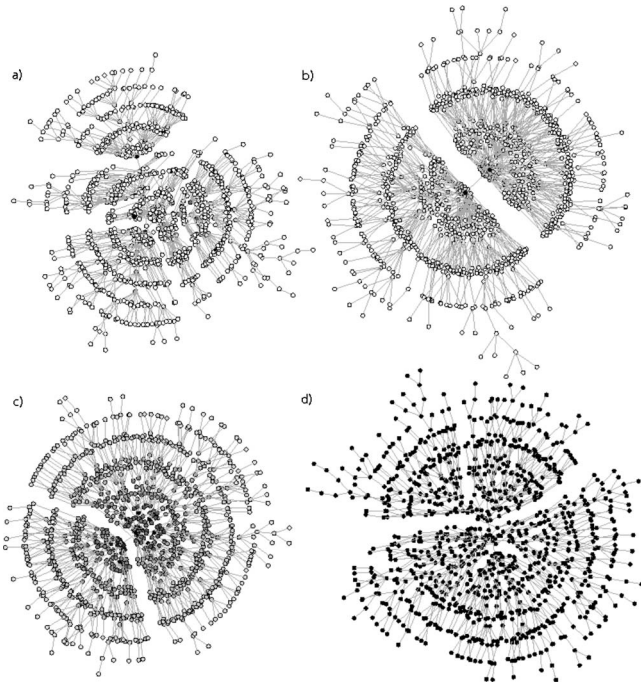


FIG. 3. Kamada-Kawai spring layout [27] for  $m=1$  and  $N = 10^3$ . Sample layout for  $\delta =$  (a) 0.01, (b) 0.2, (c) 0.7, and (d) 1. On each panel, nodes are colored by their utility on a gray scale from minimal (white) to maximal (black) utility.

the degree of the most connected node (as a fraction of the total number of links) is nonetheless monotonically decreasing with  $\delta$ , as shown in Fig. 4(d). The same behavior is observed for the utility of the most connected node (not shown here). This reveals that, as new nodes are added to the network, they do attach on average closer to the hubs as  $\delta$  increases, generating a more compact network, but not directly to them.

To gain further insights into the structural changes that take place as  $\delta$  increases from zero, we analyze the role of entry time on node connectivity by computing the ratios  $u/u_{\max}$  and  $k/k_{\max}$  for the first 100 nodes as a function of the time at which they entered the network. The ratios plotted in Fig. 5 are averaged over 30 different simulations, for networks with  $N=10^5$  and  $m=1$ . The plot shows that for any  $\delta$  the initial node is likely to acquire the highest fraction of links and utility. Moreover, for  $\delta=0.15$  both the degree and utility ratios decay faster than in the scale-free regime. This reveals that, for  $\delta \sim 0.15$ , the earlier nodes receive both a higher relative degree and utility than in the scale-free case. In other words, the earlier nodes are stronger hubs than in the scale-free case and thus the network arranges in a more starlike configuration around  $\delta=0.15$ . As  $\delta$  increases further, the slope of the utility ratio becomes lower than in the scale-free regime. In this range, the utility differences between old and new nodes are not large enough to create well-defined utility or degree hubs.

Figure 5 also highlights the fundamental mechanism of structure formation in our model: the fine relation between degree and utility as  $\delta$  is varied. In the scale-free regime, preferential attachment by utility is equivalent to preferential attachment by degree and node degree and utility assume the

same value as the network grows. At  $\delta \sim 0.15$ , we observe a gap between the two scaled quantities, suggesting a discrepancy between degree and utility for the higher-order neighbors of the utility hubs. This gap is larger for  $\delta=0.15$  than  $\delta=0.4$ , indicating that around  $\delta \sim 0.15$  the growth mechanism is generated by a variable (node utility) which is considerably independent of node degree and thus revealing why this is the region where the network displays more interesting structure. As  $\delta$  increases toward 1, the influence of random network growth becomes more important and this structure-generating mechanism disappears.

Finally, we investigate how average node utility compares in networks generated with our preferential utility attachment, the scale-free regime (here generated via the BA preferential attachment mechanism) and a star network. The average utility of a star network is given by

$$\bar{u}_*(\delta) = \delta z_1 \left( 1 + \delta \frac{N-2}{2} \right), \tag{8}$$

where  $z_1 = 2(N-1)/N$ . For  $N$  large,  $z_1 \approx 2$  and  $\bar{u}_*(\delta) \sim N\delta^2$ . In Fig. 6 we plot the average utility for networks in our model at different  $\delta$  when  $N=5 \times 10^3$  and  $z_1=2, 4$ , and 10 and compare that to the corresponding average utility in the BA model and a star with the same  $N$ . The plot shows that the BA scale-free network has a higher utility than the network generated via our preferential utility mechanism at all values of  $\delta$  and  $z_1$ . Networks in our model become more starlike for  $\delta \sim 0.5$ , but this implies an increase in the utility of only a small number of nodes (the early nodes). Therefore, the average utility of nodes in the network is still higher in the BA model than in our model for the same values of  $\delta$  due to the scale-free structure of the former. Comparisons with the star network can only be made when  $z_1=2$  as this is the average degree of the star network when  $N$  is large. Figure 6 confirms that the star network has the highest utility for this value of  $z_1$  among all the networks we study. Nevertheless, the star network can only be achieved if agents are perfectly rational and have access to full information (in which case the attachment mechanism would be deterministic). This is rarely the case in a real word situation; thus the comparison with the star is of little practical relevance.

### III. ANALYTICAL RESULTS FOR RANDOM NETWORKS

An interesting question to ask (for example, from the point of view of the social planner) is how network topologies rank against each other and which network structure maximizes the total or the average utility (networks that satisfy this condition are said to be efficient in economics). We show that it is possible to derive analytical results for the average utility in Poisson and power-law networks. By comparing average utility in different network topologies with the same size and the same average degree, we show that power-law networks are more efficient than Poisson random networks (even though less efficient than the star). The effect of costs on  $\bar{u}(\delta)$  is a constant term for all networks with the same  $N$  and  $z_1$ . Therefore, without loss of generality, we choose  $c=0$  in the analysis below.

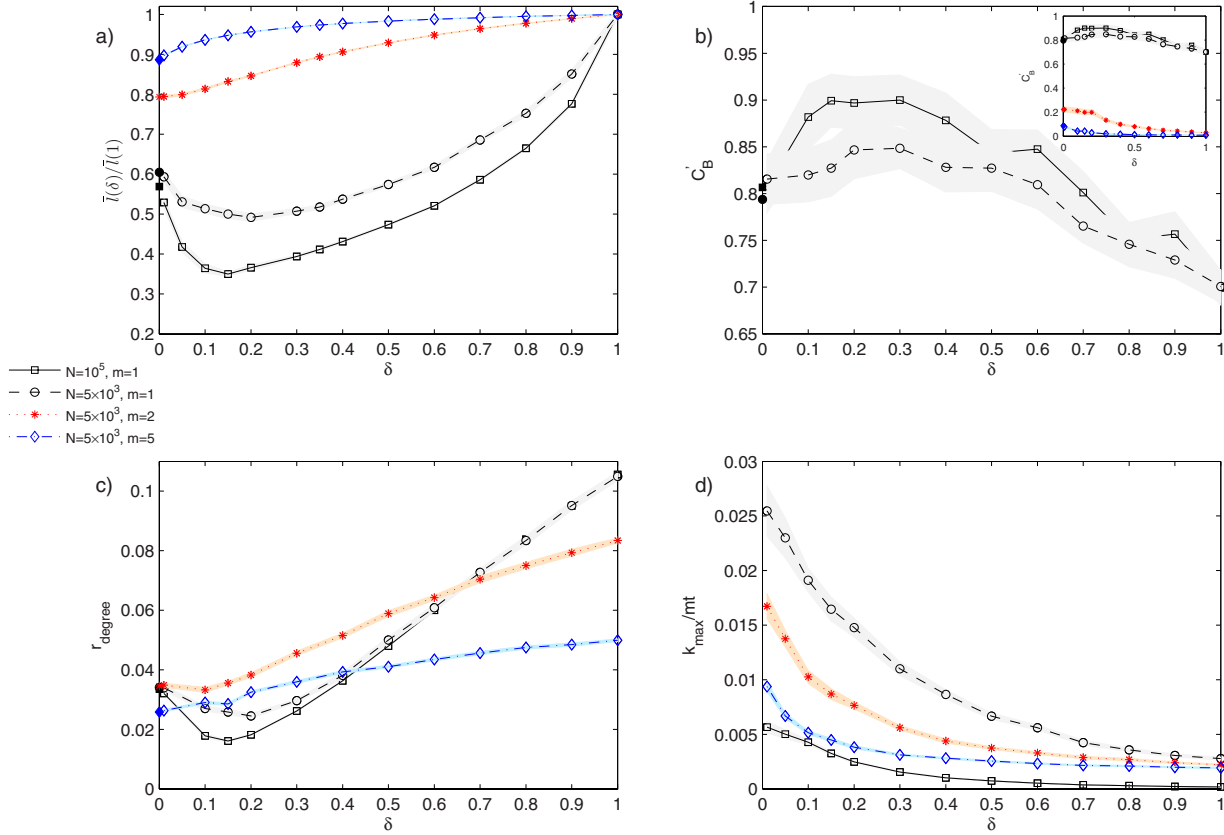


FIG. 4. (Color online) Plots of (a) average path length, (b) central point dominance, (c) assortativity coefficient, and (d) normalized maximum degree as a function of  $\delta$  for the simulation results when  $m=1, 2$ , and  $5$ . Curves in panel (a) were scaled by the values at  $\delta=1$  (exponential network). The BA values are indicated by the corresponding full symbols. Simulations were averaged over 30 runs and colored bands around the curves are 95% confidence intervals.

If the sum in Eq. (3) was to be evaluated up to distance  $l_{\max}^{(i)}=1$  for every node, expression (9) would simplify to  $\bar{u}(\delta)=\delta z_1$ , i.e., average utility would be independent of the specific network topology and all networks with the same number of nodes and links would be equally efficient. Thus we need to introduce long-range interactions ( $l_{\max}^{(i)} > 1$ ) to be able to rank networks in terms of their efficiency.

To derive an expression for average utility in generic random networks with  $N$  large, we average both sides of Eq. (3):

$$\bar{u}(\delta) = \sum_{l=1}^{\bar{l}} \delta^l z_l, \tag{9}$$

where  $z_l$  is the average number of  $l$ th neighbors of a node. Newman *et al.* [31,32] define  $\bar{l}$  via the expression

$$1 + \sum_{l=1}^{\bar{l}} z_l = N. \tag{10}$$

Now that we have expressed average utility in terms of the breadth-first search algorithm, we can derive a closed form of expression (9) if we have access to analytical expressions for  $\bar{l}$  and  $z_l$ . This can be accomplished by generating functions, which are particularly useful when determining means, standard deviations, and moments of distributions [15].

The average number of neighbors (average degree) and the average number of second neighbors of a node can be derived from the probability generating function of node degree,  $G_0(x) = \sum_{k=0}^{\infty} p_k x^k$ , as long as the degree distribution  $p_k$  is specified. The beauty of the generating function formalism is

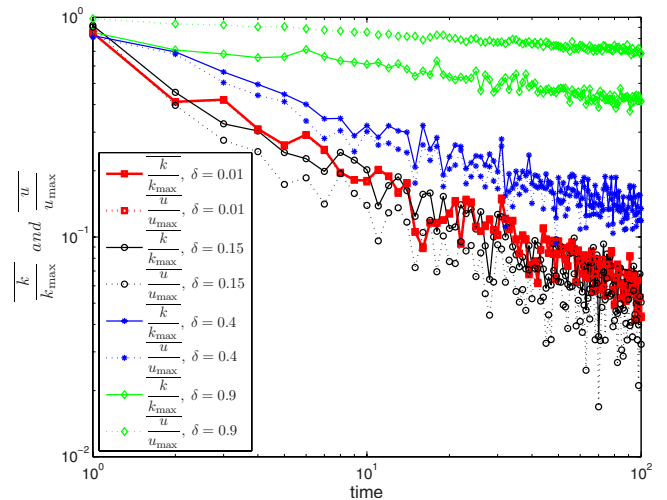


FIG. 5. (Color online) Plots of average  $k/k_{\max}$  and  $u/u_{\max}$  as a function of the entry time for the first 100 nodes when  $N=10^5$  and  $m=1$ . Simulations were averaged over 30 runs.

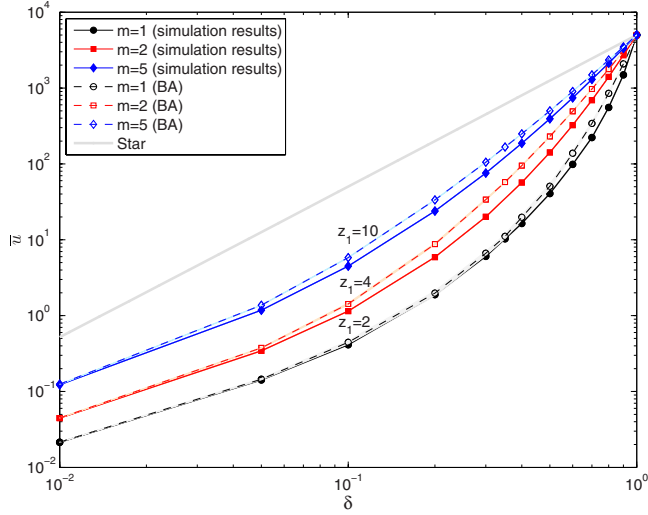


FIG. 6. (Color online) Average utility for the simulation results (solid curve and symbols) and the BA model (dashed curve, open symbols) for  $m=1$  ( $z_1=2$ ),  $2$  ( $z_1=4$ ), and  $5$  ( $z_1=10$ ) when  $N=5 \times 10^3$ . We also plot the average utility of a star with  $N=5 \times 10^3$  nodes. Simulation results were averaged over 30 runs. Colored bands around the curves are 95% confidence intervals.

that one can derive  $z_l$  as a function of  $z_1$  and  $z_2$  only [31–33]:

$$z_l = \left( \frac{z_2}{z_1} \right)^{l-1} z_1. \quad (11)$$

Replacing Eq. (11) in Eq. (9) yields

$$\bar{u}(\delta) = \delta z_1 \sum_{l=1}^{\bar{l}} (\delta Z)^{l-1} = \frac{\delta z_1 [(\delta Z)^{\bar{l}} - 1]}{\delta Z - 1}, \quad (12)$$

where  $Z = z_2/z_1$ . For  $Z > 1$  and  $N > z_1 + 1$ , which are conditions satisfied by most networks,  $\bar{l}$  can be calculated as a function of  $N$ ,  $z_1$ , and  $z_2$  from Eqs. (10) and (11) as [31]

$$\bar{l} = \frac{\ln[(N-1)(Z-1)/z_1 + 1]}{\ln(Z)}. \quad (13)$$

In what follows, we investigate the behavior of Eq. (12) for Poisson and power-law random networks.

Poisson random networks are characterized by  $z_1 = pN$  and  $z_2 = z_1^2$  [31], thus Eq. (13) yields  $\bar{l}_p = \ln(\frac{N(z_1-1)+1}{z_1}) / \ln(z_1)$ . In this case, Eq. (12) becomes

$$\bar{u}_p(N, \delta, z_1) = \frac{z_1 \delta [(\delta z_1)^{\ln(N+1-N/z_1)/\ln(z_1)} - 1]}{\delta z_1 - 1} \quad (14)$$

for  $N > z_1 + 1$ ,  $0 < \delta \leq 1$ , and  $z_1 > 1$ .

Next, we consider power-law networks with degree distribution of the form

$$p_k(\gamma, a) = \frac{1}{\zeta(\gamma, 1+a)} (a+k)^{-\gamma}, \quad a \geq 0, \quad (15)$$

where the normalizing factor  $\zeta(\gamma, a+1) = \sum_{k=0}^{\infty} (a+k)^{-\gamma}$  is the Hurwitz zeta function ( $\gamma > 1$ ). The generating function for the probability distribution is given by

$$G_0(x, \gamma, a) = \sum_{k=1}^{\infty} p_k x^k = \frac{x \Phi(x, \gamma, a+1)}{\zeta(\gamma, a+1)}, \quad (16)$$

where  $\Phi(x, \gamma, a) = \sum_{k=0}^{\infty} \frac{x^k}{(a+k)^\gamma}$  is the Lerch transcendent. For our purposes, only the first two derivatives of  $\Phi(x, \gamma, a+1)$  with respect to  $x$  are relevant, as the average number of first and second-neighbors are given, respectively, by  $z_1(\gamma, a) = \frac{\partial G_0(x)}{\partial x} \Big|_{x=1}$  and  $z_2(\gamma, a) = \frac{\partial^2 G_0(x)}{\partial x^2} \Big|_{x=1}$ . Hence

$$z_1(\gamma, a) = \frac{\Phi(1, \gamma-1, a+1) - a \Phi(1, \gamma, a+1)}{\zeta(\gamma, a+1)}, \quad (17)$$

$$\gamma > 2 \wedge a \geq 0,$$

$$z_2(\gamma, a) = \frac{\zeta(\gamma-1, a+1)}{\zeta(\gamma, a+1)} z_1(\gamma-1, a) - (a+1) z_1(\gamma, a), \quad (18)$$

$$\gamma > 3 \wedge a \geq 0.$$

Thus

$$Z(\gamma, a) = \frac{\zeta(\gamma-1, a+1)}{\zeta(\gamma, a+1)} \frac{z_1(\gamma-1, a)}{z_1(\gamma, a)} - a - 1, \quad \gamma > 3 \wedge a \geq 0. \quad (19)$$

Substituting Eqs. (17) and (19) into Eq. (13), we find

$$\begin{aligned} \bar{l}_{\text{SF}}(N, \gamma, a) &= \frac{\ln\left(-\frac{(a+2)(N-1)}{z_1(\gamma, a)} + \frac{z_1(\gamma-1, a) \zeta(\gamma-1, a+1)(N-1)}{z_1(\gamma, a)^2 \zeta(\gamma, a+1)} + 1\right)}{\ln\left(-a + \frac{z_1(\gamma-1, a) \zeta(\gamma-1, a+1)}{z_1(\gamma, a) \zeta(\gamma, a+1)} - 1\right)}, \end{aligned} \quad (20)$$

$$N > z_1(\gamma, a) + 1 \wedge \gamma > 3 \wedge a \geq 0 \wedge Z(\gamma, a) > 1,$$

and thus average utility is given by

$$\bar{u}_{\text{SF}}(N, \delta, \gamma, a) = \frac{\delta z_1(\gamma, a) \{[\delta Z(\gamma, a)]^{\bar{l}_{\text{SF}}(N, \gamma, a)} - 1\}}{\delta Z(\gamma, a) - 1}, \quad (21)$$

$$N > z_1(\gamma, a) + 1 \wedge 0 < \delta \leq 1 \wedge \gamma > 3 \wedge a \geq 0 \wedge Z(\gamma, a) > 1,$$

where  $z_1(\gamma, a)$ ,  $Z(\gamma, a)$ , and  $\bar{l}_{\text{SF}}(N, \gamma, a)$  are given by Eqs. (17), (19), and (20), respectively.

When  $a=0$ , the distribution of degree, Eq. (15), becomes a pure power law  $p_k(\gamma) = \frac{1}{\zeta(\gamma)} k^{-\gamma}$ . In this case, we have  $\zeta(\gamma, a+1)|_{a=0} = \zeta(\gamma)$  and  $\Phi(x, \gamma, a+1)|_{a=0} = \frac{\text{Li}_\gamma(x)}{x}$ ; therefore Eq. (16) becomes

$$G_0(x, \gamma) = \frac{\text{Li}_\gamma(x)}{\zeta(\gamma)}. \quad (22)$$

This generating function is also obtained for the power-law distribution with exponential cutoff, proposed in [34,35],  $p_k(\gamma, \kappa) = C k^{-\gamma} e^{-k/\kappa}$ , in the limit  $\kappa \rightarrow \infty$ .

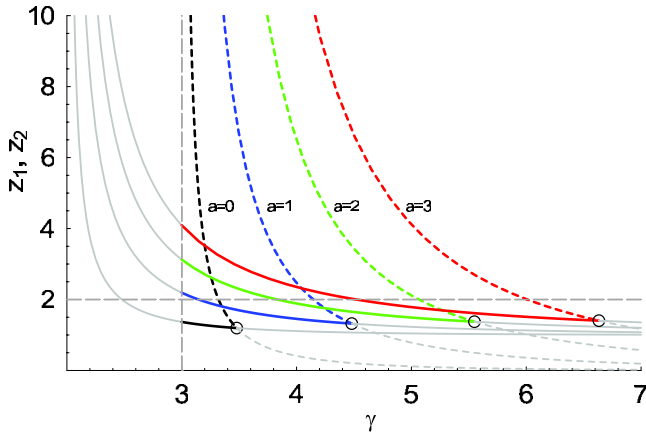


FIG. 7. (Color online) Average number of first and second neighbors  $[z_1(\gamma, a)$  and  $z_2(\gamma, a)]$  in networks with degree distribution given by Eq. (15). From left to right, we plot  $z_1(\gamma, a)$  (full curves) and  $z_2(\gamma, a)$  (dashed curves) for  $a=0$  (black), 1 (blue), 2 (green), and 3 (red). The values of  $z_1(\gamma \leq 3, a)$  is not defined are plotted in gray, as well as regions of  $\gamma$  for which  $z_1(\gamma, a) > z_2(\gamma, a)$ . The circles denote the intersection of the two curves,  $z_1(\gamma, a)$  and  $z_2(\gamma, a)$ , for each value of  $a$ .

Expression (22) implies that

$$z_1(\gamma)|_{a=0} = \frac{\zeta(\gamma-1)}{\zeta(\gamma)}, \quad \gamma > 2, \quad (23)$$

$$z_2(\gamma)|_{a=0} = \frac{\zeta(\gamma-2) - \zeta(\gamma-1)}{\zeta(\gamma)}, \quad \gamma > 3. \quad (24)$$

Therefore, in pure power-law networks, when  $N \rightarrow \infty$ , the average number of second neighbors,  $z_2(\gamma)$ , is finite only for  $\gamma > 3$ . However, the Riemann zeta function  $\zeta(\gamma)$  is a decreasing function of  $\gamma$  (for  $\gamma > 3$ ) and  $z_1(\gamma=3) = \pi^2/6\zeta(3) \approx 1.36843$ . In other words, the existence of  $z_2(\gamma)$  implies  $z_1(\gamma) < z_1(\gamma=3) \approx 1.36843$ , which is a nonrealistically low value for average degree in real networks. This explains why we have chosen the modified power-law distribution (15).

The generating function (16) encapsulates all the moments of the degree distribution [31,36,37]. Hence, the expressions for  $z_1(\gamma, a)$  and  $z_2(\gamma, a)$ , Eqs. (17) and (18), are only exact in the limit  $N \rightarrow \infty$ . Further,  $\bar{l}_{SF}(N, \gamma, a)$  and  $\bar{u}_{SF}(N, \delta, \gamma, a)$ , both of which depend on  $z_2(\gamma, a)$ , are defined only where  $z_2(\gamma, a)$  is finite, i.e., for  $\gamma > 3$ . Therefore, it is essential to understand the behavior of  $z_1(\gamma, a)$  and  $z_2(\gamma, a)$  in power-law networks. Figure 7 shows  $z_1$  (full curves) and  $z_2$  (dashed curves) within the range  $\gamma > 3 \wedge Z > 1$  [where  $\bar{l}_{SF}(N, \gamma, a)$  is defined] for, from left to right,  $a=0, 1, 2$ , and 3.

Having deduced closed-form expressions for the average utility in Poisson and power-law networks, we can now compare both networks under the condition that  $z_1$  is the same. Figure 8 is a plot of average utility versus  $\delta$  when  $z_1 = \{2, 4, 10\}$  and  $N=10^5$  for Poisson and power-law networks. The average utility of Poisson networks is completely specified by  $N, \delta$ , and  $z_1$ , but power-law networks defined by (15) have one extra degree of freedom in  $z_1(\gamma, a)$ . In this

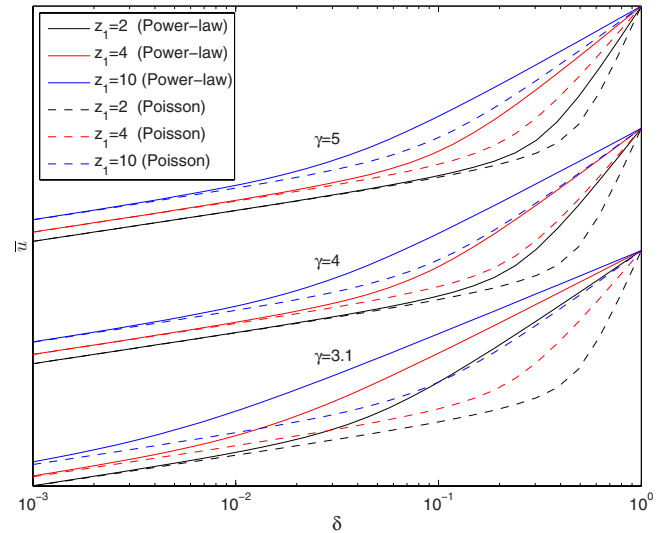


FIG. 8. (Color online) Analytical results for average utility in networks with power-law (full curves) and Poisson (dashed curves) degree distributions as a function of  $\delta, z_1$ , and  $\gamma$  for  $N=10^5$ . Curves have been shifted vertically for different values of  $\gamma$  for clarity. Values of  $z_1$  increase from bottom to top.

case, we compute  $z_1$  numerically by solving Eq. (17) for  $z_1(\gamma, a) = \{2, 4, 10\}$  when  $\gamma = \{3.1, 4, 5\}$ . For all cases studied, power-law networks are more efficient than Poisson networks.

#### IV. DISCUSSION

The growth mechanism we have proposed is a natural extension of the Barabási-Albert preferential attachment by degree to preferential attachment by node utility. Our analysis shows that, for small values of  $\delta$ , the utility decay parameter, the network retains a scale-free structure that is nonetheless destroyed when  $\delta$  increases. We have identified a regime in  $\delta$  where the network is characterized by a lower average path length and assortativity coefficient and a higher central point dominance than the scale-free network. In this regime, the distribution of utility is a steplike function and the network has a more starlike structure.

The derivation of analytical expressions for average utility in Poisson and power-law networks reveals that the latter have higher  $\bar{u}$  for the range of parameters that is of significance in real-world networks ( $z_1 \geq 2$ ). This suggests a mechanism that may explain the ubiquitous presence of power-law networks, in particular in situations where collaboration, interaction, and information sharing among the nodes are of paramount relevance.

Social networks are highly volatile. Friendships can be stable for a long time but occasional encounters may lead to creation of links that are never used again in the future. A dynamical model of network formation, where not only can links be heterogeneous, but their weights can change continuously over time, would be a more appropriate way to describe social interactions. Our assumption that  $w_{ij}=1$  for all links is obviously a first-order approximation and preferential growth, with no link rearrangements, is a crude de-



scription of social network formation. Nonetheless, even this simple mechanism can highlight surprising features of the models (as in our case a smaller network diameter for intermediate values of  $\delta$ ) and as such it is worthwhile to investigate it in more general contexts than the original BA model.

We have also assumed that the connection costs in our model are zero. This assumption is justified by the fact that if costs are node independent they do not play any role in the growing model. Similarly, costs do not play a significant role if we restrict the comparison of average utility in Sec. III to networks with the same size and the same average degree. Nonetheless in a more realistic model, where links can be rearranged over time, costs would also play an important role in determining the shape of the network. Further analysis taking into account both these effects is currently under development.

## ACKNOWLEDGMENTS

We wish to thank Ginestra Bianconi, Anirban Chakraborti, Francesco Feri, Sanjeev Goyal, Matthew Jackson, Jukka-Pekka Onnela, and Marco Patriarca for stimulating discussions. We are grateful to the organizers and participants of the EXYSTENCE Thematic Institute “From Many-Particle Physics to Multi-Agent Systems” held at the Max Planck Institute for the Physics of Complex Systems (MPIPKS) in Dresden, where the collaboration leading to this work was started. R..C would like to acknowledge support from the EPSRC Spatially Embedded Complex Systems Engineering Consortium Grant No. EP/C513703/1 and G.I. the support of the COST-P10 Action “Physics of Risk.”

- 
- [1] M. O. Jackson and A. Wolinsky, *J. Econ. Theory* **71**, 44 (1996).
- [2] V. Bala and S. Goyal, *Econometrica* **68**, 1181 (2000).
- [3] M. Marsili, F. Vega-Redondo, and F. Slanina, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1439 (2004).
- [4] D. O. Cajueiro, *Phys. Rev. E* **72**, 047104 (2005).
- [5] G. C. M. A. Ehrhardt, M. Marsili, and F. Vega-Redondo, *Phys. Rev. E* **74**, 036106 (2006).
- [6] A. D. Martino and M. Marsili, *J. Phys. A* **39**, R465 (2006).
- [7] G. Iori, S. Jafarey, and F. G. Padilla, *J. Econ. Behav. Organ.* **61**, 525 (2006).
- [8] M. O. Jackson, in *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, edited by W. N. Richard Blundell and T. Persson (Cambridge University Press, Cambridge, U.K., 2006, Vol. 1).
- [9] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
- [10] M. Newman, *SIAM Rev.* **45**, 167 (2003).
- [11] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).
- [12] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
- [13] F. Feri, *J. Econ. Theory* **135**, 442 (2007).
- [14] J. Kleinberg and E. Tardos, *Algorithm Design* (Addison-Wesley, Reading, MA, 2006).
- [15] H. S. Wilf, *Generatingfunctionology* (Academic Press, San Diego, 1994).
- [16] G. U. Yule, *Philos. Trans. R. Soc. London, Ser. B* **213**, 21 (1925).
- [17] H. A. Simon, *Biometrika* **42**, 425 (1955).
- [18] S. N. Dorogovtsev and J. F. F. Mendes, *Adv. Phys.* **51**, 1079 (2002).
- [19] R. Durrett, *Random Graph Dynamics* (Cambridge University Press, Cambridge, U.K., 2007).
- [20] C. Bedogne and G. J. Rodgers, *Phys. Rev. E* **74**, 046115 (2006).
- [21] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Munõz, *Phys. Rev. Lett.* **89**, 258702 (2002).
- [22] G. De Masi, G. Iori, and G. Caldarelli, *Phys. Rev. E* **74**, 066112 (2006).
- [23] V. D. P. Servedio, G. Caldarelli, and P. Buttà, *Phys. Rev. E* **70**, 056126 (2004).
- [24] D. Thain, T. Tannenbaum, and M. Livny, *Concurrency Comput.: Pract. Exper.* **17**, 323 (2005).
- [25] K.-I. Goh, B. Kahng, and D. Kim, *Phys. Rev. Lett.* **87**, 278701 (2001).
- [26] M. Barthélemy, *Eur. Phys. J. B* **38**, 163 (2004).
- [27] T. Kamada and S. Kawai, *Inf. Process. Lett.* **31**, 7 (1989).
- [28] M. E. J. Newman, *Phys. Rev. E* **67**, 026126 (2003).
- [29] L. C. Freeman, *Sociometry* **40**, 35 (1977).
- [30] M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).
- [31] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Phys. Rev. E* **64**, 026118 (2001).
- [32] M. Newman, D. J. Watts, and S. H. Strogatz, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2566 (2002).
- [33] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Phys. Rev. Lett.* **85**, 5468 (2000).
- [34] M. Aldana, *Physica D* **185**, 45 (2003).
- [35] M. E. J. Newman, *Phys. Rev. Lett.* **95**, 108701 (2005).
- [36] G. K. Zipf, *Human Behavior and the Principle of Least Effort* (Addison-Wesley, Reading, MA, 1949).
- [37] U. Garibaldi, D. Costantini, S. Donadio, and P. Viarengo, *Comput. Econ.* **27**, 115 (2006).