Stochastic and mixed flower graphs

C. Tyler Diggans [1,2,3] Erik M. Bollt,[1,4] and Daniel ben-Avraham [1,2]

[1]*Clarkson Center for Complex Systems Science (C³S²), Clarkson University, Potsdam, New York 13699, USA*
[2]*Department of Physics, Clarkson University, Potsdam, New York 13699, USA*
[3]*Air Force Research Laboratory: Information Directorate, Rome, New York 13441, USA*
[4]*Department of Electrical and Computer Engineering, Clarkson University, Potsdam, New York 13699, USA*

Stochasticity is introduced to a well studied class of recursively grown graphs: (u, v)-flower nets, which have power-law degree distributions as well as small-world properties (when $u = 1$). The stochastic variant interpolates between different (deterministic) flower graphs thus adding flexibility to the model. The random multiplicative growth process involved, however, leads to a spread ensemble of networks with finite variance for the number of links, nodes, and loops. Nevertheless, the degree exponent and loopiness exponent attain unique values in the thermodynamic limit of infinitely large graphs. We also study a class of mixed flower networks, closely related to the stochastic flowers, but which are grown recursively in a *deterministic* way. The deterministic growth of mixed flower-nets eliminates ensemble spreads, and their recursive growth allows for exact analysis of their (uniquely defined) mixed properties.

I. INTRODUCTION

Generative graph models that reliably display properties found in real-world networks created something of a renaissance in the study of complex networks around the beginning of the millennium [1–6]. Most of these initial models were built on the assumptions that the network had to be stochastic, produce a compact graph, and include some form of preferential attachment [2,7–9]. However, a subset of models did away with randomness, relying instead on deterministic recursive rules for their generation [10–15], and more recently this approach has been revisited in [16,17]. One important early model, the so-called DGM network, or alternatively the pseudofractal scale-free web (PSW), became a baseline example and has been studied with respect to many network properties over the years, including loopiness [18], diffusion [19], percolation [20], spectral properties [21], and minimum dominating sets [22]. A generalization of this important network: (u, v)-flower nets, was introduced in [14] and enabled the construction of pseudofractal networks with varying network properties. Of particular interest in [14] was a linearly scaling self-similarity leading to an associated transfinite dimension, which led to a relabeling of the special small-world cases as *transfractals*. A similar generalization was put forward around the same time in [9], which focused exclusively on variations of the PSW. Recursive constructions have the advantage of allowing for exact analysis of the graphs' structural and dynamical properties but are overly rigid, as they lack stochasticity. A weighted version of the PSW was introduced in [23], which did allow for asymmetry in edge weights. It is reasonable to ask whether some measure of stochasticity can be introduced without losing all of the beneficial analysis tools that have been developed for strictly recursive structures.

In this paper we introduce stochasticity to a recursive construction called the (u, v)-flowers [14], by mixing two

(or more) deterministic rules in a random fashion. Because of the random multiplicative nature of the growth process, the resulting distributions for the number of links and sites display a small but finite variance (relative to the average) even in the thermodynamic limit of infinitely large graphs. Nevertheless, some structural characteristics, such as the degree-exponent, attain a sharp, well-defined value, and can now be tuned at will by varying the randomness parameters. We also explore an alternative *deterministic* method for mixing the recursive rules, to construct (unique) graphs that mimic stochastic flowers, yet eliminate the ensemble spread while still allowing for exact analysis of their properties.

The rest of this paper is organized as follows. A brief review of (u, v)-flowers is provided in Sec. II. The stochastic growth process that mixes randomly between different recursive rules is explained in Sec. III. This section includes also an analysis of the distributions of the number of links (M) and nodes (N) and their moments, as well as of the degree exponent and the loopiness exponent of the graphs. (The loopiness exponent α measures how h_*—the most likely length of cycles in the graph—scales with its order, $h_* \sim N^\alpha$ [18].) Mixed flowers and their generation by deterministic mixing of recursive rules is described in Sec. IV, and their size, order, and statistics of cycles is analyzed exactly. We conclude with a summary and discussion of the results in Sec. V.

II. RECURSIVE FLOWER NETS

The (u, v)-flower net [14], whose *n*th generation is denoted by $F_n(u, v)$, is constructed by a recursive process. Flower-nets model many real-world properties (e.g., power-law degree distributions, small world when $u = 1$, hierarchical community structure), and they importantly allow for exact analysis of their structure and of dynamical processes on them. The
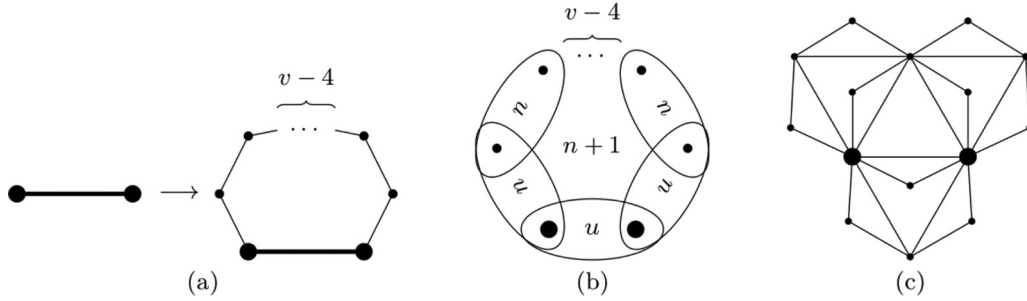
FIG. 1. The $(n+1)$th generation of $F_n(1, v)$ is constructed by using either of two methods: (a) Attach a path of v links in parallel to each link of the previous generation, or (b) paste $1 + v$ copies of the nth generation into a cycle, joining the copies hub to hub. (c) $F_3(1, 2)$, or the DGM network built to its third generation. The hubs are marked by larger full circles.

simplest case, of $(u, v) = (1, 2)$, also known as the DGM network, was first described by Dorogotsev *et al.* [12] as a *pseudofractal*. The more general subset of small world cases ($u = 1$) with generic v values were later dubbed *transfractals* in [14], due to a transfinite dimensionality defined in terms of a self-similarity that scales linearly in the distance between nodes.

The construction process begins (at generation $n = 0$) with the complete graph on two nodes, K_2. To obtain the $(n + 1)$th generation, each link in generation n is replaced by a parallel pair of paths of link lengths u and v. This process is shown in Fig. 1(a) for the case of $u = 1$. It was noted in [12] and later treatments [14,15] that the $(n + 1)$th generation could also be obtained by pasting $w = u + v$ copies of generation n at their hubs—the two original nodes from generation $n = 0$. This approach, illustrated in Fig. 1(b), is the basis to the *copy machine method* for generating adjacency matrices (see Appendix C of [19]), and is useful for computational simulations. Figure 1(c) provides an illustration of $F_3(1, 2)$, or the DGM network built to its third generation.

The number of links, M_n, and nodes, N_n, in a flower graph of generation n is [14,19,20]

$$M_n = w^n, \qquad (1)$$

$$N_n = \frac{w - 2}{w - 1} w^n + \frac{w}{w - 1}, \qquad (2)$$

where $w = u + v$ [thus, $w = 3$ for the (1,2)-flower and $w = 4$ for the (1,3)-flower].

A flower-graph of generation $n > 0$ consists of nodes of degree $k = 2, \dots, 2^n$. The number of nodes of degree 2^m in a flower graph of generation n is

$$N_{2^m, n} = \begin{cases} (w - 2)w^{n-m}, & m < n, \\ w, & m = n. \end{cases} \qquad (3)$$

In scale-free nets, in general, the number of nodes with degree larger or equal to k scales as $N_{>k} \sim k^{1-\gamma}$. According to Eq. (3), $N_{>2^l} \sim w^{-l}$ (for $n \gg l \gg 1$), so the degree exponent of flower graphs is

$$\gamma = 1 + \frac{\ln w}{\ln 2}. \qquad (4)$$

The loopiness exponent of $(1, v)$-flower graphs has been derived in [18]:

$$\alpha = \frac{\ln v}{\ln w}, \qquad (5)$$

that is, $\alpha = \ln 2 / \ln 3$ for the (1,2)-flower, and $\alpha = \ln 3 / \ln 4$ for the (1,3)-flower. Finally, in (1,2)-flowers, the clustering coefficient [1,24] of nodes of degree $k = 2^m$ is $C_k = 2/k$. Then, using the result from Eq. (3), the average clustering coefficient for $F_n(1, 2)$ is [12]

$$\langle C \rangle_n = \frac{6^n + 9}{2^{n-2} \cdot 5(3^n + 3)} \to \frac{4}{5}, \quad \text{as } n \to \infty. \qquad (6)$$

[$(1, v)$-flowers with $v > 2$ have zero clustering.] Both the decay of C_k with growing degree and the *finite* average of the overall clustering coefficient are characteristic properties of real-life complex networks.

III. STOCHASTIC FLOWER GRAPHS

Stochasticity is introduced to (u, v)-flower nets so as to interpolate between the deterministic nets and form what we call *stochastic flower graphs*. The increased flexibility gained in this way makes stochastic flower graphs better suited for modeling of real-life networks. At the same time, they are largely amenable to the type of exact analysis that made flower nets so useful. We focus on the *small-world* case (where $u = 1$), to concur with most everyday life complex nets, but similar constructions could be envisioned for $u > 1$ as well.

During the stochastic generation process, we vary v on a link-by-link basis: each edge e_i (in generation n) is replaced by two parallel paths of lengths 1 and v_i, or $(1, v_i)$-paths, where v_i is selected from $\mathbf{v} = [v_1, v_2, \dots]$ with probabilities $\mathbf{p} = [p_1, p_2, \dots]$, respectively. For simplicity, we restrict the present study to the simplest case of $\mathbf{v} = [2, 3]$ and $\mathbf{p} = [p, 1 - p]$; in other words, each link is replaced by either (1,2)-paths [as in the construction of the (1,2)-flower] with probability p, or by (1,3)-paths [as for the (1,3)-flower] with probability $1 - p \equiv q$. The first few generations of these $([2, 3]; p)$-stochastic flower graphs, denoted hereafter by $S_n([2, 3]; p)$, are shown in Fig. 2. One would expect the properties of $S_n([2, 3]; p)$ to interpolate smoothly between those of (1,2)-flowers and (1,3)-flowers, naively replacing w in Eqs. (1)–(4) with the "average" $\bar{w} = 3p + 4q = 3 + q$. We shall see that this naive expectation works quite well, to some
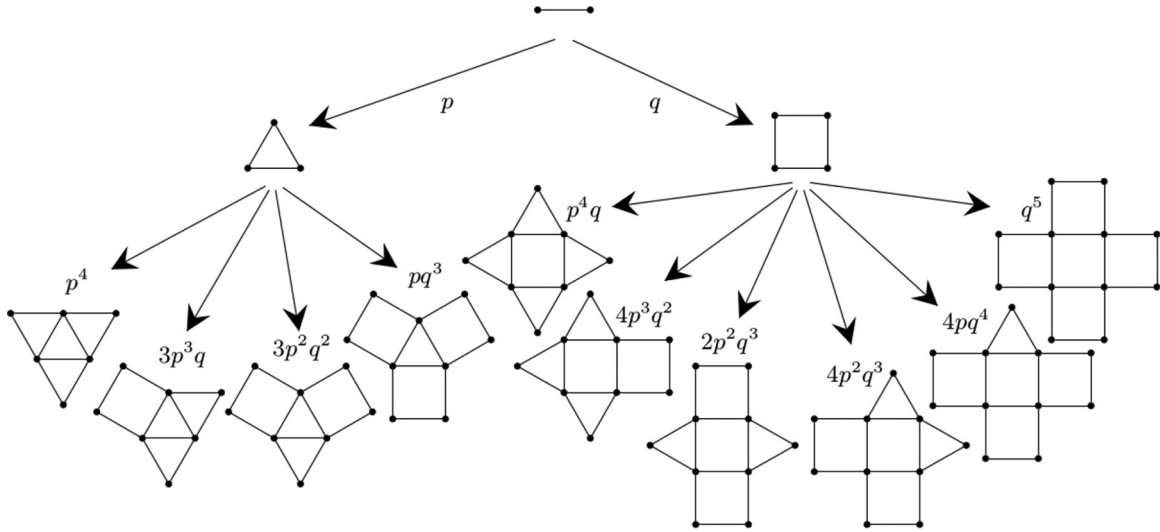
FIG. 2. Ensembles of $S_n([2, 3]; p)$, for $n = 0$, 1, and 2. The probability of obtaining any of the graphs in each ensemble is indicated. Notice the reduced symmetry for most members of the second generation. The graphs become more disordered, or asymmetric, as n gets larger.

extent. For the loopiness exponent, however, naively replacing v with $\bar{v} = 2p + 3q = 2 + q$ and w with $\bar{w}$ in Eq. (5), fails to produce the right answer.

The hub-pasting approach of Fig. 1(b) works also for the stochastic flower graphs: to obtain generation $n + 1$, simply paste either three graphs from generation n (with probability p), or four graphs (with probability q). The graphs to be pasted together are drawn randomly from the ensemble of graphs of generation n, with their respective probabilities. It can be shown by induction that the link-by-link replacement and the pasting of subgraphs result in identical stochastic ensembles. The equivalence of these two methods of construction is important because it is often advantageous to use one form or the other for the analysis of structural properties and dynamical processes by exact analytic recursions.

A. Distributions of links and nodes

Suppose $S_n([2, 3]; p)$ has M links, with probability $\mathbb{P}_n^{\text{links}}(M)$. An iteration of the graph would then yield $3M + m$ links ($m = 0, 1, \ldots, M$), with probability $\binom{M}{m} p^{M-m} q^m$. Thus,

$$\mathbb{P}_{n+1}^{\text{links}}(M) = \sum_{M'} \sum_{\substack{m \\ m + 3M' = M}} \mathbb{P}_n^{\text{links}}(M') \binom{M'}{m} p^{M'-m} q^m,$$

(7)

where the outer sum runs over all possible values of M' and the inner sum runs over all possible values of m such that $3M' + m = M$ ($\mathbb{P}_n^{\text{links}} = 0$ for unattainable values of M). One can then use Eq. (7), along with $\mathbb{P}_0^{\text{links}}(M) = \delta_{M,1}$, to generate all the $\mathbb{P}_n^{\text{links}}(M)$ recursively.

Let $\langle M^r \rangle_n = \sum_M M^r \mathbb{P}_n^{\text{links}}(M)$ denote the rth moment for the number of links in generation n. It then follows from Eq. (7) that

$$\langle M^r \rangle_{n+1} = \sum_M \sum_{m=0}^{M} \mathbb{P}_n^{\text{links}}(M) \binom{M}{m} p^{M-m} q^m (3M + m)^r. \quad (8)$$

Beginning with the first moment (setting $r = 1$), we have

$$\langle M \rangle_{n+1} = \sum_M \sum_{m=0}^{M} \mathbb{P}_n^{\text{links}}(M) \binom{M}{m} p^{M-m} q^m (3M + m)$$

$$= 3\langle M \rangle_n + \sum_M \mathbb{P}_n^{\text{links}}(M) q \frac{\partial}{\partial q} (p + q)^M$$

$$= (3 + q)\langle M \rangle_n, \quad (9)$$

which along with the initial condition $\langle M \rangle_0 = 1$, yields

$$\langle M \rangle_n = (3 + q)^n. \quad (10)$$

This is identical to Eq. (1), obtained by the naive substitution.

For the second moment, similar manipulations lead to the recursion relation

$$\langle M^2 \rangle_{n+1} = (3 + q)^2 \langle M^2 \rangle_n + pq(3 + q)^n, \quad (11)$$

and, along with $\langle M^2 \rangle_0 = 1$, we get

$$\langle M^2 \rangle_n = \left(1 + \frac{pq}{(2 + q)(3 + q)} \right) (3 + q)^{2n}$$

$$- \frac{pq}{(2 + q)(3 + q)} (3 + q)^n. \quad (12)$$

The derivation of recursion relations for higher moments is generally cumbersome, however, the first two moments suffice to explore the convergence of the distribution in the thermodynamic limit (of $n \to \infty$). Indeed, the variance of the distribution for generation n, is

$$\sigma_{M,n}^2 = \langle M^2 \rangle_n - \langle M \rangle_n^2 = qp \frac{(3 + q)^n ((3 + q)^n - 1)}{(2 + q)(3 + q)},$$

so that

$$\frac{\sigma_{M,n}^2}{\langle M \rangle_n^2} = qp \frac{1 - (3 + q)^{-n}}{(2 + q)(3 + q)} \to \frac{qp}{(2 + q)(3 + q)}$$

$$\equiv \mu(p)^2, \text{ as } n \to \infty. \quad (13)$$

In other words, the distribution of the number of links, scaled to their average, $\mathbb{P}_n^{\text{links}}(M/\langle M \rangle_n)$, converges exponentially in
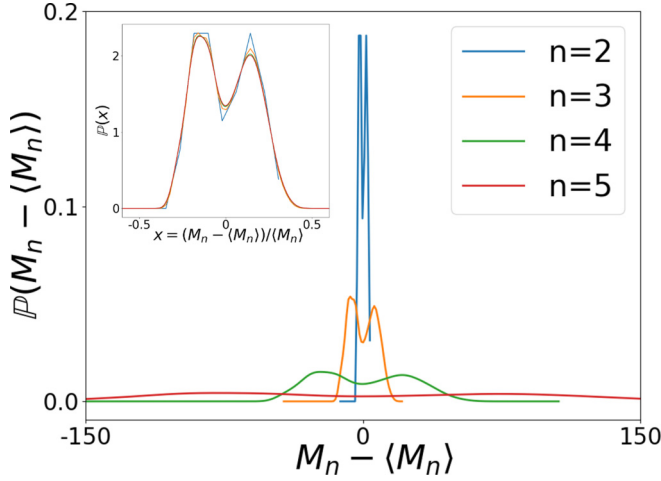
FIG. 3. The distribution of links $\mathbb{P}^{\text{links}}(M)$ when $p = 0.5$ for generations $n = 2$–5 plotted against $M - \langle M \rangle$. Inset: The scaled distribution $\mathbb{P}^{\text{links}}((M - \langle M \rangle)/\langle M \rangle)$ quickly converges to a limit curve with n.

n (or equivalently, as $1/M_n$). Notice that the standard deviation of the converged distribution is a *nonvanishing* fraction of the average: for example, $p = 0.5$ leads to $\mu(p)^2 = 1/35$. On the other hand, since the support of $\mathbb{P}_n(M)$ is $M \in [3^n, 4^n]$ and $\sigma_{M,n} \to \mu(p)(3 + q)^n$, the standard deviation is a vanishing fraction of the support in the thermodynamic limit, and in that sense the distribution can be considered "sharp."

Figure 3 shows the distributions for the number of links $\mathbb{P}_n^{\text{links}}(M)$ in generations 2 through 5 for $p = 0.5$. The bimodal feature is an artifact of the two possibilities for $n = 1$: the $(1,2)$-flower (with probability p) or the $(1,3)$-flower (probability q). If one takes either of these two cases as the starting configuration (instead of K_2), the peaks reappear, due to the split in the next generation, but are now much closer together. This trend continues as the initial seed network gets larger and larger.

Working out the distribution for the number of nodes $\mathbb{P}_n^{\text{nodes}}(N)$ proves more difficult. In particular, beginning with a graph of M_n links and N_n nodes; if m of the links are selected to generate a $(1,3)$-path [and the remaining $M_n - m$ to $(1,2)$-paths], then the resultant graph in generation $n + 1$ consists of

$$
\begin{aligned}
M_{n+1} &= 3M_n + m\,, \\
N_{n+1} &= N_n + M_n + m\,,
\end{aligned}
\tag{14}
$$

links and nodes, respectively. Thus, while links can be analyzed in closed form, as done above, the number of nodes, N_n, cannot be studied independently from the number of links, M_n.

A simple way to overcome this difficulty is to define a generating function for the joint probability distribution for the number of links and nodes in generation n, $\mathbb{P}_n(M, N)$, as

$$
f_n(x, y) = \sum_{M,N} \mathbb{P}(M, N) x^M y^N.
\tag{15}
$$

For example, for the usual initial condition of $S_0([2, 3]; p) = K_2$, $\mathbb{P}_0(M, N) = \delta_{M,1}\delta_{N,2}$, and thus $f_0(x, y) = xy^2$. Since each

link that evolves into a $(1,2)$-path (with probability p) introduces two new links and one new node, while each link that evolves into a $(1,3)$-path (with probability $q = 1 - p$) introduces three new links and two new nodes, the generating function for the evolved graph can be obtained from the mapping: $x \mapsto px \cdot x^2 y + qx \cdot x^3 y^2, y \mapsto y$ (the nodes, represented by y, do not evolve), and thus

$$
f_{n+1}(x, y) = f_n(px^3 y + qx^4 y^2, y).
\tag{16}
$$

The f_n's can then be obtained by iterating this relation.

From the joint probability distribution for links and nodes one can obtain the distribution for links or nodes alone:

$$
\mathbb{P}_n^{\text{links}}(M) = \sum_N \mathbb{P}(M, N); \quad \mathbb{P}_n^{\text{nodes}}(N) = \sum_M \mathbb{P}(M, N).
\tag{17}
$$

The summation over N can be effected by setting $y = 1$, i.e., $f_{n+1}(x, 1) = f_{n+1}^{\text{links}}(x) = f_n^{\text{links}}(px^3 + qx^4)$, leading to the results for $\mathbb{P}^{\text{links}}$ discussed above. Summing over M in this way is more complex, because due to the coupling between M and N the substitution of $x = 1$ should be done only at the end, after having iterated the full expression with x and y to the n-th generation. Instead, we derive equations for the moments of N.

Using Eqs. (15) and (16), we have

$$
\begin{aligned}
\langle N \rangle_{n+1} &= y \frac{\partial}{\partial y} \sum_{M,N} \mathbb{P}_n(M, N)(px^3 y + qx^4 y^2)^M y^N \big|_{\substack{x=1 \\ y=1}} \\
&= (1 + q)\langle M \rangle_n + \langle N \rangle_n.
\end{aligned}
\tag{18}
$$

Putting in the result from Eq. (10) for $\langle M \rangle_n$, along with $\langle N \rangle_0 = 2$, and solving for $\langle N \rangle_n$, we find

$$
\langle N \rangle_n = \frac{1 + q}{2 + q}(3 + q)^n + \frac{3 + q}{2 + q},
\tag{19}
$$

which again coincides with the result for (u, v)-flowers, Eq. (2), with the expected substitution $\bar{w} = (3 + q)$.

For the second moment, we begin similarly with

$$
\begin{aligned}
\langle N^2 \rangle_{n+1} &= y \frac{\partial}{\partial y} y \frac{\partial}{\partial y} \sum_{M,N} \mathbb{P}_n(M, N)(px^3 y + qx^4 y^2)^M y^N \big|_{\substack{x=1 \\ y=1}} \\
&= qp\langle M \rangle_n + (1 + q)^2 \langle M^2 \rangle_n \\
&\quad + 2(1 + q)\langle MN \rangle_n + \langle N^2 \rangle_n.
\end{aligned}
\tag{20}
$$

Thus, in order to solve for $\langle N^2 \rangle_n$, we require a formula for $\langle MN \rangle_n$, which we get from an additional recursion relation:

$$
\begin{aligned}
\langle MN \rangle_{n+1} &= x \frac{\partial}{\partial x} y \frac{\partial}{\partial y} \sum_{M,N} \mathbb{P}_n(M, N)(px^3 y + qx^4 y^2)^M y^N \big|_{\substack{x=1 \\ y=1}} \\
&= (3 + q)\langle MN \rangle_n + (1 + q)(3 + q)\langle M^2 \rangle_n \\
&\quad + pq\langle M \rangle_n.
\end{aligned}
\tag{21}
$$

Starting with the initial condition, $\langle MN \rangle_0 = 2$, and the results of Eqs. (10) and (12) one can then solve explicitly for $\langle MN \rangle_n$; armed with this result and the initial condition of $\langle N^2 \rangle_0 = 4$, one can then use Eq. (20) to obtain an explicit solution for $\langle N^2 \rangle_n$. The actual expressions are lengthy and we instead

present only the large-n asymptotic limits:

$$\langle MN \rangle_n \sim \frac{6(1+q)^2}{(2+q)^2(3+q)}(3+q)^{2n}, \quad \text{as } n \to \infty,$$

$$\langle N^2 \rangle_n \sim \frac{6(1+q)^3}{(2+q)^3(3+q)}(3+q)^{2n}, \quad \text{as } n \to \infty. \quad (22)$$

We conclude that, as in the case for links, $\mathbb{P}_n^{\text{nodes}}(N)$ too has a standard deviation proportional to its average

$$\frac{\sigma_{N,n}^2}{\langle N \rangle_n^2} \to \frac{qp}{(2+q)(3+q)} \equiv \nu(p)^2, \quad (23)$$

and with the same fractional ratio, i.e., $\nu(p) = \mu(p)$ [Eq. (13)].

In what follows we show that despite the spread in the distributions of structural properties such as the number of links and nodes, characteristic exponents of stochastic flowers, such as the degree exponent and the loopiness exponent, attain perfectly well-defined sharp limits.

B. Degree exponent

For a stochastic flower of generation $n > 0$, $S_n([2,3];p)$, the degree sequence is still $\{k\} = \{2, 2^2, \ldots, 2^n\}$, same as for n-generation (u,v)-flowers in general. The average number of nodes with degree larger than 2^m scales like $N_{>2^m} \sim \bar{w}^{n-m}$. However, as implied by the foregoing discussion on the distribution of N, the typical fluctuation in $N_{>2^m}$ is of the same order as its average, that is,

$$N_{>2^m} \sim (1+\sigma)\bar{w}^{n-m}, \quad (24)$$

where $-\nu(p) \lesssim \sigma \lesssim \nu(p)$ is a random variable [c.f. Eq. (23)]. Thus, a plot of $\ln N_{>k}$ vs. $\ln k$ would result in a curve that meanders typically between the two parallel lines of $-m \ln \bar{w} + \ln(1 \pm \nu(p))$ plotted against $m \ln 2$. A linear fit to that curve, as $n \to \infty$, would therefore converge to the degree exponent

$$\gamma = 1 + \frac{\log \bar{w}}{\log 2} = 1 + \frac{\ln(3+q)}{\ln 2} \quad (25)$$

with an error that vanishes as $1/n$ [or more precisely, as $\ln[(1+\nu)/(1-\nu)]/n \ln 2$]. Thus, the degree exponent γ achieves a well-defined value in the thermodynamic limit, despite the nonvanishing fluctuations in N.

C. Loopiness and clustering

The *loopiness* exponent α characterizes how the most likely length of a cycle in the graph, h_*, scales with its order (number of nodes, N) [18]:

$$h_* \sim N^\alpha. \quad (26)$$

The statistics of loops was also explored for two types of random scale-free graphs in [25], and subsequent work explored efficient algorithms for obtaining the statistics of cycles in generic graphs [26,27]. Following the development in [18], we denote the number of loops, or cycles of length h in a graph of generation n, by $C_n(h)$; and the number of paths of length h connecting the two hub nodes in generation n, by $L_n(h)$. Using the recursive construction of pasting copies of generation n to produce generation $n+1$, one can write exact
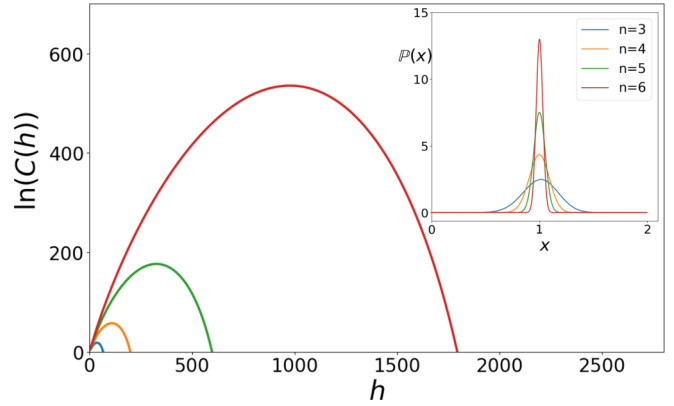

FIG. 4. Plot of $\ln C_n(h)$ vs. h for generations $n = 3, 4, 5$, and 6 in stochastic flower graphs with $p = 1/2$. The inset shows $\mathbb{P}(h/h_*)$— the probability distribution density of h-cycles scaled to their most likely length, h_*.

recursion relations for these quantities. In a (1,2)-flower, for example,

$$L_{n+1}(h) = L_n(h) + \sum_{j+k=h} L_n(j)L_n(k),$$

$$C_{n+1}(h) = 3C_n(h) + \sum_{j+k+l=h} L_n(j)L_n(k)L_n(l). \quad (27)$$

The first equation expresses the fact that a path between the two hubs in generation $n+1$ could go through the single subunit of generation n, or span the other two v-subunits; similarly, cycles of length h could be found in each of the three subunits, or made up of three paths spanning the generation-n subunits across their hubs (second equation). Once again, the generating functions

$$L_n(z) = \sum_h L_n(h)z^h, \quad C_n(z) = \sum_h C_n(h)z^h, \quad (28)$$

simplify the analysis, as the convolution terms become then simple products. Applying these ideas to the $S_n([2,3];p)$ stochastic flowers, we get

$$L_{n+1}(z) = L_n(z) + pL_n^2(z) + qL_n^3(z)$$

$$C_{n+1}(z) = (3+q)C_n(z) + pL_n^3(z) + qL_n^4(z). \quad (29)$$

The first line, in this case, denotes the fact that apart from spanning a path through the single n-subunit, the $(n+1)$-path could go through either two v-subunits (with probability p), or three v-subunits (probability $q = 1 - p$), depending on which v was selected in the recursive construction; and similarly for the second line.

Starting with the initial condition $L_0(z) = z$ and $C_0(z) = 0$, Eqs. (29) can be iterated and $C_n(h)$ can then be obtained from the coefficient of z^h in $C_n(z)$. In Fig. 4 we plot the statistics of cycles obtained in this way for generations $n = 3, 4, 5, 6$ and $p = 1/2$. The results show that the most probable cycle length h_* grows by a factor of 3 from one generation to the next. That is also true for other values of $p < 1$ (not shown). Since $h_* \sim 3^n$ and $N_n \sim \bar{w}^n$, we conclude that the loopiness
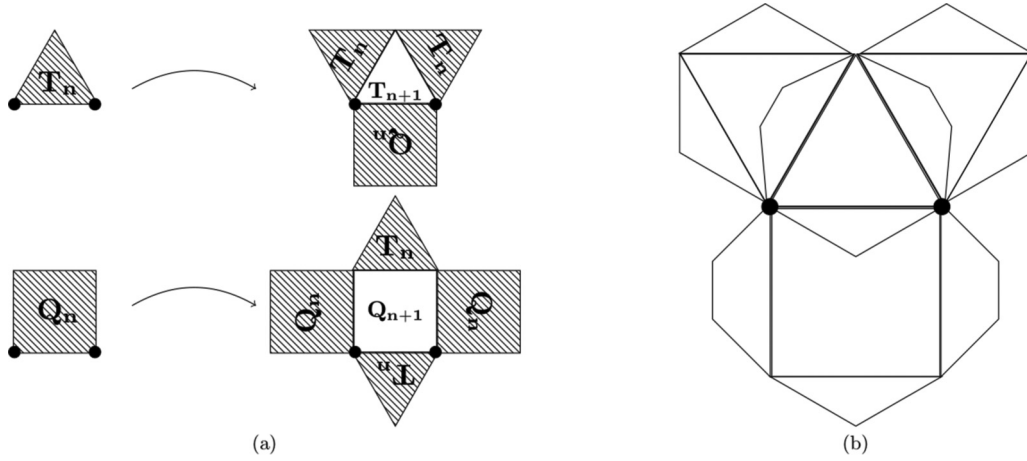
FIG. 5. (a) Generation of T_{n+1} and Q_{n+1} from T_n's and Q_n's by the rule set $\{QTT; TQTQ\}$. The hubs are indicated by large solid circles. (b) T_3, obtained by the rule set $\{QTT; TQTQ\}$.

exponent of $S_n([2, 3]; p)$ is

$$\alpha = \frac{\ln 3}{\ln \bar{w}} = \frac{\ln 3}{\ln (3 + q)}, \quad p < 1. \tag{30}$$

[For the deterministic case of $p = 1$, one obtains (1,2)-flowers, with $\alpha = \ln 2 / \ln 3$.] We see that naively replacing v with $\bar{v} = 2p + 3q$ in Eq. (5) does not work, and instead any *finite* q yields a loopiness exponent characteristic of (1,3)-flowers.

Note that while the recursions (27) for $F_n(1, 2)$ are exact, the same is *not* true for the analogous recursions (29) for the $S_n([2, 3]; p)$. The problem is that $L_n(h)$ and $C_n(h)$ in this case do not have a single sharp value (as for deterministic flowers), but represent a distribution of values. Furthermore, the replacement of the distributions by their average does not work, as these distributions do not converge to δ functions as $n \to \infty$, and the equations are nonlinear [in general, $\langle f(L_n) \rangle \neq f(\langle L_n \rangle)$]. Nevertheless, the naive approach followed here, of replacing the distributions by their average, yields the correct result (30) for α, as shown in the next section.

Clustering in stochastic flowers can be dealt with, at least qualitatively, by following the same procedure of replacing distributions with their averages. Following this approach for $S_n([2, 3]; p)$, we find

$$\langle C \rangle_n = \frac{p}{2^{n-1}} \frac{\bar{w} - 1}{(\bar{w} - 2)\bar{w}^n + \bar{w}} \left[\bar{w} + \frac{\bar{w} - 2}{2\bar{w} - 1}((2\bar{w})^n - 2\bar{w}) \right], \tag{31}$$

and in the limit of $n \to \infty$

$$\langle C \rangle_\infty = \frac{2p(\bar{w} - 1)}{2\bar{w} - 1} = \frac{2p(3 - p)}{7 - 2p}. \tag{32}$$

Thus, $S_n([2, 3]; p)$ retains the feature of finite average clustering in the thermodynamic limit. As expected, $\langle C \rangle_\infty$ interpolates between 0 for $p = 0$, or $F(1, 3)$, to 4/5 for $p = 1$, or $F(1, 2)$.

To better address the tricky issues arising from ensemble spread, we next introduce a class of *mixed* flower-graphs that are fully deterministic, yet closely mimic typical members of the stochastic ensemble of $S_n([2, 3]; p)$.

IV. MIXED FLOWER GRAPHS

We now introduce a class of *mixed flower nets* whose recursive construction is completely deterministic; there is a *single* instance of each mixed flower net. The recursive construction generates two networks in tandem: T_n, that starts from an initial three-cycle seed, T_1, in generation $n = 1$, and Q_n, that starts from a four-cycle seed, Q_1. T_{n+1} and Q_{n+1} are obtained by pasting three or four networks of generation n at their hubs, respectively, according to some prescribed rule. Figure 5(a) shows a specific example of such a rule: T_{n+1} is obtained by pasting Q_n, T_n, T_n together at the hubs, starting with Q_n and proceeding counterclockwise. The hubs of T_{n+1} are the hubs of the first subnet in the sequence (Q_n, in this case). Similarly, Q_{n+1} is obtained by pasting T_n, Q_n, T_n, Q_n at the hubs, starting with T_n and proceeding counterclockwise. The hubs of Q_{n+1} are the hubs of the first subnet in the sequence (T_n). This particular set of rules can then be denoted more succinctly as $\{QTT; TQTQ\}$. Figure 5(b) shows T_3, obtained by this rule set. Since mixed flowers are deterministic constructs, their structural properties can be obtained from *exact* recursion relations, as shown by the few examples below.

A. Size and order of mixed flower graphs

Focusing, as an example, on the mixed flower nets resulting from the rule set $\{QTT; TQTQ\}$, the number of links in T_n ($M_{\triangle,n}$) and in Q_n ($M_{\square,n}$) obey the recursion relation

$$\begin{bmatrix} M_{\triangle,n+1} \\ M_{\square,n+1} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} M_{\triangle,n} \\ M_{\square,n} \end{bmatrix}. \tag{33}$$

This, together with the initial conditions $M_{\triangle,1} = 3, M_{\square,1} = 4$, yields

$$M_{\triangle,n} = \left(\frac{3}{2} + \sqrt{2} \right) r_+^{n-1} + \left(\frac{3}{2} - \sqrt{2} \right) r_-^{n-1},$$

$$M_{\square,n} = \left(2 + \frac{3\sqrt{2}}{2} \right) r_+^{n-1} + \left(2 - \frac{3\sqrt{2}}{2} \right) r_-^{n-1}, \tag{34}$$

where

$$r_\pm = 2 \pm \sqrt{2}$$

are the eigenvalues of the matrix $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix}$. For the number of nodes, $N_{\triangle,n}$ and $N_{\square,n}$, we have

$$\begin{bmatrix} N_{\triangle,n+1} \\ N_{\square,n+1} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} N_{\triangle,n} \\ N_{\square,n} \end{bmatrix} - \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \tag{35}$$

the difference being that here we subtract the nodes identified in the pasting of the subgraphs, to avoid overcounting. In view of the initial conditions $N_{\triangle,1} = 3$, $N_{\square,1} = 4$, the solutions are

$$N_{\triangle,n} = \tfrac{1}{2}(r_+^n + r_-^n) + 1,$$

$$N_{\square,n} = (1 + \sqrt{2})r_+^{n-1} + (1 - \sqrt{2})r_-^{n-1} + 2. \tag{36}$$

We see that the growth of links and nodes in all cases is dominated by the larger eigenvalue of $\mathbf{A}$, r_+: $M_{\triangle,n} \sim M_{\square,n} \sim N_{\triangle,n} \sim N_{\square,n} \sim r_+^n$, as $n \to \infty$. Comparing this observation with the result for stochastic flowers, $\langle M \rangle_n \sim \langle N \rangle_n \sim \bar{w}^n$; $\bar{w} = 3 + q$, we deduce that the rules of growth in our example result in a (deterministic) mixed flower (either T_n or Q_n) that mimics stochastic flowers with $p = 2 - \sqrt{2} \equiv \tilde{p} \approx 0.5858$, $S_n([2, 3]; 2 - \sqrt{2})$. From a different perspective, $\tilde{p} M_{\triangle,n} + \tilde{q} M_{\square,n} \equiv \bar{M}_n$ and $\tilde{p} N_{\triangle,n} + \tilde{q} N_{\square,n} \equiv \bar{N}_n$ are L_1-norms of eigenvectors of $\mathbf{A}$ corresponding to r_+, and $\bar{M}_n$ and $\bar{N}_n$ are identical to $\langle M \rangle_n$ and $\langle N \rangle_n$ of Eqs. (10) and (19), for $p = \tilde{p}$.

B. Degree exponent of mixed flowers

The degree sequence of both T_n and Q_n is $\{k\} = \{2, 2^2, \ldots, 2^n\}$, same as for (u, v)-flower nets and stochastic flowers in general. The number of nodes of degree $k = 2^m$ scales as

$$N_{2^m,n} = N_{n-m+1} - N_{n-m} \sim r_+^{n-m},$$

according to Eq. (35). Therefore, $N_{>2^m} \sim r_+^{-m} \sim (2^m)^{1-\gamma}$, for $1 \ll m \ll n$, and

$$\gamma = 1 + \frac{\ln \tilde{w}}{\ln 2}, \tag{37}$$

where $\tilde{w} = r_+ = 3\tilde{p} + 4\tilde{q} = 3 + \tilde{q}$, exactly as expected.

It is interesting to note that the *ordering* of the pasting of the various copies in the recursive construction has no effect on $\mathbf{A}$, nor on M_n, N_n, $\tilde{p}$, and γ. For example, the rule $\{TQT; QQTT\}$ yields exactly the same results as discussed above for the rule $\{QTT; TQTQ\}$. The deterministic networks resulting from these two rule sets are different nevertheless—the difference manifests in other structural properties of the graphs.

C. Statistics of cycles and clustering of mixed flower graphs

As we have seen, the mixed flower $\{QTT; TQTQ\}$ mimics $S_n([2, 3]; \tilde{p})$, while remaining deterministic. It therefore affords us an opportunity to explore the statistics of cycles in an exact fashion—without the replacing of distributions by their averages, used in the analysis of Sec. III C.

Following a similar notation to that of Sec. III C, the generating functions for the statistics of paths and loops obey
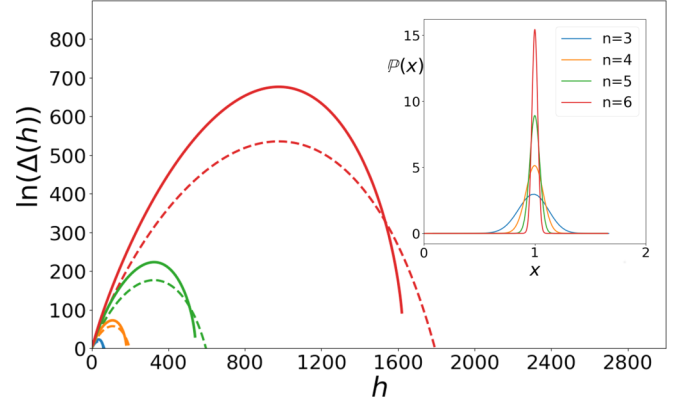

FIG. 6. $\ln \triangle_n(h)$ vs. the cycle length h for the $\{QTT; TQTQ\}$ flower, as computed from Eqs. (38) for generations $n = 3, 4, 5, 6$ (solid curves). For comparison, we also plot the statistics of cycles for the analogous $S_n([2, 3]; \tilde{p})$ flower, as obtained from the approximate Eqs. (29) (broken curves). While disagreement is expected, the identical scaling, of $h_*(n) \sim 3^n$, in both cases, is a nice feature of the approximation. Inset: The distribution of $\triangle_n(h/h_*)$ tends to a δ function as $n \to \infty$.

the recursion relations:

$$
\begin{aligned}
\wedge_{n+1} &= \wedge_n^2 + \sqcap_n && ; && \wedge_1 = z + z^2, \\
\triangle_{n+1} &= \wedge_n^2 \sqcap_n + 2\triangle_n + \square_n && ; && \triangle_1 = z^3, \\
\sqcap_{n+1} &= \wedge_n \sqcap_n^2 + \wedge_n && ; && \sqcap_1 = z + z^3, \\
\square_{n+1} &= \wedge_n^2 \sqcap_n^2 + 2\triangle_n + 2\square_n && ; && \square_1 = z^4.
\end{aligned}
\tag{38}
$$

(Here, we have replaced $L_{\triangle,n}$, $C_{\triangle,n}$, $L_{\square,n}$, and $C_{\square,n}$ with the more visually obvious notations $\wedge_n$, $\triangle_n$, $\sqcap_n$, and $\square_n$, respectively.) We stress that the above relations are exact, as the graphs involved are unique and there is no ensemble spread in the statistics of loops of a given length h; rather, $\wedge_n(h)$, $\triangle_n(h)$, $\sqcap_n(h)$, and $\square_n(h)$ and their corresponding generating functions are all deterministically determined quantities.

In Fig. 6 we present results for $\triangle_n(h)$ for generations $n = 3, 4, 5, 6$. Once again, the most likely length for cycles, h_*, increases by a factor of 3 from one generation to the next, leading to the same loopiness exponent, $\alpha = \ln 3/\ln \tilde{w}$, as was found for stochastic $S_n([2, 3]; p)$ flowers. For comparison, we also show in the plot the statistics of $C_n(h)$ (broken lines), obtained for $S_n([2, 3]; \tilde{p})$ and the initial conditions of a triangle cycle in generation $n = 1$ [$L_1(z) = z + z^2$, $C_1(z) = z^3$]. That the approximate curves for $S_n([2, 3]; \tilde{p})$ and the exact curves for the analogous $\{QTT; TQTQ\}$ do not match comes as no surprise, but it is still reassuring that the approximate approach for stochastic flowers yields the correct scaling and the exact loopiness exponent, nevertheless.

The statistics of cycles provides an example to the importance of the mixed-flowers rule ordering. Indeed, the rule $\{TQT; QQTT\}$ leads to identical statistics for M_n and N_n as the rule $\{QTT; TQTQ\}$, as previously observed, but to different recursion relations for cycles: for example $\wedge_{n+1} = \wedge_n \sqcap_n + \wedge_n$, instead of the first relation in Eq. (38), and likewise for the other relations. This results in the same scaling and loopiness exponent, but the $\triangle_n(h)$ are not equal in the two cases.

Clustering in mixed flower graphs can be studied analytically with (exact) recursion relations. The results are similar to those of deterministic $F_n(1, 2)$ and stochastic $S_n([2, 3]; p)$ with $C_k \sim 1/k$ and $\langle C \rangle_n > 0$ as $n \to \infty$. Consider, for example, the clustering coefficient of nodes of degree $k = 2$ in a deterministic graph of generation n. These nodes arise only from the iteration of links in generation $n - 1$: a link evolving into a triangle yields one two-degree node with clustering coefficient $C = 1$, while each link evolving into a square yields two two-degree nodes with $C = 0$. In the thermodynamic limit of $n \to \infty$ the fraction of nodes evolving into triangles (squares) is $\tilde{p}(1 - \tilde{p})$. In that case, the average clustering of two-degree nodes is $C_2 \to \tilde{p}/(2 - \tilde{p})$. Since nodes of degree 2 constitute a finite fraction of the nodes in the network, even as $n \to \infty$, the overall clustering coefficient is finite $\langle C \rangle_\infty > 0$ (for $\tilde{p} > 0$). At the opposite end, consider nodes of degree $k = 2^n$ in an n-generation network. For the rule $\{QTT; TQTQ\}$ we get

$$C_{2^n}^{\triangle} = \frac{2^n + 2/3}{2^n(2^n - 1)}, \quad C_{2^n}^{\square} = \frac{2^{n-1}}{2^{n-1}(2^n - 1)}, \qquad (39)$$

for nets starting from a triangle or square seed, respectively. In either case, $C_{2^n}^{\triangle, \square} \to 1/(2^n - 1) \sim 1/k$ as $n \to \infty$. The details of how C_k converges to its $1/k$ behavior and the actual value for $\langle C \rangle_\infty$ depend on the initial seed and the ordering in the rules of mixing. Similar remarks can be made for diffusion times between nodes [19,28], to mention one other obvious example where the ordering of mixing is of consequence.

V. CONCLUSION

In conclusion, we have explored two different ways to interpolate between the structural properties of (u, v)-flower graphs. *Stochastic flowers* are obtained by selecting u and v in a *random* fashion, according to a prescribed set of probabilities. The closely related *mixed flower graphs*, on the other hand, are obtained by mixing between the rules of growth of different (u, v)-flowers in a *deterministic* fashion.

Both stochastic and mixed flowers yield themselves to analysis, by exploiting their recursive constructions. Some structural properties of stochastic flowers, such as their size and order, fail to converge in the thermodynamic limit of infinitely large graphs—a property we referred to as "ensemble spread." Nevertheless, characteristic exponents of structural properties, such as the degree exponent and loopiness exponent, converge nicely and might be found by replacing the distributions in the recursion relations with their averages, even when the recursion relations are nonlinear.

Mixed flowers circumvent the problem of ensemble spread altogether, as there is a unique, deterministic configuration for each mixed flower of any size. On the other hand, our examples suggest that they mimic stochastic flowers very closely. It can in fact be shown (future work) that some mixed flowers are members of the ϵ-typical ensemble of stochastic flowers, with respect to the asymptotic equipartition property [29].

For the ease of exposition, we have focused on the simplest examples of stochastic and mixed flowers. Our examples for each can be generalized in several obvious ways: Stochastic flowers could be obtained by choosing randomly between more than two sets of (u, v) values in each recursive growth. Mixed flowers, likewise, could be obtained by mixing more than two sets of rules for (u, v)-flowers, and a finer gradation in the mixing can be effected by designing rules to produce generation $n + m$ from generation n (with $m > 1$, instead of $m = 1$). We have also limited our study to a dearth of structural properties: order, size, degree exponent, statistics of cycles, and clustering. Many other structural properties, such as assortativity [30] and dynamical properties, such as diffusion between nodes [19,28,31], can be studied analytically by similar recursive means.

[1] D. J. Watts and S. H. Strogatz, Nature **393**, 440 (1998).

[2] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[3] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).

[4] S. N. Dorogovtsev and J. F. F. Mendes, Adv. Phys. **51**, 1079 (2002).

[5] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[6] M. E. J. Newman, SIAM Review **45**, 167 (2003).

[7] G. Bianconi and A.-L. Barabási, Europhys. Lett. **54**, 436 (2001).

[8] P. L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000).

[9] Z.-Z. Zhang, L.-L. Rong, and F. Comellas, J. Phys. A **39**, 3253 (2006).

[10] A.-L. Barabási, E. Ravasz, and T. Vicsek, Physica A **299**, 559 (2001).

[11] F. Comellas and M. Sampels, Physica A **309**, 231 (2002).

[12] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Phys. Rev. E **65**, 066122 (2002).

[13] F. Comellas, G. Fertin, and A. Raspaud, Phys. Rev. E **69**, 037104 (2004).

[14] H. D. Rozenfeld, S. Havlin, and D. ben-Avraham, New J. Phys. **9**, 175 (2007).

[15] Z.-Z. Zhang, L.-L. Rong, and S. Zhou, Physica A **377**, 329 (2007).

[16] Y. Sun, K. Hou, and Y. Zhao, Int. J. Mod. Phys. C **24**, 50062 (2013).

[17] L. Barrière, F. Comellas, C. Dalfó, and M. A. Fiol, J. Phys. A **49**, 225202 (2016).

[18] H. Rozenfeld, J. E. Kirk, E. M. Bollt, and D. ben-Avraham, J. Phys. A **38**, 4589 (2005).

[19] E. M. Bollt and D. ben-Avraham, New J. Phys. **7**, 26 (2005).

[20] H. D. Rozenfeld and D. ben-Avraham, Phys. Rev. E **75**, 061102 (2007).

[21] P. Xie, Z.-Z. Zhang, and F. Comellas, Appl. Math. Computation **273**, 1123 (2016).

[22] L. Shan, H. Li, and Z.-Z. Zhang, Theor. Comput. Sci. **677**, 12 (2017).

[23] Y. Zhang, Z.-Z. Zhang, S. Zhou, and J. Guan, Physica A **389**, 3316 (2010).

[24] D. Barmpoutis and R. M. Murray, arXiv:1007.4031 (2010).

[25] G. Bianconi and M. Marsili, J. Stat. Mech. (2005) P06005.

[26] K. Klemm and P. F. Stadler, Phys. Rev. E **73**, 025101(R) (2006).

[27] E. Marinari, G. Semerjian, and V. Van Kerrebroeck, Phys. Rev. E **75**, 066708 (2007).

[28] I. J. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, Phys. Rev. E **64**, 026704 (2001).

[29] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley, New York, 2006).

[30] M. E. J. Newman, Phys. Rev. E **67**, 026126 (2003).

[31] S. S. Manna, Phys. Rev. E **68**, 027104 (2003).