Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

Loops of any size and Hamilton cycles in random scale-free networks

Ginestra Bianconi[1,2] **and Matteo Marsili**[1]

[1] The Abdus Salam International Center for Theoretical Physics, Strada Costiera 11, 34014 Trieste, Italy
[2] INFM, UdR Trieste, via Beirut 2-4, 34014, Trieste, Italy
E-mail: gbiancon@ictp.trieste.it and marsili@ictp.trieste.it

Received 23 February 2005
Accepted 19 May 2005
Published 7 June 2005

Online at stacks.iop.org/JSTAT/2005/P06005
doi:10.1088/1742-5468/2005/06/P06005

Abstract. Loops are subgraphs responsible for the multiplicity of paths going from one to another generic node in a given network. In this paper we present an analytic approach for the evaluation of the average number of loops in random scale-free networks valid at fixed number of nodes N and for any length L of the loops. We bring evidence that the most frequent loop size in a scale-free network of N nodes is of the order of N as in random regular graphs, while small loops are more frequent when the second moment of the degree distribution diverges. In particular, we find that finite loops of sizes larger than a critical one almost surely pass from any node, thus casting some doubts on the validity of the random tree approximation for the solution of lattice models on these graphs. Moreover, we show that Hamiltonian cycles are rare in random scale-free networks and may fail to appear if the power-law exponent of the degree distribution is close to two even for minimal connectivity $k_{\rm min} \geq 3$.

Keywords: random graphs, networks

ArXiv ePrint: cond-mat/0502552

 1742-5468/05/P06005+14$30.00

Contents

The scale-free network structure has been found in a number of social, technological and biological networks as the skeleton of their interaction [1]–[3]. The main property of scale-free networks is to have a power-law degree distribution $P(k) \sim k^{-\gamma}$ and second diverging moment, i.e. $\gamma \in (2, 3]$. To distinguish between different scale-free networks, recently, much attention has been devoted to network motifs [4]–[6], i.e. subgraphs that recur with higher frequency than in maximally random graphs with the same degree distribution. Among these, the most simple types of subgraphs are loops [7]–[9], i.e. closed paths of various length that visit each node only once. Loops (or cycles) are interesting because they account for the multiplicity of paths between any two nodes. Therefore, they encode the redundant information in the network structure. Another discriminant aspect of real scale-free networks is the presence of degree correlations between linked nodes. Characteristic motifs in a graph and degree correlations are in many real graphs not independent phenomena but they depend on each other as has been shown for small (up to maximal connectivity) size subgraphs in [11, 10, 12, 13]. Last but not least, it has been observed that the distribution of loop sizes is intimately connected with the thermal properties of lattice models defined on that graph [18, 17]. On the other hand, the analytic approach to these models relies on the assumption that locally a random graph can be considered to have a tree-like structure [19]–[21], i.e. that loops of finite size are rare.

In this paper we present an analytic derivation of the average number of loops of any size in a random scale-free network. Our motivation is that the results on random networks provide a reference picture which often captures key intuition which extends to correlated networks (see e.g. [14, 15]). In addition, it provides valuable information for the statistical mechanics of the lattice model on random graphs [18, 17].

J. Stat. Mech. (2005) P06005

Let us first recall the classic results for regular random graphs, i.e. random graphs with fixed connectivity of the nodes $k_i = c$ for each node i. A regular random graph contains a finite number of small loops of size $L \ll \log(N)$, with average expected number

$$\mathcal{N}_L = \frac{1}{2L}(c-1)^L \tag{1}$$

and Poisson fluctuations around the mean. In contrast, for large loop sizes $L \sim \mathrm{O}(N)$ the number of loops goes as

$$\mathcal{N}_L = \exp(N\sigma(\ell)) \tag{2}$$

where $\ell = L/N$ and $\sigma(\ell)$ is a function having the maximum at $\ell_{\max} = c/(c+1)$ whose expression can be found in the literature [16, 17]. Regarding Hamilton cycles, i.e. loops that span the entire network $L = N$, their expected number for a large regular random graph is diverging with the system size as long as $c \geq 3$. For $c = 2$ the average number of Hamilton cycles goes to zero as the system size diverges [16]. Coming to the scale-free network literature, [9] analyses the number of loops of any size on a pseudo-fractal scale-free graph and reports the scaling behaviour

$$\log \mathcal{N}_L = L f(L/L^*) \tag{3}$$

with $L^* = N^{1/(\gamma-1)}$. No result have been presented on Hamilton cycles, to our knowledge, so far.

In the following we characterize the statistics of loops in random scale-free networks. We find a larger number of small loops with respect to regular random graphs. In particular, we compute the expected number of loops of a given size passing through a node and find that when $\gamma \in (2,3)$ this number diverges with the network size, beyond a finite loop size. This raises some doubts on the solution of lattice models on these graphs based on the local tree approximation [19]–[21]. We also find that loops have a characteristic size $L^* \sim N$. In other words, our results are consistent with the scaling (3) with $L^* \sim N$ and with equation (2) for regular graphs. This suggests that the result of [9] crucially depends on the peculiar correlations of the ensemble they consider. Special attention will be given to Hamilton cycles that in networks with a small γ exponent can fail to exist unless the lower cut-off of the distribution is large enough.

There are different ensembles of random networks one can consider. The classical one follows the prescription of Molloy and Reed [24]: first, to each node i of the network is assigned a connectivity k_i drawn from the chosen probability distribution; second, edges are randomly matched. This ensemble indeed generates networks of given degree distribution but it may yield networks with multiply occupied links. More precisely, for the distribution of the links between two nodes of connectivity k_i, k_j is a Poisson variable with mean $k_i k_j/(cN)$, where henceforth $c = \langle k \rangle$ will denote the average connectivity. Hence the probability of no multiply occupied links is

$$\Pi_{i>j}\left(1 + \frac{k_i k_j}{cN}\right)\exp\left(-\frac{k_i k_j}{cN}\right) \simeq \exp\left(-\frac{1}{4}\left(\frac{\gamma-2}{3-\gamma} m^{\gamma-2} K^{(3-\gamma)}\right)^2\right) \tag{4}$$

where the right-hand side refers to a scale-free random graph with degree distribution $P(k) = ak^{-\gamma}$ with $k \in [m, K]$. Taking a structural cut-off $K \sim N^{1/2}$ [25, 26] for $\gamma \in (2, 3]$,

J. Stat. Mech. (2005) P06005

we conclude that double links appear with probability one as $N \to \infty$ in the Molloy–Reed ensemble. When counting loops of a network this effect becomes relevant and undesirable. Thus we also consider another ensemble where double links cannot appear: the static fitness network [22, 23]. In the fitness ensemble nodes are assigned a random variable (fitness) q drawn from a $\rho(q)$ distribution function and every couple of nodes is linked with a probability depending on the fitness of the considered nodes $p(q, q')$. When $\rho(q)$ is power-law distributed and $p(q, q') = qq'/\langle q\rangle N$ the resulting graph is a random scale-free graph characterized by the same exponent of the fitness distribution. In these graphs the connectivity of every node is a Poisson variable with expected value $\langle k(q)\rangle = q$. This ensemble does not allow for networks with double links but instead may give rise to networks with isolated nodes ($k_i = 0$) or to nodes connected with a single link ($k_i = 1$) to the others. The presence of such nodes rules out the possibility to find Hamilton cycles, hence we shall take this effect into account when discussing Hamilton cycles.

Consequently, the Molloy–Reed ensemble and the fitness ensemble are not equivalent and have intrinsic properties that could sensitively perturb the counting of the number of loops. In order to understand the dependence on the details of how graphs are generated in the following we are going to study the expected number of loops in the two ensembles.

1. Loops in the fitness ensemble

The prescription of [22] to generate a class of random scale-free networks with exponent γ is the following: (i) assign to each node i of the graph a hidden continuous variable q_i distributed according to a scale-free distribution $\rho(q) = \rho_0 q^{-\gamma}$ for $q \in [m, Q]$ with $\rho_0 = (\gamma - 1)/(m^{1-\gamma} - Q^{1-\gamma})$ the normalizing constant. Then (ii) each pair of nodes with hidden variables q, q' is linked with probability $qq'/(cN)$, where $c = \langle q\rangle$ is the expected value of q. The cut-off $Q \sim \min(N^{1/2}, N^{1/(\gamma-1)})$ is needed to keep the linking probability smaller than one, i.e. $Q^2/(cN) < 1$. By construction the expected value of the connectivity of a node with hidden variable q is $\langle k|q\rangle = q$ and there are no multiple connections between nodes. Notice that the average connectivity of the graph is given by

$$\langle k\rangle = \langle q\rangle = c \to \frac{\gamma - 1}{\gamma - 2} m \tag{5}$$

in the limit $N \to \infty$.

A loop of size L is an ordered set of distinct nodes $\{1_i, \ldots, i_L\}$. For each choice of the nodes, the probability that they are connected in a loop is

$$\frac{q_{i_1} q_{i_2}}{Nc} \frac{q_{i_2} q_{i_3}}{Nc} \cdots \frac{q_{i_L} q_{i_1}}{Nc} = \prod_{\ell} \frac{q_{i_\ell}^2}{Nc}.$$

The total number of possible loops joining these L nodes in any possible way is $L!/(2L)$ where the factor $2L$ comes from the fact that the initial node of the loop can be chosen in L ways and that there are two orientations. In order to count loops, let us lump together nodes with hidden variable $q_i \in [q, q + \Delta q)$, where Δq is a small interval of q. In each interval of q there are $N_q \simeq NP(q)\Delta q$ nodes of the network. For each choice of the L nodes, let n_q be the number of nodes with $q_{i_\ell} \in [q, q + \Delta q)$. Then the average number of loops of size L in the graph is given by the number of ways we can choose $\{n_q\}$

J. Stat. Mech. (2005) P06005

nodes multiplied by the probability that these nodes are connected in all distinguishable orderings. Consequently we have

$$\mathcal{N}_L = \frac{L!}{2L}\sum_{\{n_q\}}\prod_q \frac{N_q!}{n_q!(N_q-n_q)!}\left(\frac{q^2}{Nc}\right)^{n_q} \tag{6}$$

where the sum is extended over all $\{n_q\}$ such that $\sum_q n_q = L$. Introducing this constraint with a delta function and using its integral representation, we find

$$\mathcal{N}_L = \frac{L!}{2L}\int_{-\infty}^{\infty}\frac{\mathrm{d}x}{2\pi}\exp\left(\mathrm{i}Lx + N\left\langle \log\left[1+q^2\mathrm{e}^{-\mathrm{i}x}/(Nc)\right]\right\rangle\right). \tag{7}$$

Notice that in equation (7) one can safely take the limit $\Delta q \to 0$ and that the average over the $P(q)$ distribution is taken assuming that we focus on the limit $N\to\infty$. In what follows, we will evaluate equation (7) in different ranges of $L \le N$ in the limit $N\to\infty$.

1.1. Small loops

For L finite but large, the integral in equation (7) is dominated by values $x \simeq -\mathrm{i}z^*$ where

$$\mathrm{e}^{-z^*} \simeq \frac{\langle q^2\rangle}{Lc}\left(1 - \frac{\langle q^4\rangle}{\langle q^2\rangle^2}\frac{L}{N} + \cdots\right) \tag{8}$$

where we have neglected all terms beyond the first leading correction when $N\to\infty$. The argument of the exponential in equation (7) can be expanded around $x \sim -\mathrm{i}z^*$ yielding $N\langle\log[1+q^2\mathrm{e}^{-\mathrm{i}x}/cN]\rangle + Li x \simeq L[1 - z^* - \frac{1}{2}(x-\mathrm{i}z^*)^2 + \mathrm{O}(x-\mathrm{i}z^*)^3]$. Hence the integral can be estimated by the saddle point for L large. Using the asymptotic expression $L! \simeq \sqrt{2\pi L}L^L\mathrm{e}^{-L}$, we find to leading order

$$\mathcal{N}_L \simeq \frac{1}{2L}\left(\frac{\langle q^2\rangle}{c}\right)^L. \tag{9}$$

This approximation is valid as long as the leading correction in equation (8) is small. Using that $\langle q^n\rangle = \rho_0(Q^{n-\gamma+1} - m^{n-\gamma+1})/(n-\gamma+1)$ for $\gamma \neq n$ and that $Q \sim N^{1/2}$ we find that the expression above for $\mathcal{N}_L$ holds when

$$L \ll N\frac{\langle q^2\rangle^2}{\langle q^4\rangle} \sim \begin{cases} N & \gamma > 5 \\ N^{2\frac{\gamma-3}{\gamma-1}} & 3<\gamma<5 \\ N^{(3-\gamma)/2} & 2<\gamma<3 \end{cases} \tag{10}$$

with logarithmic corrections for $\gamma = 3$ and 5. Note that strictly speaking the expansion (8) is converging only for $N\langle q^2\rangle/L \gg Q^2$, i.e. $L \ll N^{(3-\gamma)/2}$ for $2<\gamma<3$ and $L \ll N^{(\gamma-3)/(\gamma-1)}$ for $\gamma > 3$. Nevertheless, equation (9) remains valid in the limits (10) as an asymptotic expansion. For $\gamma > 3$ we obtain a result very similar to equation (1) for regular graphs. In contrast, for $2<\gamma<3$ we have $\langle q^2\rangle \simeq aN^{(3-\gamma)/2}$, with a a constant, hence the number of finite loops

$$\mathcal{N}_L \simeq \frac{1}{2L}\left(\frac{a}{c}\right)^L N^{((3-\gamma)/2)L} \tag{11}$$

diverges as $N\to\infty$.

J. Stat. Mech. (2005) P06005

1.2. Intermediate loop sizes and the most frequent loops

It is convenient, at this point, to write equation (7) as

$$\mathcal{N}_L \simeq \int_{-\infty}^{\infty} \frac{\mathrm{d}x}{2\sqrt{\pi L}} \exp(Nf(Nc\,\mathrm{e}^{\mathrm{i}x}, L/N)) \tag{12}$$

where we have used Stirling's approximation and

$$f(y,\ell) = \left\langle \log\left[1+q^2/y\right]\right\rangle + \ell\log(\ell y/c) - \ell. \tag{13}$$

The integral can be computed by the saddle point method, deforming the contour of integration so as to pass from the point where f is stationary. The condition $\partial_y f(y,\ell)=0$ yields

$$\left\langle \frac{q^2}{q^2+y} \right\rangle = \ell. \tag{14}$$

Let $y^*(\ell)$ be the value of y which solves this equation. We can expand $f(Nc\,\mathrm{e}^{\mathrm{i}x},\ell)$ around the corresponding (complex) value x^* of x

$$\begin{aligned} f(Nc\mathrm{e}^{\mathrm{i}x},\ell) &= f(y^*,\ell) - \frac{y^{*2}}{2}\frac{\partial^2 f}{\partial y^2}(x-x^*)^2 + \cdots \\ &= f(y^*,\ell) - \frac{y^*}{2}\left\langle \frac{q^2}{[q^2+y^*]^2}\right\rangle (x-x^*)^2 + \cdots. \end{aligned} \tag{15}$$

As long as

$$Ny^*\left\langle \frac{q^2}{[q^2+y^*]^2}\right\rangle \gg 1$$

we can neglect higher order terms. This yields the leading behaviour

$$\mathcal{N}_L \simeq \frac{1}{2}\left[LNy^*\left\langle \frac{q^2}{[q^2+y^*]^2}\right\rangle\right]^{-1/2} \mathrm{e}^{Nf(y^*,\ell)}. \tag{16}$$

The number of loops $\mathcal{N}_L$ takes its maximum for loops of length $L=N\ell_{\max}$ where

$$\left\langle \frac{q^2}{c+q^2\ell_{\max}} \right\rangle = 1. \tag{17}$$

The solution $\ell_{\max}$ is plotted in figure 1 against γ for scale-free graphs and different values of m, for $N=10^6$. Notice that as $\gamma\to 2^+$ the size of most probable loops vanishes as $\ell_{\max}\sim\gamma-2$. Around the maximum, $\mathcal{N}_L$ takes a form similar to that for regular graphs (see equation (2)), which is consistent with the scaling form equation (3) with $L^*\sim N$.

For scale-free random graphs with $2<\gamma<3$ there is an intermediate range of loop sizes $L\sim N^{(3-\gamma)/2}$ which is related to the solutions with $y^*=\mu N$ with $\mu>1$. More precisely, we find that for loops of size $L=\chi(\mu)N^{(3-\gamma)/2}$ we have

$$\mathcal{N}_L \simeq \frac{G(\mu)}{L}\exp(L[\log(\mu L/c) - 1 + H(\mu)])$$

J. Stat. Mech. (2005) P06005

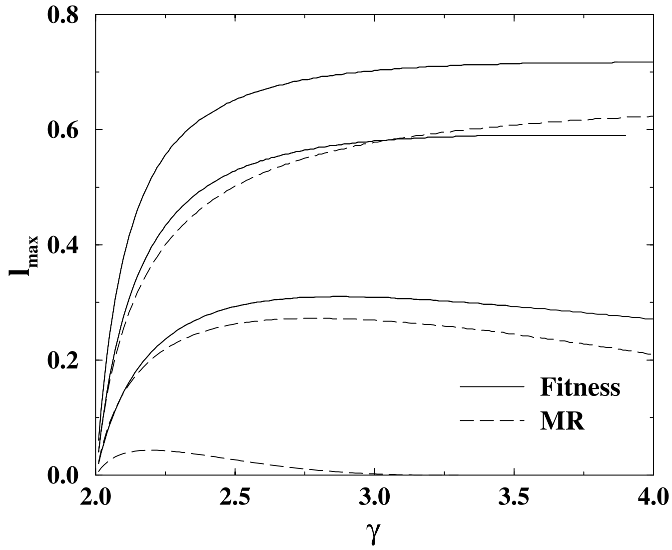

Figure 1. The behaviour of $\ell_{\max}$ as a function of γ for $m = 1, 2, 3$ and $N = 10^6$ for the fitness ensemble (solution of equation (17) in solid lines) and for the Molloy–Reed ensemble (solution of equation (32) in dashed lines).

with $G(\mu)$ a function of μ and

$$\begin{aligned}\chi(\mu) &= (\gamma-1)m^{\gamma-1}\sum_{n=0}^{\infty}(-1)^n\frac{\mu^{-n-1}}{3-\gamma+2n}\\ H(\mu) &= \frac{1}{\chi(\mu)}(\gamma-1)m^{\gamma-1}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}\mu^{-n}}{n(1-\gamma+2n)}.\end{aligned} \tag{18}$$

Notice that $\mathcal{N}_L$ does not satisfy a simple scaling form such as equation (3) in this intermediate region.

1.3. Hamilton cycles

From equation (6) we can easily calculate the number of Hamilton cycles. Indeed for $L = N$ we have the asymptotic behaviour

$$\mathcal{N}_N = \sqrt{\frac{\pi}{2N}}\exp(N[2\langle\log q\rangle - 1 - \log c]). \tag{19}$$

This is the expected number of Hamiltonian cycles over all the networks of the fitness ensemble including networks with nodes of low degree $k_i = 0, 1$, which by definition cannot have a Hamilton cycle. It seems a sensible thing to compute the number of Hamilton cycles in networks with a minimal degree connectivity greater than three, i.e. $k_i >= 3$. In fact it is well known that for a regular random graph of connectivity $c = 2$ the expected number of Hamilton cycles goes to zero in the $N \to \infty$ limit whereas regular graphs are Hamiltonian when $c \geq 3$. Taking this as a reference result, we normalize $\mathcal{N}_N$ by the probability π that all the nodes have at least three connections. Since the connectivity of

J. Stat. Mech. (2005) P06005

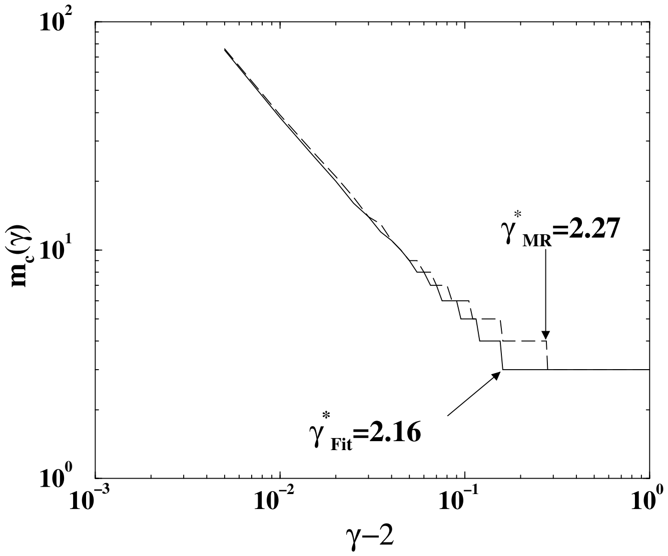

Figure 2. Dependence of $m_c(\gamma)$ on γ in the fitness (solid line) and Molloy–Reed ensemble (dashed line) in the limit $N \to \infty$. Observe that for $\gamma < \gamma^*$ we have $m_c(\gamma) > 3$ in both ensembles.

each node in the fitness network is a Poisson variable with expected value q the probability that all the nodes have connectivity $k \geq 3$ is simply given by $\pi = \mathrm{e}^{N\lambda(m,\gamma)}$, where

$$\lambda(m, \gamma) = \langle \log(1 - (1 + q + q^2/2)\mathrm{e}^{-q}) \rangle. \tag{20}$$

In the limit $N \to \infty$ we find

$$\frac{1}{N} \log \frac{\mathcal{N}_N}{\pi} \to \log\left(\frac{\gamma-2}{\gamma-1} m\right) + \frac{3-\gamma}{\gamma-1} - \lambda(m, \gamma). \tag{21}$$

This implies that random scale-free graphs have Hamilton cycles only for $m > m_c(\gamma)$, where $m_c(\gamma)$ is the value of m for which equation (21) vanishes. Conversely, for $m < m_c(\gamma)$ a random scale-free graph has almost surely no Hamilton cycle.

In figure 2 we report the critical value $m_c(\gamma)$ as a function of γ. Notice that $m_c \sim 1/(\gamma - 2) \to \infty$ as $\gamma \to 2^-$. Consequently, if we consider only the networks of the ensemble with $k_{\min} \geq 3$, we find that random graphs with $\gamma < \gamma^*_{\mathrm{Fit}} = 2.16\ldots$ (where $m_c(\gamma^*_{\mathrm{Fit}}) = 3$) do not have Hamiltonian cycles. Considering that regular random graphs with $k = c \geq 3$ are Hamiltonian, this may seem a surprising result, at first sight. The basic intuition to explain this apparent paradox is that most paths pass through well connected nodes. Hence even if $k_i \geq 3$ it is very unlikely to have a path spanning the entire network which is not passing through the most connected nodes more than once.

1.4. Loops passing through a node

In order to count of the number of loops of size L passing through a given node, with fitness value q_i, we can repeat the previous outlined above, without taking the average over q_i. For $\gamma < 3$ and short loop sizes $L \ll N^{(3-\gamma)/2}$ this gives the expected number

$$\mathcal{N}_L(q_i) \simeq \frac{q_i^2}{cN} \frac{1}{2L} \left(\frac{a}{c}\right)^{(L-1)} N^{((3-\gamma)/2)(L-1)}. \tag{22}$$

J. Stat. Mech. (2005) P06005

Focusing on nodes with $q_i \simeq N^{\alpha}$, we find that the number of loops of size

$$L \geq L_0 \equiv 1 + \frac{2}{3-\gamma}(1-2\alpha) \tag{23}$$

diverges with the network's size N. For example, nodes with a finite q_i have an infinite number of $L = 5$ loops passing through them in networks with $\gamma < 2.5$ but at most a finite number of loops of size $L = 3$ if $\gamma > 2$. The most connected nodes ($\alpha = 1/2$) instead have an infinite number of loops of any size $L \geq 3$ passing through them. Notice that $L_0 \to \infty$ as $\gamma \to 3$ in order to match the behaviour $L_0 \sim \log N$ of regular graphs. Conversely, in a finite graph of N nodes, only the large fitness nodes with

$$q_i \gg N^{(1/2)-((3-\gamma)/2)(L-1)} \tag{24}$$

belong to a significant number of loops of size L.

2. Loops in the Molloy–Reed ensemble

The counting of the number of loops in the Molloy–Reed [24] ensemble follows a procedure very similar to the one considered for the fitness ensemble, nevertheless giving different results. To construct a Molloy–Reed network one proceeds as follows. (i) A degree is assigned to each node of the network following the desired degree distribution with cutoff $K \sim \min(N^{1/2}, N^{1/(\gamma-1)})$. Degree distributions which do not satisfy the parity of $cN = \sum_i k_i$ are disregarded. (ii) The edges coming out of the nodes are randomly matched until all edges are connected. When this procedure ends with nodes having links to themselves (tadpoles), the whole network is rejected and the procedure is started anew.

To calculate $\mathcal{N}_L$ in this ensemble first one has to count in how many ways it is possible to have a loop of size L in the network and weight the results with the fraction of possible networks in the ensemble which contains the loop. Let us first state that the total number of graphs in the Molloy–Reed ensemble is given by $(cN-1)!!$. Indeed when constructing the network by linking cN unconnected edges one start by taking one edge at random and connecting it to one of the $(cN-1)$ possible connections. Then one proceeds taking another edge and linking it to one of the remaining $(cN-3)$ possible connections, thus giving rise of one of the $(cN-1)!!$ possible networks. By similar arguments one shows that the total number of networks containing a given loop of size L is $(cN-2L-1)!!$. On the other side the total number of loops of size L in the Molloy–Reed ensemble is given by the number of ways one can choose an ordered set of L nodes $\{1_i, \ldots, i_L\}$ of connectivity $\{k_1, k_2, \ldots, k_L\}$ and connect them on a loop. As for the fitness network, the total number of possible loops joining these L nodes in any possible way is $L!/(2L)$. The number of ways one can choose the edges coming out of the nodes to form the loop is given by

$$\Pi_{i=1}^{L} k_i(k_i-1).$$

Consequently, the average number of loops in the Molloy–Reed ensemble will be given by

$$\mathcal{N}_L = \frac{L!}{2L} \sum_{\{n_k\}} \prod_{k=m}^{K} \frac{N_k!}{n_k!(N_k-n_k)!} \left(k(k-1)\right)^{n_k} W_{N,L} \tag{25}$$

J. Stat. Mech. (2005) P06005

where $N_k = NP(k)$ (n_k) is the number of nodes with connectivity k present in the network (loop), K is the structural cut-off and the sum over $\{n_k\}$ is restricted to $\{n_k\}$ such that $\sum_k n_k = L$. Moreover we use the definition $W_{N,L} = (cN-2L-1)!!/(cN-1)!!$. If we use the Stirling approximation for $W_{N,L}$ we get the expression

$$W_{N,L} \sim (cN)^{-L} \mathrm{e}^{Ng(\ell)} \tag{26}$$

with $\ell = L/N$ and

$$g(\ell) = \frac{1}{2}(c-2\ell)\log\left(\frac{c-2\ell}{c}\right) + \ell. \tag{27}$$

Thus we get

$$\mathcal{N}_L = \frac{L!}{2L}\sum_{\{n_k\}}\prod_{k=m}^{K}\frac{N_k!}{n_k!(N_k-n_k)!}\left(\frac{k(k-1)}{cN}\right)^{n_k}\mathrm{e}^{Ng(\ell)} \tag{28}$$

which except for the substitution $q^2 \to k(k-1)$ and the factor $\exp(Ng(\ell))$ is equivalent to the expression (7) of the average number of loops of size L in the fitness ensemble. Following the same steps as in the fitness ensemble, we get

$$\mathcal{N}_L = \frac{L!}{2L}\int_{-\infty}^{\infty}\frac{\mathrm{d}x}{2\pi}\exp\left(\mathrm{i}Lx + N\left\langle\log\left[1+k(k-1)\mathrm{e}^{-\mathrm{i}x}/(Nc)\right]\right\rangle + Ng(\ell)\right) \tag{29}$$

with $g(\ell)$ given by equation (27).

2.1. Small loop size

The number of small loops in the Molloy–Reed ensemble is given by

$$\mathcal{N}_L \simeq \frac{1}{2L}\left(\frac{\langle k(k-1)\rangle}{c}\right)^L \tag{30}$$

where this approximation is valid asymptotically for loop sizes satisfying equation (10). Note that as in the fitness ensemble for $\gamma \in (2,3)$ short loops diverge as $\mathcal{N}_L \sim N^{(3-\gamma)/2}$.

2.2. Intermediate loop sizes

For intermediate loops in the Molloy–Reed ensemble a similar expression to equation (13) holds with

$$f'(y,\ell) = \langle\log[1+k(k-1)/y]\rangle + \ell\log(\ell y/c) - \ell + g(\ell). \tag{31}$$

The calculations of the average number of loops are very similar for the Molloy–Reed and fitness ensembles, with a difference for the equation of the loops of maximal size which in the Molloy–Reed ensemble satisfy

$$\left\langle\frac{k(k-1)}{c-2\ell_{\max}+k(k-1)\ell_{\max}}\right\rangle = 1. \tag{32}$$

In figure 1 we report the value of $\ell_{\max}$ in the Molloy–Reed networks as a function of γ for different values of the minimal connectivity m for $N=10^6$.

J. Stat. Mech. (2005) P06005

2.3. Hamiltonian cycles

Starting with expression (29) one can easily evaluate the expected number of Hamiltonian cycles in Molloy–Reed networks. Indeed for $L = N$ and $c > 2$ one can use the Stirling approximation to find the asymptotic behaviour ($N \to \infty$)

$$\frac{1}{N}\log(\mathcal{N}_N) = \langle \log(k(k-1)/c)\rangle + \frac{1}{2}(c-2)\log(1-2/c). \tag{33}$$

If we approximate $k(k-1)$ with k^2 which is possible close to $\gamma \to 2$ in the limit $c \to \infty$ we recover the same behaviour as in the fitness ensemble: if the minimal connectivity m is smaller than the value $m_c(\gamma)$ for which equation (33) vanishes, then a scale-free network is typically not Hamiltonian. In figure 2 we report $m_c(\gamma)$ for $2 < \gamma < 3$ and we confirm the behaviour $m_c \sim 1/(\gamma-2)$ for $\gamma \to 2$. For example, we find that Molloy–Reed random scale-free graphs with minimal connectivity $m = 3$ are typically not Hamiltonian if $\gamma < \gamma^*_{\mathrm{MR}} = 2.27\ldots$. As for the fitness ensemble, the intuition is that it is not possible to extract from a random scale-free graph a subgraph which is a regular random graph with fixed connectivity $c \geq 3$ if $m < m_c(\gamma)$.

2.4. Loops passing through a node

To count the loops of size L passing through a given node of connectivity k we must fix it and choose another $L-1$ nodes to form the loop. For short loop sizes and exponent $\gamma < 3$ this gives the expected number

$$\mathcal{N}_L(k) \simeq \frac{k(k-1)}{cN}\frac{L-1}{L}\mathcal{N}_{L-1} \tag{34}$$

with $\mathcal{N}_L \sim N^{((3-\gamma)/2)L}$. The same results as derived for the fitness ensemble hold: There is a critical finite loop size $L_0(k)$ such that there are infinitely many loops of size $L > L_0$ passing through a given node of connectivity k. In contrast, in a finite but large network of N nodes, loops of size L becomes significant for nodes of connectivity

$$k \gg N^{(1/2)-((3-\gamma)/4)(L-1)}. \tag{35}$$

3. Numerical results

We compare the analytic results derived so far with the direct count of the number of loops in a sample of computer generated random graphs in both the fitness and the Molloy–Reed ensemble. This is important because $\mathcal{N}_L$ is a fluctuating quantity which takes exponentially large values. In other words, the analytic calculation of the expected number of loops may be dominated by (exponentially in N) rare realization of graphs with an exponentially large number of loops. In this case the number of loops of a typical realization of a graph would differ from our estimate. We have chosen the fastest known algorithm for calculating the total number of loops exactly [27] as in [17]. This algorithm has a upper time bound of $\mathrm{O}(N\mathcal{L})$ where $\mathcal{L}$ is the total number of loops in the network. The simulations performed in this way enable one only to consider small network sizes $N < 50$ and small $m \leq 3$, as the total number of loops in such graphs increases exponentially with the system size. Note that for such small sizes the degree distribution contains nodes of

J. Stat. Mech. (2005) P06005

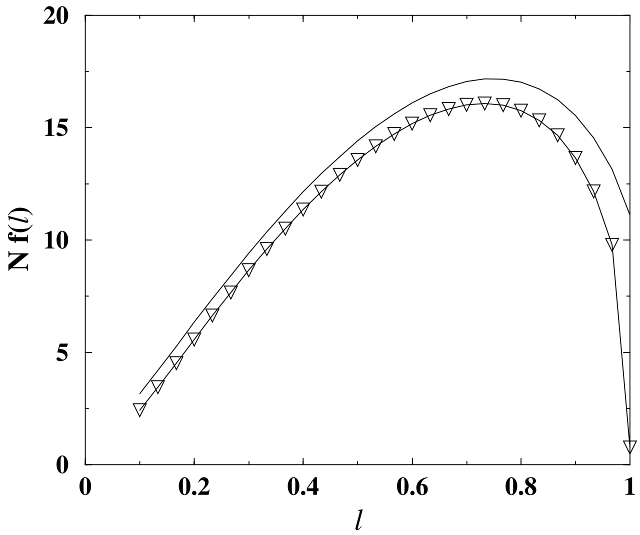

Figure 3. Number of loops for fitness networks of $N = 30$ nodes and given distribution $\rho(q)$ with $m = 3$. The average is taken over 50 realizations.

very similar degree since the upper cut-off is $K \sim 6$ for $\gamma = 2.1$. Moreover, in order to compare the direct counting with the analytical calculation, we have chosen a fixed degree (fitness) distribution $N_k = NP(k)$ to reduce fluctuations that become relevant for such small sizes. In figure 3 we report the analytic prediction of the average number of loops of a given size in a fitness network of $N = 30$ nodes. These results are compared with direct counting of the loops in computer realizations of these networks where data are averaged over 50 realizations. We found strong sample to sample fluctuations which we believe are responsible for the deviation from the analytical results.

In contrast, for the Molloy–Reed networks of the same system size the direct count of loops is very close to the analytic prediction. Figure 4 reports the direct count of loops for Molloy–Reed networks[3] with $N = 30$ and several degree distributions and it compares it with the corresponding analytic prediction.

4. Conclusions

In conclusion, we have computed analytically the expected number of loops of any size in a scale-free network. We found that scale-free graphs have a very large number of small loops compared to regular random graphs. In contrast we have shown that, also with a minimal connectivity $k_{\min} \geq 3$ the expected number of Hamilton cycles can be zero in the $N \to \infty$ limit provided that γ is sufficiently close to two. The reason for this is that paths connecting many nodes need to pass frequently on nodes with high connectivity. Put differently, it is not possible to embed a regular graph of connectivity $c \geq 3$, which would have a Hamilton cycle, in scale-free networks if γ is too small, even if all nodes have $k_i \geq c$. In the intermediate region of relatively large loops we found that the expected

[3] Note that in the considered Molloy–Reed networks double links are not allowed. This correction is supposedly quite small for the considered system sizes.

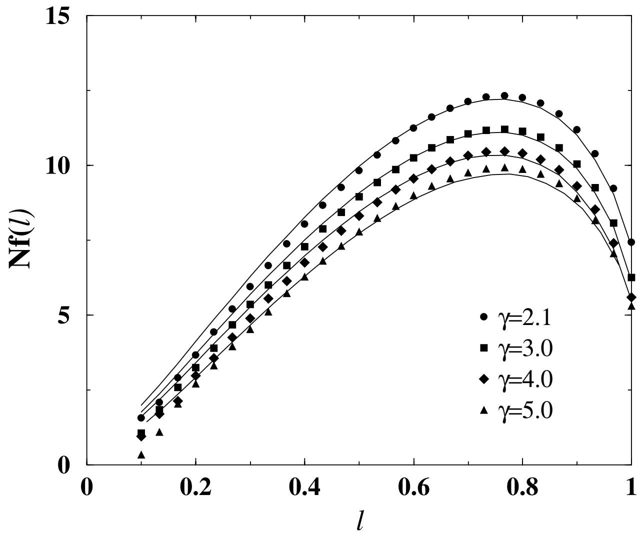

Figure 4. Number of loops for MR networks without double links of $N = 30$ nodes and different γ. The direct count averaged on ten realizations is compared to the analytic prediction for the same degree distribution (full lines).

number of loops attains its maximum for loops of size $L \sim N$. These results are derived both in the fitness and in the Molloy–Reed ensembles. While the generic picture is the same, the results in the two ensembles differ quantitatively, highlighting that the loop size distribution is somewhat sensitive to the precise prescription for drawing random graphs. Moreover, we have checked the results with direct counting of computer generated scale-free networks belonging to the two ensembles. It would be desirable to derive similar results for ensembles of correlated scale-free networks.

Acknowledgments

GB was partially supported by EVERGROW and by EU grant HPRN-CT-2002-00319, STIPCO. We wish to thank F Tria for useful discussions.

References

[1] Albert R and Barabási A-L, *Statistical mechanics of complex networks*, 2002 *Rev. Mod. Phys.* **74** 47
[2] Dorogovtsev S N and Mendes J F F, 2003 *Evolution of Networks* (Oxford: Oxford University Press)
[3] Pastor-Satorras R and Vespignani A, 2004 *Evolution and Structure of the Internet* (Cambridge: Cambridge University Press)
[4] Itzkovitz S, Milo R, Kashtan N, Ziv G and Alon U, *Subgraphs in random networks*, 2003 *Phys. Rev.* E **68** 026127
[5] Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D and Alon U, *Network motifs: simple building blocks of complex networks*, 2002 *Science* **298** 824
[6] Dobrin R, Beg Q K, Barabási A-L and Oltvai Z N, *Aggregation of topological motifs in the Escherichia coli transcriptional network*, 2004 *BMC Bioinformatics* **5** 10
[7] Bianconi G and Capocci A, *Number of loops of size h in growing scale-free networks*, 2003 *Phys. Rev. Lett.* **90** 078701
[8] Caldarelli G, Pastor-Satorras R and Vespignani A, *Structure of cycles and local ordering in complex networks*, 2002 *Eur. Phys. J.* B **38** 183

[9] Rozenfeld H D, Kirk J E, Bollt E M and ben-Avraham D, *Statistics of cycles: how loopy is your network?*, 2004 *Preprint* cond-mat/0403536
[10] Dorogovstev S, *Clustering of correlated networks*, 2004 *Phys. Rev.* E **69** 027104
[11] Bianconi G, Caldarelli G and Capocci A, *Loops structure of the internet at the autonomous system level*, 2004 *Preprint* cond-mat/0408349
[12] Vazquez A, Dobrin R, Sergi D, Eckmann J-P, Oltvai Z N and Barabási A-L, *The topological relationship between the large-scale attributes and local interactions patterns of complex networks*, 2004 Proc. Nat. *Acad. Sci.* **101** 17940
[13] Vazquez A, Oliveira J G and Barabási A-L, *The inhomogeneous evolution of subgraphs and cycles in complex networks*, 2005 *Preprint* cond-mat/0501399
[14] Pastor-Satorras R and Vespignani A, *Epidemic spreading in scale-free networks*, 2001 *Phys. Rev. Lett.* **86** 3200
[15] Boguña M, Pastor-Satorras R and Vespignani A, *Epidemic spreading in scale-free networks*, 2003 *Phys. Rev. Lett.* **90** 028701
[16] Janson S, Luczak T and Rucinski A, 2000 *Random Graphs* (New York: Wiley)
[17] Marinari E and Monasson R, *Circuits in random graphs: form local trees to global loops*, 2004 *J. Stat. Mech.* P09004
[18] Parisi G, 1988 *Statistical Field Theory* (Redwood City: Addison-Wesley)
[19] Dorogovtsev S N, Goltsev A V and Mendes J F F, *Potts model on complex networks*, 2004 *Eur. Phys. J.* B **38** 177
[20] Leone M, Vazquez A, Vespignani A and Zecchina R, *Ferromagnetic ordering in graphs with arbitrary degree distribution*, 2002 *Eur. Phys. J.* B **28** 191
[21] Ehrhardt G and Marsili M, *Potts model on random trees*, 2005 *J. Stat. Mech.* P02006
[22] Caldarelli G, Capocci A, De Los Rios P and Muñoz M A, *Scale-free networks form varying vertex intrisic fitness*, 2002 *Phys. Rev. Lett.* **89** 258702
[23] Boguña M and Pastor-Satorras R, *Class of correlated random networks with hidden variables*, 2003 *Phys. Rev.* E **68** 036112
[24] Molloy M and Reed B, *A critical point for random graphs with a given degree sequence*, 1995 *Random Struct. Algorithms* **6** 161
[25] Burda Z and Krzywicki A, *Uncorrelated random networks*, 2003 *Phys. Rev.* E **67** 046118
[26] Boguña M, Pastor-Satorras R and Vespignani A, *Cut-offs and finite size effects in scale-free networks*, 2004 *Eur. Phys. J.* C **38** 205
[27] Johnson D B, *Finding elementary circuits of a directed graph*, 1975 *SIAM J. Comput.* **4** 77

J. Stat. Mech. (2005) P06005