Influence of reciprocal edges on degree distribution and degree correlations

Vinko Zlatić[1,2] and Hrvoje Štefančić[1]
[1]*Theoretical Physics Division, Rudjer Bošković Institute, P.O. Box 180, HR-10002 Zagreb, Croatia*
[2]*INFM–CNR Centro SMC Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Moro 5, 00185 Roma, Italy*
(Received 19 March 2009; revised manuscript received 10 June 2009; published 28 July 2009)

Reciprocal edges represent the lowest-order cycle possible to find in directed graphs without self-loops. Representing also a measure of feedback between vertices, it is interesting to understand how reciprocal edges influence other properties of complex networks. In this paper, we focus on the influence of reciprocal edges on vertex degree distribution and degree correlations. We show that there is a fundamental difference between properties observed on the static network compared to the properties of networks, which are obtained by simple evolution mechanism driven by reciprocity. We also present a way to statistically infer the portion of reciprocal edges, which can be explained as a consequence of feedback process on the static network. In the rest of the paper, the influence of reciprocal edges on a model of growing network is also presented. It is shown that our model of growing network nicely interpolates between Barabási-Albert (BA) model for undirected and the BA model for directed networks.

DOI: 10.1103/PhysRevE.80.016117 PACS number(s): 89.65.−s, 89.20.Hh, 05.65.+b, 89.75.−k

I. INTRODUCTION

Most of real networks as, for example, Internet [1], www [2], or biological webs [3], etc., show interesting topological properties compared with simple models of random graphs. Today, it is well known that most of the real networks exhibit certain properties such as fat tail degree distributions or small-world effect, etc [4,5]. Reading literature on different types of networks, a reader will find a huge number of papers describing different types of correlations in complex networks. Assortativity [6], clustering coefficient [7], reciprocity [8], *k*-cores [9], rich-club coefficient [10], triad significance profile [11], and many other measures related to correlations are frequently reported. Although the identification of these correlations and their reporting in various empirical complex networks has significantly improved our understanding of the field, the question of interrelations of correlation measures naturally emerges. Is it really surprising to find, for example, both the strong rich-club behavior and strong degree correlations in the network? The answer is—clearly not. Today, it is a well-known fact that most of the real networks are correlated. Nevertheless, there is still a huge gap in our understanding of how exactly certain types of correlation-related measures influence other correlation-related measures. In this paper, we will try to bridge a part of that gap relating the reciprocity measure to degree sequence and to degree correlations.

The directed network represents an interesting subgroup of real networks, which allow movement in just one direction. Reciprocity [8] of complex networks is a fraction of directed edges, which have their counterparts showing in the opposite direction compared to the total number of edges, i.e., every bidirectional arrow in the directed graph is considered as composed of two reciprocal edges. It can be said that it is in fact a measure of how much is directed network similar to undirected one. Reciprocity was also shown as an important feature for percolation on directed networks [12]. In previous work, we have also shown that the reciprocity is a very stable correlation measure of all the investigated measures in the case of Wikipedia networks ensemble [13].

In the paper [14], the influence of the broad class of degree correlations on the reciprocity measure is described and quantified. In this paper, our aim is exactly the opposite, i.e., to find a way to quantify the influence of reciprocal edges on degree correlations and degree distributions of complex networks. More precisely, first we focus on random addition of reciprocal edges in the underlying static network. This process of transformation of unidirectional edges into the bidirectional ones can be justified in many ways. First, it is the simplest possible choice of creating reciprocal links in the network, which already has some structure. Second, it is the logical model of information return in the case of the information networks such as an e-mail or a www network.

II. INFLUENCE OF RECIPROCITY ON DEGREE CORRELATIONS IN STATIC NETWORKS

As a null hypothesis, it is reasonable to suppose that the mutual functional relationships between vertices are distributed completely randomly over the whole network. In other words that means that we suppose that the mutual functional relationship does not depend in any way on the degrees of vertices or any other measurable network quantity. We suppose that reciprocal edges can form only between vertices which are already connected. The question we address is the following: how reciprocal edges formed in this way transform the degree distribution and correlations between degrees in a given complex network? The model is defined with the initial network as an input. On the starting network, the unidirectional edges are transformed with probability p into reciprocal ones and with probability $1-p$ are left unchanged. After this process, the properties of the new network are measured again. We show that using the inversion of the transformation process we can infer the most probable starting configuration of the network.

In the following, we distinguish between bidirectional edge as a single edge, which is pointing in two directions contrary to some other analyses in which reciprocal edges are represented as two unidirectional edges connecting two vertices i and j in opposing directions [8,14]. Every vertex of the network is described by three numbers: the first represents exclusive in degree of the vertex, the second represents exclusive out degree of the vertex, and the third represents its bidirectional degree. Exclusive in (or out) degree is the number of unidirectional in (or out) edges which are attached to a given vertex. In the following, these degrees will be designated as $\boldsymbol{k} \equiv (k_i, k_o, k_r)$.

From the initial network, we can extract the following information as initial conditions for the observed transformation process: L number of edges; N number of vertices; $L^{\rightarrow}$ number of *strictly* unidirectional edges; $L^{\leftrightarrow}$ number of *strictly* bidirectional edges; $L(\boldsymbol{k} \rightarrow \boldsymbol{q})$ number of unidirectional edges, which are pointing from the vertex of degrees $\boldsymbol{k}$ to the vertex of degrees $\boldsymbol{q}$; $L(\boldsymbol{k} \leftrightarrow \boldsymbol{q})$ number of bidirectional edges, which are connecting the vertices of degrees $\boldsymbol{k}$ and degrees $\boldsymbol{q}$; and $N(\boldsymbol{k})$ number of vertices with degrees $\boldsymbol{k}$. In the following text, the convention will be that if we observe unidirectional edges with $\boldsymbol{k}$ are designated degrees of the starting vertex while with $\boldsymbol{q}$ are designated the degrees of the ending vertex.

With these properties, it is possible to represent adequately maximally random graph, as well as graphs with any given degree distribution and correlations between degrees of neighboring vertices. Information on correlations between degrees of neighboring vertices existing in the network is given by the frequency of edges, which connect different vertices,

$$\mathcal{P}(\boldsymbol{k}, \rightarrow, \boldsymbol{q}) = \frac{L(\boldsymbol{k} \rightarrow \boldsymbol{q})}{L},$$

$$\mathcal{P}(\boldsymbol{k}, \leftrightarrow, \boldsymbol{q}) = \frac{L(\boldsymbol{k} \leftrightarrow \boldsymbol{q})}{L}. \tag{1}$$

Probabilities $\mathcal{P}(\boldsymbol{k}, \rightarrow, \boldsymbol{q})$ and $\mathcal{P}(\boldsymbol{k}, \leftrightarrow, \boldsymbol{q})$ are defined as joint probabilities that the vertex of degrees $\boldsymbol{k}$ is *pointing to/is connected to* the vertex of degrees $\boldsymbol{q}$, with unidirectional edge in the former and with the bidirectional edge in the latter case. The proper summation of these joint probabilities is

$$\sum_{\substack{q_{i,o}=0\\ q_r=1}}^{\infty} \sum_{\boldsymbol{k} \geq \boldsymbol{q}}^{\infty} \mathcal{P}(\boldsymbol{k}, \leftrightarrow, \boldsymbol{q}) + \sum_{\substack{k_{i,r}=0\\ k_o=1}}^{\infty} \sum_{\substack{q_{o,r}=0\\ q_i=1}}^{\infty} \mathcal{P}(\boldsymbol{k}, \rightarrow, \boldsymbol{q}) = 1, \tag{2}$$

where $\boldsymbol{k} \geq \boldsymbol{q}$ means that every degree of the vector with degrees $\boldsymbol{k}$ is greater or equal to the corresponding degree of the vector with degrees $\boldsymbol{q}$. The summations are different for the bidirectional edges compared with unidirectional edges because $\mathcal{P}(\boldsymbol{k}, \leftrightarrow, \boldsymbol{q}) = \mathcal{P}(\boldsymbol{q}, \leftrightarrow, \boldsymbol{k}) = \frac{L(\boldsymbol{k} \leftrightarrow \boldsymbol{q})}{L}$. Although the statistics of degrees of neighboring vertices gives relevant information on correlation structure of the given network and the one-vertex statistics can, in principle, be easily calculated from that information, from an analytical aspect we will show that is much easier to explicitly calculate one-vertex degree correlations described with

$$P(\boldsymbol{k}) = \frac{N(\boldsymbol{k})}{N}, \tag{3}$$

where $P(\boldsymbol{k})$ represents the joint probability that the vertex has degrees k_i, k_o i k_r.

In the studied model, every unidirectional edge is transformed in a bidirectional one with the probability p. The equation which expresses a new joint probability that a vertex of degrees $\boldsymbol{k}'$ is pointing to a vertex of degrees $\boldsymbol{q}'$ via unidirectional edge is

$$\mathcal{P}'(\boldsymbol{k}', \rightarrow, \boldsymbol{q}') = \sum_{\mathcal{C}} \mathcal{T}(\boldsymbol{k}', \rightarrow, \boldsymbol{q}' | \boldsymbol{k}, \rightarrow, \boldsymbol{q}) \mathcal{P}(\boldsymbol{k}, \rightarrow, \boldsymbol{q}), \tag{4}$$

where $\mathcal{T}$ represents the transition probability for the given process. A prime on the probabilities means that they are calculated after the transformation process; while the absence of a prime means that the probabilities are calculated from the given starting network. The summation is run over the set $\mathcal{C}$ of unidirectional edges, which fulfill the following conditions. (i) The number of neighbors $S^{(j)} = k_i^{(j)} + k_o^{(j)} + k_r^{(j)}$ is conserved for every vertex j because transformation process does not create new edges between vertices which are not neighbors already; (ii) before and after the transformation process, the following relations hold: $k_i'^{(j)} \leq k_i^{(j)}$, $k_o'^{(j)} \leq k_o^{(j)}$, and $k_r'^{(j)} \geq k_r^{(j)}$. The transition probability $\mathcal{T}$ written in a more detail is

$$\begin{aligned} &\mathcal{T}(\boldsymbol{k}', \rightarrow, \boldsymbol{q}' | \boldsymbol{k}, \rightarrow, \boldsymbol{q}) \\ &\quad = (1-p) \mathcal{T}(k_i' | k_i) \mathcal{T}(k_o' - 1 | k_o - 1) \mathcal{T}(q_i' - 1 | q_i - 1) \mathcal{T}(q_o' | q_o). \end{aligned} \tag{5}$$

The first part of Eq. (5) is the probability that the unidirectional edge stays unidirectional after the transformation process. Other unidirectional edges attached to the vertices can be changed with probability p or stay unidirectional with probability $1-p$. The fact that in this case only other edges are monitored is represented in equation by subtracting one edge from the out degree of the out vertex and the in degree of the in vertex. Probabilities of the transition $\mathcal{T}(x' | x)$, where x represents any of the aforementioned degrees are binomial probabilities, i.e.,

$$\mathcal{T}(x' | x) = \binom{x}{x'} p^{x-x'} (1-p)^{x'}. \tag{6}$$

New joint probability distribution of degrees of the vertices connected via the bidirectional edge $\mathcal{P}'(\boldsymbol{k}, \leftrightarrow, \boldsymbol{q})$ is

$$\begin{aligned} \mathcal{P}'(\boldsymbol{k}', \leftrightarrow, \boldsymbol{q}') = \sum_{\mathcal{C}} [&\mathcal{T}(\boldsymbol{k}', \leftrightarrow, \boldsymbol{q}' | \boldsymbol{k}, \leftrightarrow, \boldsymbol{q}) \mathcal{P}(\boldsymbol{k}, \leftrightarrow, \boldsymbol{q}) \\ &+ \mathcal{T}(\boldsymbol{k}', \leftrightarrow, \boldsymbol{q}' | \boldsymbol{k}, \rightarrow, \boldsymbol{q}) \mathcal{P}(\boldsymbol{k}, \rightarrow, \boldsymbol{q})] \\ &+ \sum_{\mathcal{C}'} \mathcal{T}(\boldsymbol{q}', \leftrightarrow, \boldsymbol{k}' | \boldsymbol{k}, \rightarrow, \boldsymbol{q}) \mathcal{P}(\boldsymbol{k}, \rightarrow, \boldsymbol{q}). \end{aligned} \tag{7}$$

The set $\mathcal{C}$ is fulfilling all the conditions as in Eq. (4), while for the set $\mathcal{C}'$ the following relations hold. (i) The number of neighbors $S^{(j)}=k_i^{(j)}+k_o^{(j)}+k_r^{(j)}$ is conserved for every vertex j because transformation process does not create new edges between vertices, which are not neighbors already; (ii) before and after the transformation process, the following relations hold: $q_i'^{(j)}\le k_i^{(j)}$, $q_o'^{(j)}\le k_o^{(j)}$, and $q_r'^{(j)}\ge k_r^{(j)}$. In this equation, the probabilities of transition have the similar meaning as in Eq. (5),

$$\mathcal{T}(\boldsymbol{k}', \leftrightarrow ,\boldsymbol{q}'|\boldsymbol{k}, \leftrightarrow ,\boldsymbol{q}) = \mathcal{T}(k_i'|k_i)\mathcal{T}(k_o'|k_o)\mathcal{T}(q_i'|q_i)\mathcal{T}(q_o'|q_o), \tag{8}$$

while

$$\mathcal{T}(\boldsymbol{k}', \leftrightarrow ,\boldsymbol{q}'|\boldsymbol{k}, \rightarrow ,\boldsymbol{q}) = p\mathcal{T}(k_i'|k_i)\mathcal{T}(k_o'|k_o-1)\mathcal{T}(q_i'|q_i-1)\mathcal{T}(q_o'|q_o), \tag{9}$$

and

$$\mathcal{T}(\boldsymbol{q}', \leftrightarrow ,\boldsymbol{k}'|\boldsymbol{k}, \rightarrow ,\boldsymbol{q}) = p\mathcal{T}(k_o'|q_o)\mathcal{T}(k_i'|q_i-1)\mathcal{T}(q_o'|k_o-1)\mathcal{T}(q_i'|k_i). \tag{10}$$

It is important to notice that in Eq. (8) we do not need to worry about the edge which connects neighboring vertices because it is a process invariant bidirectional edge. Parameter p in Eqs. (9) and (10) represents the probability of the transformation of unidirectional edge, which connects neighbors to bidirectional edge. Individual probabilities of transition in Eqs. (8)–(10) are again binomial (6).

Similar equations are easy to write for the transformation process of one-vertex statistics. More precisely, equation

$$P'(\boldsymbol{k}') = \sum_{\mathcal{C}} \binom{k_i}{k_i'}\binom{k_o}{k_o'} p^{k_i-k_i'+k_o-k_o'}(1-p)^{k_i'+k_o'}P(\boldsymbol{k}) \tag{11}$$

describes their probability of transformation of joint probability of one-vertex degrees described with $P(\boldsymbol{k})$ into the joint probability of one-vertex degrees $P'(\boldsymbol{k}')$.

In all the aforementioned equations for the transformation process, we used the joint probability statistics because the total number of edges over which this statistics is obtained does not change with the process. However, to calculate the correlations existing in the network it is convenient to use the conditional probabilities with respect to the type of the edge which connects two neighboring vertices. The usual equation for conditional probability can be employed as

$$\mathcal{P}'(\boldsymbol{k}',\boldsymbol{q}'|\rightarrow) = \mathcal{P}'(\boldsymbol{k}', \rightarrow ,\boldsymbol{q}')/P'(\rightarrow). \tag{12}$$

The probability that two neighboring vertices are connected with unidirectional edge is $P'(\rightarrow)=P(\rightarrow)(1-p)$ and the probability that a unidirectional edge exists before the transformation process is $P(\rightarrow)=L^{\rightarrow}/L$, where $L^{\rightarrow}$ is the number of unidirectional edges before the transformation process. Similar equation holds for the conditional probability that two vertices are connected via bidirectional edge,

$$\mathcal{P}'(\boldsymbol{k}',\boldsymbol{q}'|\leftrightarrow) = \mathcal{P}'(\boldsymbol{k}', \leftrightarrow ,\boldsymbol{q}')/P'(\leftrightarrow), \tag{13}$$

where $P'(\leftrightarrow)=P(\leftrightarrow)+pP(\rightarrow)$ and $P(\leftrightarrow)=L^{\leftrightarrow}/L$. The $L^{\leftrightarrow}$ is the number of bidirectional edges before the transformation process.

In the literature on complex networks [4], it is usual to use statistics of average degree of neighbors of a given vertex to represent the correlations of degrees in the network. Such a measure is usually represented with figures of the average neighboring degree dependence on the degree of the monitored vertex. It is easy to verify if the network is correlated or not by simple inspection of such a figure. In order to analytically describe degree-degree correlations resulting from this process, we will use a different measure much more common in usual statistical analysis. The observed and calculated correlations are just the noncentralized product moments of independent variables. In the following, we will loosely use the term correlations for all of the calculated statistical quantities both for statistics obtained on one vertex via Eq. (11) or for statistics of degrees on connected pairs of vertices calculated via Eqs. (4) and (7).

The equation for calculation of one-vertex statistics is

$$\langle k_i'k_o'\rangle = \sum_{\boldsymbol{k}'} k_i'k_o'P'(\boldsymbol{k}'), \tag{14}$$

for the case of in-out degree correlations. All other one-vertex degree correlations are calculated in a similar way. There are two different equations with which we calculate two-vertex degree correlations. The first one is for the calculation of two-vertex degree correlations connected via unidirectional edges. This type of correlations is designated as $\langle\cdot|\rightarrow\rangle$ in order to distinguish them from the two-vertex degree correlations calculated via bidirectional edges $\langle\cdot|\leftrightarrow\rangle$. The equation for the in-out degree correlations of the vertices connected via unidirectional edge is

$$\langle k_i'q_o'|\rightarrow\rangle = \sum_{\boldsymbol{k}'\boldsymbol{q}'} k_i'q_o'\mathcal{P}'(\boldsymbol{k}',\boldsymbol{q}'|\rightarrow). \tag{15}$$

Using Eqs. (4) and (12), we can calculate degree correlations of unidirectionally connected pairs of vertices. The calculation of degree correlations of bidirectionally connected pairs of vertices is a bit trickier. The approximate equation for in-out degree correlations in this case is

$$2\langle k_i'q_o'|\leftrightarrow\rangle \simeq \sum_{\boldsymbol{k}'\boldsymbol{q}'} k_i'q_o'\mathcal{P}'(\boldsymbol{k}',\boldsymbol{q}'|\leftrightarrow). \tag{16}$$

This equation is just an approximation because in order to semianalytically calculate expected correlations after the transformation process, we have to sum over all degrees $\boldsymbol{k}'$ and $\boldsymbol{q}'$, thus, including every bidirectionally connected pair two times, except for the pairs which have *exactly the same* degrees. The equation could be improved by taking into account a new class of correlations just between the bidirectionally connected pairs of vertices, which have the same degrees; but as it will be shown later, this approximation is more than good enough for estimating expected correlations for most of large enough networks. Using Eqs. (7) and (13), it is possible to calculate degree correlations of bidirectionally connected pairs of vertices.

If the calculated correlations are less or greater than the expected in the random network of the same degree distribution (configuration model) [15,16] then the network exhibits a structural tendency that the vertices of larger degrees are mutually connected less or more frequently. For example, the expected in-out degree correlations of unidirectionally connected pairs of vertices are

$$\langle k_i' q_o' | \rightarrow \rangle_{Rand} = \sum_{\boldsymbol{k}'\boldsymbol{q}'} k_i' q_o' \mathcal{P}'(\boldsymbol{k}'|\rightarrow)\mathcal{P}'(\boldsymbol{q}'|\rightarrow). \quad (17)$$

In Eq. (17), we use the conditional probabilities that the two neighboring vertices are connected via unidirectional edge. The $\mathcal{P}'(\boldsymbol{k}'|\rightarrow)$ designates that the vertex of degrees $\boldsymbol{k}'$ is the starting vertex of the conditioned edge, while $\mathcal{P}'(\boldsymbol{q}'|\rightarrow)$ designates that the vertex with degrees $\boldsymbol{q}'$ is the end vertex of that edge. In the case of vertex statistics, it can be written as

$$\mathcal{P}'(\boldsymbol{k}'|\rightarrow) = \frac{k_o}{\langle k_o \rangle} P'(\boldsymbol{k}') \quad (18)$$

because of the fact that the vertex certainly has an outgoing edge, which connects it to a neighboring vertex.

It is important to understand the fine difference between the correlations between degrees of neighboring vertices and correlations of degrees of one vertex. In the following, the correlations of degrees of neighboring vertices will be designated with the conditional type of edge to make the distinction from one-vertex correlations.

For example, the in-in-degree correlations of neighboring vertices can be calculated using the expression

$$\langle k_i' q_i' | \rightarrow \rangle = \sum_{\boldsymbol{k}'\boldsymbol{q}'} k_i' q_i' \mathcal{P}'(\boldsymbol{k}', \boldsymbol{q}'|\rightarrow), \quad (19)$$

and using Eq. (4) and (12) the final solution is

$$\langle k_i' q_i' | \rightarrow \rangle = (1-p)^2 \langle k_i q_i | \rightarrow \rangle + p(1-p)\langle k_i | \rightarrow \rangle. \quad (20)$$

The other examples are in-in-degree correlations of bidirectionally connected vertices calculated with the presented scheme,

$$\langle k_i' q_i' | \leftrightarrow \rangle = \frac{(1-p)^2 \langle k_i q_i | \leftrightarrow \rangle P(\leftrightarrow)}{P(\leftrightarrow) + pP(\rightarrow)} + \frac{(1-p)^2 p P(\rightarrow)(\langle k_i q_i | \rightarrow \rangle - \langle k_i | \rightarrow \rangle)}{P(\leftrightarrow) + pP(\rightarrow)}. \quad (21)$$

In this case, the factors $P(\rightarrow)$ and $P(\leftrightarrow)$ are also present because they did not cancel out as they did in Eq. (20).

To compute Eq. (19) and other possible correlations, the following set of relations is useful:

$$\sum_{l=0}^{n} \binom{n}{l} l^2 \left(\frac{1-p}{p}\right)^l = \frac{n(1-p)}{p^n}[(1-p)n + p],$$

$$\sum_{l=0}^{n} \binom{n}{l} l \left(\frac{1-p}{p}\right)^l = \frac{n(1-p)}{p^n},$$

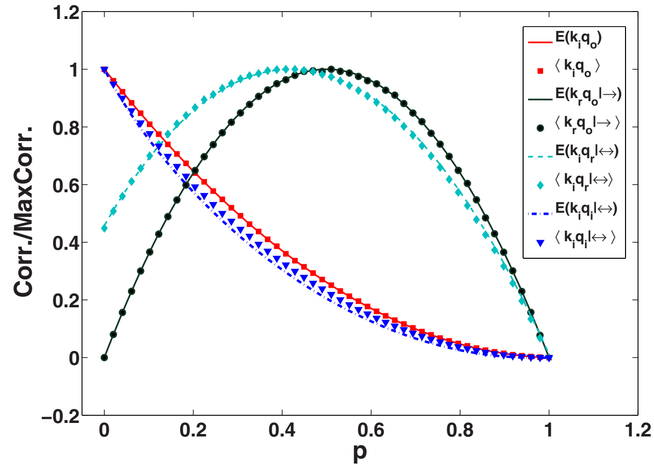

FIG. 1. (Color online) In this figure, we present an excellent agreement between simulations of transformation process and our semianalytical treatment. On x coordinate is parameter p, which represents the probability of edge transformation. The y-coordinate represents the ratio of the numerical value of the given correlations and the maximum value of that correlation, in such a way that we can present the correlations with very different magnitudes in the same figure. With $E(\cdot)$ are designated expectations calculated from equations similar to Eqs. (20) and (21), as an input we used the measured correlations in the network before the start of the process. Simulations are designated with $\langle \cdot \rangle$. The initial networks we used for this figure are a Barabási-Albert directed network of 10^5 vertices for the case of one-vertex degree correlations and for the correlations of unidirectionally connected vertices and a Spanish Wikipedia for the case of the degree correlations of bidirectionally connected vertices. The simulations are averaged over 1000 realizations of the process in the case of a directed BA networks and over 100 realizations in the case of the Spanish Wikipedia. It can be seen that the expected values of degree correlations between the bidirectionally connected vertices deviate a bit from simulational results, which is probably a consequence of approximation (16).

$$\sum_{l=0}^{n} \binom{n}{l} \left(\frac{1-p}{p}\right)^l = \frac{1}{p^n},$$

$$\sum_{l=1}^{n} \binom{n-1}{l-1} l \frac{(1-p)^{l-1}}{p^l} = \frac{n(1-p)+p}{p^n},$$

$$\sum_{l=1}^{n} \binom{n-1}{l-1} \frac{(1-p)^{l-1}}{p^l} = \frac{1}{p^n}. \quad (22)$$

All the correlations are calculated semianalytically and checked with numerical simulations on networks of different sizes, density of edges, and starting correlation structure. In Fig. 1, we present some one-vertex, two-vertex unidirectional, and two-vertex bidirectional correlations calculated semianalytically and compared to simulations. The computation of all elementary correlations can be more elegantly described with matrices of transformation $\mathbf{T}$. If the observed correlations are represented as components of a "correlation vector," the studied process can be described with two different transformation matrices—one which transforms vector of one-vertex correlations $\mathbf{T}_{1v}$ and the second which transforms the vectors of the neighboring pairs correlations $\mathbf{T}_{2v}$.

III. TRANSFORMATION MATRIX

Two transformation matrices differ one from another. The matrix of one-vertex correlations is the square matrix of rank 8. The complete one-vertex statistics of interest can be written as a vector $\mathbf{S}$ with components $\mathbf{S}^{\mathbf{T}}=\{\langle k_i\rangle =\langle k_o\rangle,\langle k_r\rangle,\langle k_i^2\rangle,\langle k_o^2\rangle,\langle k_r^2\rangle,\langle k_ik_o\rangle,\langle k_ik_r\rangle,\langle k_ok_r\rangle\}$. The expected correlations calculated after the transformation process for one-vertex correlations can now be written as a simple linear equation,

$$\langle\mathbf{S}'(p)\rangle=\mathbf{T}_{1v}(p)\mathbf{S}(0), \tag{23}$$

where $\mathbf{S}'(p)$ represents the vector of correlations after fraction p of unidirectional edges is transformed into bidirectional edges. Average in degree is always equal to the average out degree and is therefore eliminated from the matrix. The transformation matrix for one-vertex correlations $\mathbf{T}_{1v}$ is

$$\mathbf{T}_{1v}=\begin{pmatrix} (1-p) & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 2p & 1 & 0 & 0 & 0 & 0 & 0 & 0\\ p(1-p) & 0 & (1-p)^2 & 0 & 0 & 0 & 0 & 0\\ p(1-p) & 0 & 0 & (1-p)^2 & 0 & 0 & 0 & 0\\ 2p(1-p) & 0 & p^2 & p^2 & 1 & 2p^2 & 2p & 2p\\ 0 & 0 & 0 & 0 & 0 & (1-p)^2 & 0 & 0\\ -p(1-p) & 0 & p(1-p) & 0 & 0 & p(1-p) & (1-p) & 0\\ -p(1-p) & 0 & 0 & p(1-p) & 0 & p(1-p) & 0 & (1-p) \end{pmatrix}. \tag{24}$$

This matrix also has its inverse,

$$\mathbf{T}_{1v}^{-1}=\begin{pmatrix} \frac{1}{1-p} & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ \frac{2p}{p-1} & 1 & 0 & 0 & 0 & 0 & 0 & 0\\ -\frac{p}{(1-p)^2} & 0 & \frac{1}{(1-p)^2} & 0 & 0 & 0 & 0 & 0\\ -\frac{p}{(1-p)^2} & 0 & 0 & \frac{1}{(1-p)^2} & 0 & 0 & 0 & 0\\ -\frac{2p}{(1-p)^2} & 0 & \frac{p^2}{(1-p)^2} & \frac{p^2}{(1-p)^2} & 1 & \frac{2p^2}{(1-p)^2} & \frac{2p}{(1-p)} & \frac{2p}{(1-p)}\\ 0 & 0 & 0 & 0 & 0 & \frac{1}{(1-p)^2} & 0 & 0\\ \frac{p}{(1-p)^2} & 0 & -\frac{p}{(1-p)^2} & 0 & 0 & -\frac{p}{(1-p)^2} & \frac{1}{(1-p)} & 0\\ \frac{p}{(1-p)^2} & 0 & 0 & -\frac{p}{(1-p)^2} & 0 & -\frac{p}{(1-p)^2} & 0 & \frac{1}{(1-p)} \end{pmatrix}. \tag{25}$$

The inverse matrix can be of interest for the statistical analysis of real networks. If there is a reason to believe that the bidirectional edges are completely random consequence of the mentioned transformation process and if one has a network model, which does not take into account the bidirectional edges, it can be tested using the inverse transformation matrix. It is easy to calculate the parameter p as $p=\frac{L^{\leftrightarrow}}{L}$, where $L^{\leftrightarrow}$ represents the number of bidirectional edges in the network of interest, while L is the total number of edges. Then using equation

$$\langle\mathbf{S}(0)\rangle=\mathbf{T}_{1v}^{-1}(p)\mathbf{S}'(p), \tag{26}$$

one can find the vector of expected degree correlations before the transformation process. Comparing then $\langle\mathbf{S}(0)\rangle$ with the vector of correlations obtained by the model, one can gain additional information on the structural role of bidirectional edges and/or quality of the studied model. Such assumptions could be a good null model for a number of real world applications such as the analysis of communication or traffic networks. In the companion paper [17], we will present an application of this framework to the Wikipedia networks as a case study.

The transformation of two-vertex correlations is given by the expression

$$\langle \mathbf{S}'_{2v}(p)\rangle = \mathbf{T}_{2v}(p)\mathbf{S}_{2v}(0) + \mathbf{b}(p), \tag{27}$$

where $\mathbf{S}_{2v}$ presents vector of two-vertex product moments and $\mathbf{b}(p)$ additional vector containing terms such as p^2 given in Eq. (28). The matrix of two-vertex degree correlations is too big to be presented. For example, the equation for the expected correlation of two bidirectional degrees of nodes connected via unidirectional edge is

$$\begin{aligned}\langle k'_r q'_r | \rightarrow \rangle = & \langle k_r, q_r | \rightarrow \rangle - p(\langle q_r | \rightarrow \rangle + \langle k_r | \rightarrow \rangle) \\ & + p(\langle k_i q_r | \rightarrow \rangle + \langle k_o q_r | \rightarrow \rangle + \langle k_r q_o | \rightarrow \rangle \\ & + \langle k_r q_i | \rightarrow \rangle) + p^2(\langle k_i q_i | \rightarrow \rangle + \langle k_i q_o | \rightarrow \rangle \\ & + \langle k_o q_i | \rightarrow \rangle + \langle k_o q_o | \rightarrow \rangle - \langle k_i | \rightarrow \rangle - \langle k_o | \rightarrow \rangle \\ & - \langle q_o | \rightarrow \rangle - \langle q_i | \rightarrow \rangle + 1). \end{aligned} \tag{28}$$

Nevertheless, there is enough information for interested reader to be able to reconstruct the two-vertex transformation matrix completely.

It is important to note that correlations $\langle k_i | \rightarrow \rangle$ and $\langle k_o | \rightarrow \rangle$ of the exit vertex are very different from the correlations obtained with one-vertex statistics. It can be written using usual one-vertex statistics as $\langle k_i | \rightarrow \rangle = \frac{\langle k_i k_o \rangle}{\langle k_o \rangle}$, while $\langle k_o | \rightarrow \rangle = \frac{\langle k_o^2 \rangle}{\langle k_o \rangle}$. Similarly, the correlations of in vertex written by means of one-vertex statistics are $\langle q_i | \rightarrow \rangle = \frac{\langle q_i^2 \rangle}{\langle k_o \rangle}$, while $\langle q_o | \rightarrow \rangle = \frac{\langle q_i q_o \rangle}{\langle k_o \rangle}$.

It can be shown that the correlations arising from the transformation process are different from those that we would expect from the noncorrelated network. The comparison between real correlations in the network and the ones expected from the configuration model is shown in Fig. 2.

Up to now, we have shown that degree correlations can be strongly influenced by the addition of bidirectional edges. If the initial network is already very correlated, the transformation process tends to amplify these correlations compared to the configuration model. An example of such strongly correlated networks is the Barabási-Albert (BA) directed network that we used for comparison [18]. In this model, new vertices are attached to the old ones proportionally to the sum of in degree of the old vertices and some parameter a. In our case, the parameter $a=1$ was chosen in the simulations and the starting network for BA evolution was Gilbert network of 10^3 vertices connected with probability 0.01. Other values of parameter a were tested as well. It is known that the properties of the directed BA network do not depend on size or degree sequence of initial network in the thermodynamical limit. The final size of the simulated networks was 10^6.

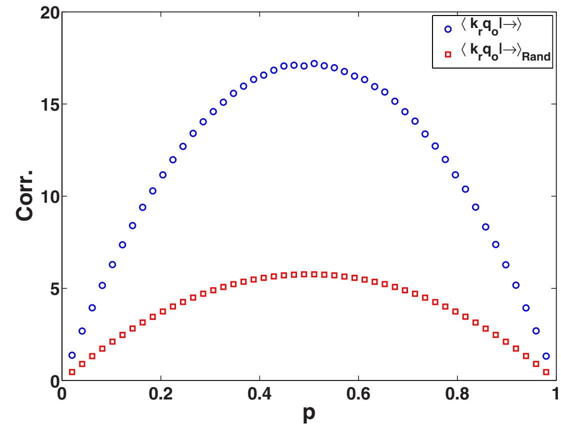

FIG. 2. (Color online) The correlations $\langle k'_r q'_o | \rightarrow \rangle$ resulting from the transformation process compared with the expected correlations of configuration model $\langle k'_r q'_o | \rightarrow \rangle_{Rand} = \frac{\langle k_r k_o \rangle \langle k_i k_o \rangle}{\langle k_o \rangle^2}$. In this case, the transformation process clearly amplifies observed degree correlations.

A very important assumption of this analysis is that the initial network does not mutate/evolve in any other way during the transformation process. However, the reality is that many complex networks evolve during the course of time and are highly nonequilibrium systems [4]. There is a myriad of different rules one can think of in order to simulate some aspect of network growth and to study the influence of the addition of the bidirectional edges in all these cases is impossible. We decided to study how the addition of reciprocal edges changes some very well-studied growth process. An obvious candidate for the study was the preferential attachment growth with the addition of reciprocal edges during the evolution process.

IV. GROWTH MODEL

In this model, the crucial idea is that at the formation of directed edge between new and old vertices there is a transfer of information about that event from the pointing vertex to the pointed vertex. In that case, the old vertex can return the newly formed edge to the new vertex, thus, forming a bidirectional edge. In the model, this process of information return will be modeled by the probability r that the old vertex points back to the new vertex.

The model can be described as a variant of the directed network growth by means of preferential attachment and formation of reciprocal edges. More precisely, in every time step t there are t vertices labeled from $0,\ldots,t-1$ present in the network and a new vertex labeled with t attaches to the network with m outgoing edges. Each of those m edges is attached to some already present vertex s with probability proportional to the in degree of the old vertex, i.e., $P(t \rightarrow s) \simeq \frac{k_i(s)}{\langle k_i \rangle t}$. If the network is grown only using this rule, the model is a variant of the BA model for the growth of directed network with the attractiveness parameter $A=0$. The additional rule is that each of m new edges, with probability r, can receive a reciprocal edge from the old vertex. With this additional rule, the model is completely described. It is useful to note that based on previous work [19,20], one can expect that for the value of parameter $r \approx 0$ the network will have an in-degree distribution with exponent $\gamma \approx -2$ and, for value of parameter $r=1$, the network will have an in-degree distribution with exponent $\gamma = -3$.

Although we will later calculate the analytical expression for the joint degree probability distribution for general parameter m, we will first present the solution for $m=1$ because it is easier to write it in a closed form. We will use the master equation to calculate joint degree probability distribution. Let $p(k_i,k_o,s,t)$ present the probability that the vertex introduced to the network at the moment s, at time t possesses in degree k_i and out degree k_o. In this treatment, for simplicity we will not use bidirectional degree because it would unnecessarily complicate the calculation. As the initial condition at time $t=1$, we choose the network of two reciprocally connected vertices $s=0$ i $s=1$,

$$p(k_i,k_o,0,1) = \delta_{k_i,1}\delta_{k_o,1},$$

$$p(k_i,k_o,1,1) = \delta_{k_i,1}\delta_{k_o,1}. \tag{29}$$

The probability $p(k_i,k_o,s,t)$ for $k_o \geq 1$, $k_i > 0$, and $s < t$ is

$$\begin{aligned} p(k_i,k_o,s,t) = {} & \frac{k_i-1}{L_{in}(t)}(1-r)p(k_i-1,k_o,s,t-1) \\ & + \frac{k_i-1}{L_{in}(t)} r p(k_i-1,k_o-1,s,t-1) \\ & + \left(1-\frac{k_i}{L_{in}(t)}\right)p(k_i,k_o,s,t-1). \end{aligned} \tag{30}$$

The function $L_{in}(t)$ is a random variable, which is equal to the sum of all degrees present in the network at time t,

$$L_{in}(t) = \sum_{s=0}^{t-1} k_i(s) = \sum_{k_i} k_i \sum_{s,k_o} p(k_i,k_o,s,t-1). \tag{31}$$

The following approximation for the $L_{in}(t)$ is very reasonable for a very big network,

$$L_{in}(t) \simeq \langle L_{in}(t)\rangle = (1+r)t, \tag{32}$$

i.e., we assume that the random variable $L_{in}(t)$ is well described by its expected value.

The equation for the vertex t, which is just attaching to the network at the time t is

$$p(k_i,k_o,t,t) = r\delta_{k_i,1}\delta_{k_o,1} + (1-r)\delta_{k_i,0}\delta_{k_o,1}. \tag{33}$$

Probability that the vertex s does not have any ingoing edge is

$$p(0,k_o,s,t) = (1-r)\delta_{k_o,1}. \tag{34}$$

We sum the obtained joint probabilities $p(k_i,k_o,s,t)$ that the vertex s at time t have degrees k_i and k_o, over all present vertices s to get the probability $P(k_i,k_o,t)$, that the randomly chosen vertex at time t has degrees k_i and k_o, i.e., $P(k_i,k_o,t) = \Sigma_{s=0}^{t} p(k_i,k_o,s,t)/(t+1)$. We also assume

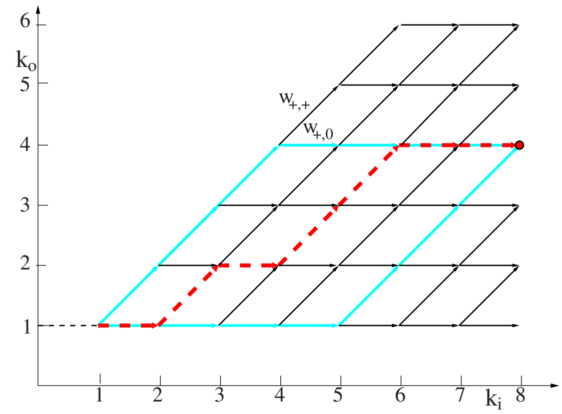

FIG. 3. (Color online) The event-set lattice used to calculate the stable distribution. The borders of the lattice segment in which paths contribute to the probability are designated with cyan color. Dashed red designates one of the possible paths and arrows represent the possible directions of paths.

that the distribution will be stable for large t, i.e., $P(k_i,k_o,t) \xrightarrow{t\to\infty} P(k_i,k_o)$. The described procedure results with the equations

$$P(0,k_o) = (1-r)\delta_{k_o,1}, \tag{35}$$

$$P(k_i \geq 1,1) = \frac{r(1+r)\delta_{k_i,1} + (k_i-1)(1-r)P(k_i-1,1)}{1+r+k_i}, \tag{36}$$

and

$$\begin{aligned} P(k_i \geq 1,k_o > 1) = {} & \frac{k_i-1}{1+r+k_i}[(1-r)P(k_i-1,k_o) \\ & + rP(k_i-1,k_o-1)]. \end{aligned} \tag{37}$$

Equation (36) shows that $P(1,1) = r(1+r)/(2+r)$. The simplest way to solve this set of equations is to sum contributions of all possible paths for probability distribution $P(k_i,k_o)$. Equations (36) and (37) can be easily represented as the walk on the event lattice shown in Fig. 3. The nodes of this lattice represent all the possible *events* (degree combinations) of randomly choosing a vertex from the ensemble of networks generated by the studied process. Every movement to the right from the site $\{k_i-1,k_o\}$ to the site $\{k_i,k_o\}$ is multiplying the probability distribution attached to the site with the factor $w_{+,0} = (k_i-1)(1-r)/(1+r+k_i)$, while every diagonal movement from the site $\{k_i-1,k_o-1\}$ to the site $\{k_i,k_o\}$ represents multiplying the probability distribution with the factor $w_{+,+} = (k_i-1)r/(1+r+k_i)$. The value of the joint degree probability distribution $P(k_i,k_o)$ is therefore equal to the sum of the contributions of all possible paths from site $\{1,1\}$ to the site $\{k_i,k_o\}$. Every path has k_i-k_o movements to the right and k_o-1 diagonal movements and every one of them has the same contribution

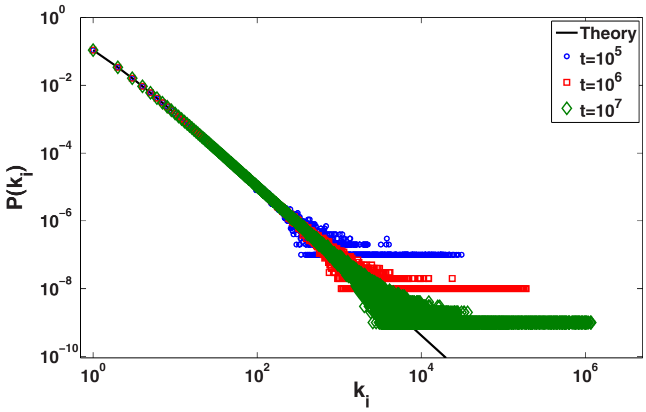

FIG. 4. (Color online) In this figure, it can be seen that the agreement between distributions of in degrees of the process simulated over 100 different realizations (markers) and the analytical solution (full line). The parameter r has the value of 0.2.

$$r^{k_o-1}(1-r)^{k_i-k_o}P(1,1)\prod_{n=2}^{k_i}\frac{n-1}{1+r+n}. \tag{38}$$

The number of distinct paths is $\binom{k_i-1}{k_o-1}$, which is equal to the number of combinations of factors $w_{+,0}$ and $w_{+,+}$. The general expression for the joint degree distribution is therefore equal to

$$P(k_i,k_o)=\Theta(k_i-k_o)\binom{k_i-1}{k_o-1}r^{k_o-1}(1-r)^{k_i-k_o}\times\frac{r(1+r)}{2+r}\frac{(k_i-1)!}{(r+3)_{k_i-1}}, \tag{39}$$

where denominator in the last factor $(r+3)_{k_i-1}$ represents the Pochammer symbol, defined with relation $(x)_n=x(x+1)\ldots(x+n-1)$. The nice property of this solution is that the correlations between degrees of one vertex are exactly computed and easily checked. In the limit of the big in degree, using the representation of the Pochammer symbol by means of Gamma functions $(x)_n=\frac{\Gamma(x+n)}{\Gamma(x)}$, it is easy to show that the asymptotic behavior of the last factor is $\lim_{k_i\to\infty}\frac{(k_i-1)!}{(r+3)_{k_i-1}}=\Gamma(r+3)k_i^{-(2+r)}$. By simple summation over the possible values of k_o, we can see that the marginal distribution $P(k_i)$ also has a power-law behavior with the same exponent.

To check analytical solutions, we have performed a series of simulations for different values of parameter r and different network sizes. For all monitored parameters and sizes of network, we found a nice agreement between analytical solution and simulations (Fig. 4).

V. GROWTH MODEL FOR A GENERAL PARAMETER *m*

It is possible to calculate the joint degree distribution of the model for the general parameter m. We again use the master equation,

$$p(k_i,k_o,s,t)=\sum_{l=0}^{m}\binom{m}{l}\left(\frac{k_i-l}{L_{in}(t)}\right)^{l}\left(1-\frac{k_i-l}{L_{in}(t)}\right)^{m-l}\Theta(k_i-l)\times\sum_{n=0}^{l}\binom{l}{n}r^n(1-r)^{l-n}\times p(k_i-l,k_o-n,s,t-1)\Theta(k_o-m-n), \tag{40}$$

where $\Theta(x)$ represents usual Heaviside Theta function with convention $\Theta(0)=1$. Indices m, l, and n combine all the possible combinations of the number of outgoing edges formed on the new vertex, the number of new edges attached to the old vertex, and the number of formed reciprocal edges from which it is possible to create vertex with given degrees. To ease the calculation, we allowed the formation of multiple edges between two vertices, which in the thermodynamical limit does not influence the exact solution. The boundary condition of this set of equations is

$$p(k_i,k_o,t,t)=\sum_{n=0}^{m}\binom{m}{n}r^n(1-r)^{m-n}\delta_{k_i,n}\delta_{k_o,m}. \tag{41}$$

We again sum over all vertices and approximate the function $L_{in}(t)$ with its expected value $L_{in}(t)\simeq\langle L_{in}(t)\rangle=(1+r)mt$. Using Eq. (41) and assuming a stable degree distribution in the thermodynamical limit, the following equation for the joint degrees distribution is obtained:

$$\frac{1+r+k_i}{1+r}P(k_i,k_o)=\sum_{n=0}^{m}\binom{m}{n}r^n(1-r)^{m-n}\delta_{k_i,n}\delta_{k_o,m}+\frac{k_i-1}{1+r}\Theta(k_i-1)\Theta(k_o-m-1)rP(k_i-1,k_o-1)+\frac{k_i-1}{1+r}\Theta(k_i-1)\Theta(k_o-m)(1-r)P(k_i-1,k_o). \tag{42}$$

This equation can be solved in a manner similar to Eqs. (35)–(37). In Fig. 5 we present the event-set lattice with the contributions of the paths to the joint probability attached to every site. The contribution of every path is again identical for every lattice bond as soon as the path detaches from the line $k_o=m$. The total contribution of the paths differs only in the number of steps made on line $k_o=m$ for the $1<k_i\leq m$. For $1<k_i\leq m$ and $k_o=m$, the joint degree probability is

$$P(k_i,m)=\frac{1+r}{1+r+k_i}\sum_{l=1}^{k_i-1}a_l(1-r)^{k_i-l}\prod_{j=l}^{k_i-1}\frac{j}{1+r+j}+\frac{1+r}{1+r+k_i}a_{k_i}, \tag{43}$$

and a_l represents the probability of binomial distribution $a_l=\binom{m}{l}r^l(1-r)^{m-l}$. The equation for the case $k_i>m$ and $k_o=m$ is

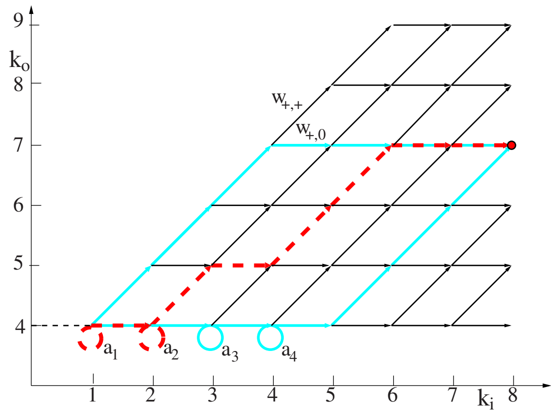

FIG. 5. (Color online) The lattice used to calculate the stable distribution of the model for the parameter $m=4$. Cyan color designates the borders of the lattice area within which the paths contribute to the joint probability $P(k_i=7,k_o=8)$, while the dashed red path represents one of the possible contributing paths. Arrows represent the allowed directions of movement on the lattice, while the loops on the sites represent the additional coefficients a_l contributing to the joint degree probability on these sites.

$$P(k_i,m) = (1-r)^{k_i-m} P(m,m) \prod_{j=m+1}^{k_i} \frac{j-1}{1+r+j}. \tag{44}$$

The contribution of paths which separated from the line $k_o = mi1 < k_i \leq m$ at the site $(k_i=k_i', k_o=m)$ is

$$\begin{aligned} P(k_i,k_o)_{\pi(k_i',m)} &= P(k_i',m)\binom{k_i+k_o-m-k_i'-2}{k_i-k_i'-1} \\ &\times (1-r)^{k_o-m} r^{k_i-k_i'-k_o+m} \frac{(k_i-1)!}{(k_i'-1)!}\frac{(r+3)_{k_i'-1}}{(r+3)_{k_i-1}}, \end{aligned} \tag{45}$$

where $\pi(k_i',m)$ represents the sum over the possible paths after detachment from the line $k_o=m$. The whole solution is now easy to write using Eqs. (43)–(45), but it is complicated and not very informative. Nevertheless, it is interesting to monitor the behavior of Eq. (43). It can be verified that for $P(k_i<m,k_o)$, the joint degree distribution can increase depending on the initial parameters of the model. On the other hand, in the limit $k_i \gg m$ we expect the fall of the joint probability distribution. This implies that for certain range of parameters, the distribution has a nontrivial mode. Such a behavior can be easily checked if Eq. (42) is summed over all out degrees $k_o \in |m,\infty\rangle$ to obtain the marginal distribution of in degree. In the range $k_i \in [1,m]$, the solution is

$$P(k_i) = \frac{1+r}{1+r+k_i}\left(a_{k_i} + \sum_{l=1}^{k_i-1} a_l \frac{(k_i-1)!}{(l-1)!}\prod_{j=l}^{k_i-1}\frac{1}{1+r+j}\right). \tag{46}$$

The marginal in-degree distribution obtained analytically coincides with the simulations rather well as shown on Fig. 6. The modal character of the in-degree distribution is easily observed in this equation. The dependence of mode on the parameters r and m is shown in Figs. 7 and 8.

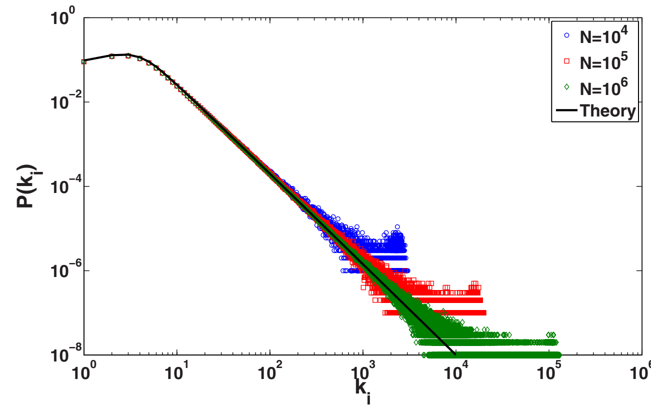

FIG. 6. (Color online) The marginal distribution of the in degree calculated from theory (full line) and simulations (markers), for three different sizes of final networks. The averaging was performed over 100 simulation realizations and parameters used are $r=0.15$ and $m=18$. There is an excellent agreement between theory and simulations.

The other important property of the distribution is its power-law behavior in the tail. As can be seen in Fig. 8, the exponent of the tail does not depend on the parameter m. Power-law behavior of the tail is governed only by the parameter r as shown in Fig. 7. Indeed in the continuum approximation, valid for $k_i \gg 1$, the equation for the in-degree marginal distribution is

$$(1+r)P(k_i) \sim -\frac{d[k_i P(k_i)]}{dk_i}. \tag{47}$$

The solution of this equation is

$$P(k_i) \sim k_i^{-(2+r)}, \tag{48}$$

and the dependence of the value of the power-law exponent with respect to parameter r is very clear. The equation for the power-law exponent $\gamma=-2-r$ also confirms our claim that

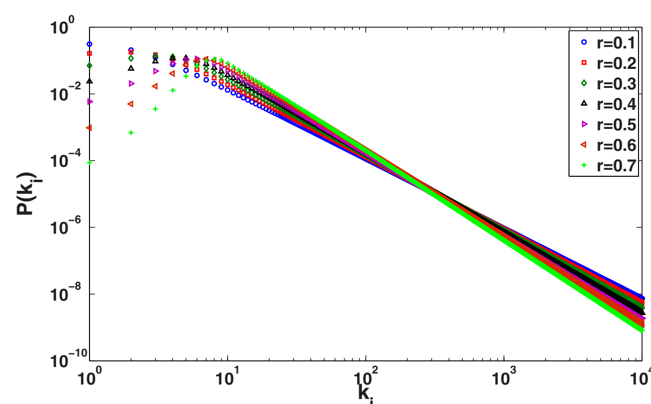

FIG. 7. (Color online) The marginal distribution of the in degree calculated from the theory for $m=10$ and different values of parameter r. The tail of distribution clearly depends on the parameter r and the power-law character of the tail is easily checked. The shape and the position of the mode also depend on the value of r.

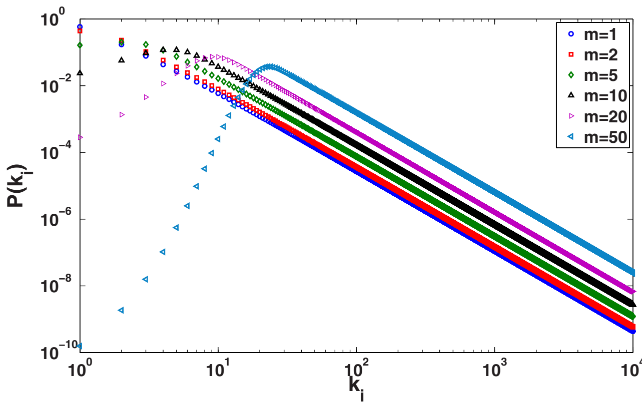

FIG. 8. (Color online) The marginal distribution of the in degree calculated from the theory for $r=0.4$ and different values of parameter m. The existence and the position of the mode strongly depend on the value of parameter m. The tail of distribution is independent of the parameter m and the power-law character of the tail is easily checked.

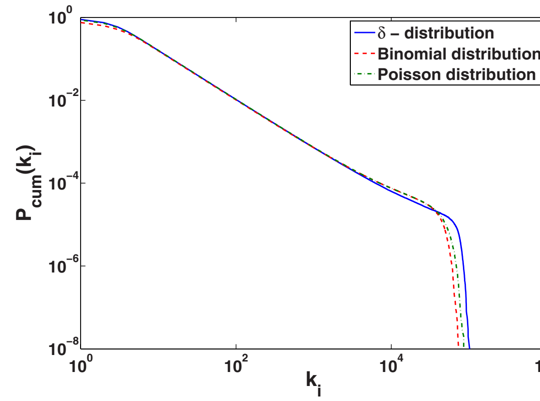

FIG. 9. (Color online) In this figure, we present cumulative in-degree distributions for the different choices of the initial out-degree distribution of the vertex introduce at time t. The simulations are presented for the parameters $m=10$, $r=0.2$, $N=10^6$, and 100 realizations. It is clear that different choices of out-degree distributions for the new vertices do not change the general behavior of the in-degree distribution. The broader distributions of initial out degree result with a lower maximal degree, which we explain by a stronger competition for the new vertices.

the model behavior interpolates between usual BA model with the exponent -3 and the directed BA model with the exponent close to -2. While agreement between our model for $r=1$ and the usual BA model for the undirected networks is trivial, the relation with the directed version of BA model is a bit more complicated. As can be seen in asymptotic form of our model for $m=1$, the marginal in-degree distribution vanishes for $r=0$ just as in case of directed BA model [20]. In their model, the asymptotic behavior is governed by the relation $P(k_i)\sim k_i^{-(2+a)}$, where a is a fraction of attractiveness A and m number of outgoing edges. For very small a, $P(k_i)\approx k_i^{-2}$, as is the case in our model, if we let r to be very small. Therefore, the asymptotic behaviors of both models are very similar in the case of $a\approx r\approx 0$. In the case $a=r=0$, resulting marginal in-degree distribution changes dramatically in both models. In this case, the resulting network is similar to star with the center of the star equivalent to the size of initial core.

It is important to mention that this analytical discussion is valid up to a certain point also for a little bit broader class of growth models. In the analytic treatment, the distribution of the outgoing edges of the new vertex t [Eq. (41)] at the time t is a delta function. We can expect that this reasoning can be applied for a more general class of distributions for the outgoing edges of the vertex t at time t with the assumption that the mean-field approximation is valid. In particular, we expect that this consideration will be valid for all unimodal discrete distributions with fast decaying tails. To test this assumption, we examined cases in which the out degree of the vertex t at time t is drawn from the binomial and Poisson distributions.

The Poisson distribution

$$P(k_o|m) = \frac{m^{k_o}e^{-m}}{k_o!} \tag{49}$$

is determined only by parameter m. For every monitored m of the original model, we made a new set of simulations with Poisson distribution with the same m. For binomial distribution,

$$P(k_o|m,Z) = \binom{Z}{k_o}\left(\frac{m}{Z}\right)^{k_o}\left(1-\frac{m}{Z}\right)^{Z-k_o}, \tag{50}$$

the case is a little bit more complicated because it is defined with two parameters: m the expected number of outgoing edges and Z maximal allowed out degree of the vertex t at time t. For a broad choice of values of parameter Z, the results were very similar to the ones expected from the original model as $N\gg Z$, as can be seen in Fig. 9.

VI. CONCLUSION

We have shown that reciprocal edges can significantly influence the degree correlations in complex networks. In the first part of the paper, we laid down a way to investigate the influence of randomly distributed bidirectional edges on the overall degree correlations and have shown how this hypothesis can be tested in the case of real networks. We also studied a simple model of the network growth, which conserves an expected fraction of reciprocal edges.

The analysis laid out in the first part of this paper focuses on the degree correlations represented as average product moments. This analysis has its positive and negative side. The average product moments clearly do not contain as much information as the average neighbor degree functions [4] and this is an obvious shortcoming of such a measure. On the other hand, it is exactly the reason why product moments can be very useful for case studies. In the case of very correlated networks, sometimes the frequency of degree statistics for the large degrees is so scarce that it effectively shrinks the available configuration space for the null models, which are trying to preserve correlations found in the network. This reduction in the available configuration space can sometimes be so huge that for connected pairs of vertices with large degrees any result different from the already observed in the network is almost impossible to realize. In this case, the product moments incorporate in themselves much larger number of viable network realizations so that the analysis of network with correlated null models is much better founded.

In the companion paper [17], we apply the theory presented in this paper to show that Wikipedia networks cannot be explained by the random distribution of bidirectional edges on the static network. In the same paper, we used the presented growth model to explain the in-degree distribution of the Wikipedia networks with very good results.

From the theoretical point of view, this model helps to understand possible mechanisms, which create modes in the degree distributions of different scale-free-directed networks. Furthermore, this model is a good candidate to explain other empirical directed networks, which combine power-law tails and nontrivial mode of the degree distribution, and the future work in this direction is clearly needed. It also represents one of the simplest growth models, which preserves some type of local correlations. It is our opinion that the understanding of interrelations between different types of correlations in complex networks heavily depends on such growth models. The validation of this claim is an important task for our future research.

ACKNOWLEDGMENTS

This work was financed by the Ministry of Education, Science and Sports of the Republic of Croatia under Contract No. 098-0352828-2863 and by the INFN, Italy. V.Z. is thankful for the support of G. Caldarelli and would like to thank A. Gabrielli for reading the manuscript.

[1] G. Caldarelli, R. Marchetti, and L. Pietronero, Europhys. Lett. **52**, 386 (2000).
[2] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).
[3] D. Garlaschelli, G. Caldarelli, and L. Pietronero, Nature (London) **423**, 165 (2003).
[4] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[5] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).
[6] M. E. J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).
[7] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[8] D. Garlaschelli and M. I. Loffredo, Phys. Rev. Lett. **93**, 268701 (2004).
[9] B. Bollobas, in *Graph Theory and Combinatorics*, Proceedings of the Cambridge Combinatorial Conference in Honor of Paul Erdős (Academic Press, New York, 1984), pp. 35–57.
[10] V. Colizza, A. Flammini, M. A. Serrano, and A. Vespignani, Nat. Phys. **2**, 110 (2006).
[11] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon, Science **303**, 1538 (2004).
[12] M. Boguñá and M. A. Serrano, Phys. Rev. E **72**, 016106 (2005).
[13] V. Zlatić, M. Božičević, H. Štefančić, and M. Domazet, Phys. Rev. E **74**, 016115 (2006).
[14] G. Zamora-López, V. Zlatić, Ch. Zhou, H. Štefančić, and J. Kurths, Phys. Rev. E **77**, 016106 (2008).
[15] A. Bekessy, V. Bekessy, and J. Komlos, Stud. Sci. Math. Hung. **7**, 343 (1972).
[16] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995).
[17] V. Zlatić and H. Štefančić, e-print arXiv:0902.3548.
[18] B. Bollobás, C. Borgs, J. Chayes, and O. Riordan, *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms* (SIAM, Philadelphia, 2003), pp. 132–139.
[19] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[20] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000).