J Stat Phys (2015) 161:965–985
DOI 10.1007/s10955-015-1360-4

The Configuration Model for Partially Directed Graphs

Kristoffer Spricer[1] · **Tom Britton**[1]

Received: 16 March 2015 / Accepted: 24 August 2015 / Published online: 7 September 2015

Abstract The configuration model was originally defined for undirected networks and has recently been extended to directed networks. Many empirical networks are however neither undirected nor completely directed, but instead usually partially directed meaning that certain edges are directed and others are undirected. In the paper we define a configuration model for such networks where vertices have in-, out-, and undirected degrees that may be dependent. We prove conditions under which the resulting degree distributions converge to the intended degree distributions. The new model is shown to better approximate several empirical networks compared to undirected and completely directed networks.

Keywords Configuration model · Partially directed · Semi-directed · Degree distribution · Asymptotic convergence

1 Introduction

Graphs appear in many current applications. In social sciences groups of people are often modeled by letting the vertices in the graph represent persons and edges represent the interactions or relationships between them. Edges can be directed or undirected, the latter indicating a reciprocal relationship between the vertices.

Usually the graphs created from such datasets are simplifications of the original dataset. One typical simplification is to allow only directed or only undirected edges. However, in real world graphs it is common to find a combination of directed and undirected edges. In [3] we find some examples of empirical graphs where the proportion of directed edges is in the range 0.26–0.85, the rest being undirected edges. Additional examples are shown in Table 1 where the proportion of directed edges has been calculated for some social networks that can be found in [9]. We expect such graphs to be better represented by *partially directed graphs*,

✉ Kristoffer Spricer
spricer@math.su.se

[1] Department of Mathematics, Stockholm University, SE-106 91 Stockholm, Sweden

Table 1 Proportion of directed edges for some data sets from [9], when viewed as partially directed graphs

Data set	# vertices	# edges	Proportion directed
soc-LiveJournal1	4, 847, 571	42, 851, 237	0.402
soc-Epinions1	75, 879	506, 585	0.996
soc-Pokec	1, 632, 803	22, 301, 964	0.627
soc-Slashdot0922	82, 168	504, 230	0.274
email-EuAll	265, 214	310, 006	0.851
wiki-Vote	7, 115	100, 762	0.971
wiki-Talk	2, 394, 385	4, 659, 565	0.922

We see that several of these graphs have a substantial proportion of undirected edges and of directed edges, such that neither type should be ignored

where we allow both directed and undirected edges. For instance, results in [2] show that that epidemic spread on such partially directed graphs is different than e.g. on undirected graphs.

The *configuration model* has been used extensively to model undirected networks [4,5]. It has also been been adapted to work for directed graphs [1]. In the configuration model the graph is constructed by first assigning a degree to each vertex of the graph and then connecting the edges uniformly at random. The degrees of the vertices of the graph are either given as a degree sequence or the degrees are drawn from some given degree distribution. Graphs created in this way will share some properties with real world graphs, but will be different in other aspects. E.g. the configuration model for directed networks will have a very low proportion of reciprocal edges, i.e. two parallel directed edges in opposite directions. This is an effect of connecting edges uniformly at random in this type of graph, resulting in a low probability of achieving a reciprocal connection between vertices. The low proportion of undirected edges in the resulting configuration model graph can be undesirable if we wish to use it as a *null reference* to compare with a real-world graph. While we wish to connect the edges uniformly at random, we may want to preserve the degree distribution, including any dependence between the indegrees, outdegrees and undirected degrees.

In this paper we consider a partially directed configuration model where we allow both directed and undirected edges. Any vertex in such a partially directed configuration model graph can have all three types of edges: *incoming*, *outgoing* and *undirected*. We select the degree of each vertex from a given joint, three dimensional degree distribution and we do *not* assume or require the in-, out- and undirected degrees to be independent. When connecting the stubs, the yet unconnected edges, outgoing stubs can only connect to incoming stubs and undirected stubs can only connect to undirected stubs. Once all possible connections are made we want the graph to be *simple* and thus do not allow unconnected stubs, self loops or parallel edges of any type. We make the graph simple by erasing unconnected stubs, self loops and parallel edges, and by converting parallel directed edges in opposite directions into undirected edges. Since this process modifies the degree of some of the vertices, it is not certain that the empirical degree distribution converges to the given degree distribution. However, in Sect. 2 we show that, with suitable restrictions on the first moments of the degree distribution, the empirical degree distribution asymptotically converges to the desired one.

Note that, by selecting a joint degree distribution in the proper way we can also create completely directed graphs or completely undirected graphs, with or without any dependence between the degrees. Thus the presented partially directed configuration model incorporates several of the already existing models.

In Sect. 2 we present definitions and state the main result of the paper. Detailed derivations and proofs have been postponed to Sect. 4. To illustrate how these graphs work, Sect. 3 is devoted to some simulations of partially directed graphs, showing results for small and for large n. The latter is to give an intuitive feeling for the asymptotic results and the former is to illustrate that significant deviations from these asymptotic results are possible for small n. A comparison with an empirical social network is also done. Conclusions and discussion can be found in Sect. 5.

2 Definitions and Results

In this section we define the configuration model for partially directed graphs. We define the terminology used, how the graph is created from a degree distribution, how the graph is made simple and finally show, with suitable restrictions on the first moments of the degree distribution, that the degree distribution of the partially directed configuration model graph asymptotically converges to the desired distribution. Proofs are left for Sect. 4.

2.1 Terminology

A graph consists of vertices and edges. The size of the graph, the number of vertices, is denoted n. Here we will specifically study the case when $n \to \infty$. We work with graphs that are *partially directed*, meaning that any vertex can have incoming edges, outgoing edges and undirected edges. We distinguish between edges and *stubs*. By stubs we mean yet unconnected half-edges of a vertex. Corresponding to directed edges we have in-stubs and out-stubs, and corresponding to undirected edges we have undirected stubs. The number of stubs of the different types is the degree of a vertex and will be denoted $\underline{d} = (d^{\leftarrow}, d^{\rightarrow}, d^{\leftrightarrow})$, where the individual terms represent the indegree, outdegree and undirected degree, respectively. When the degree of the vertex is a random quantity, it is denoted $\underline{D} = (D^{\leftarrow}, D^{\rightarrow}, D^{\leftrightarrow})$.

A degree sequence that is non random is denoted $\mathbf{d} = \{\underline{d}_r\} = \{(d_r^{\leftarrow}, d_r^{\rightarrow}, d_r^{\leftrightarrow})\}$, $r = 1, \ldots, n$, where n is the number of vertices in the graph. When these degree sequences are random vectors they are denoted $\mathbf{D} = \{\underline{D}_r\} = \{(D_r^{\leftarrow}, D_r^{\rightarrow}, D_r^{\leftrightarrow})\}$.

Degrees can be assigned to the vertices from some given joint degree distribution with distribution function F for which the probability of a specific combination of indegree, outdegree and undirected degree is called $p_{\underline{d}} = p_{ijk} = P(\underline{D}{=}(i, j, k))$. We will also use the marginal distributions. We have $p_i^{\leftarrow} = p_{i..} = \sum_{jk} p_{ijk}$ for the incoming edges, $p_j^{\rightarrow} = p_{i.k} = \sum_{ik} p_{ijk}$ for the outgoing edges and $p_k^{\leftrightarrow} = p_{..k} = \sum_{ij} p_{ijk}$ for the undirected edges. The corresponding random variables, i.e. the number of edges of each type, will be denoted $D^{\leftarrow}$, $D^{\rightarrow}$ and $D^{\leftrightarrow}$.

Other quantities of interest are the moments of the distribution. Here we will consider the first moments $\mu^{\leftarrow} = \mathrm{E}[D^{\leftarrow}] = \sum i p_i^{\leftarrow}$, $\mu^{\rightarrow} = \mathrm{E}[D^{\rightarrow}] = \sum j p_j^{\rightarrow}$ and $\mu^{\leftrightarrow} = \mathrm{E}[D^{\leftrightarrow}] = \sum k p_k^{\leftrightarrow}$.

A graph is *simple* if there are no unconnected stubs, no self-loops and no parallel edges.

For a finite graph of size n we also want to count the number of vertices with a certain degree $\underline{d}$. We call this quantity $N_{\underline{d}}^{(n)}$. Dividing by n we can calculate $N_{\underline{d}}^{(n)}/n$, the proportion of vertices that have degree $\underline{d}$. Whenever the graph is created by some random process, we can also consider the expectation of this random quantity $p_{\underline{d}}^{(n)} := \mathrm{E}\left[N_{\underline{d}}^{(n)}/n\right]$, which defines the distribution function $F^{(n)}$.

2.2 Defining the Model

We define the partially directed configuration model as follows:

(1) We start with a graph with n vertices, but without any edges or stubs.
(2) For each vertex, we independently draw a degree $\underline{D}_r$ from F at random.
(3) We connect undirected stubs with other undirected stubs. We do this by picking two undirected stubs uniformly at random and connecting them. We repeat this with the remaining unconnected undirected stubs until there is at most one undirected stub left, which happens if the number of undirected stubs is odd.
(4) We connect directed incoming stubs with directed outgoing stubs. We do this by picking one directed incoming stub and one directed outgoing stub, both independently and uniformly at random and then connecting them. We repeat this with the remaining unconnected directed stubs until we are out of incoming stubs or outgoing stubs (or both). Unless, in the given degree distribution, the number of in-stubs is equal to the number of out-stubs for every degree that has a probability that is not zero, the probability that the number of in-stubs is equal to the number of out-stubs in the graph will go to zero as the size of the graph goes to infinity. Since the typical case for a partially directed graph is that in-degrees are different from out-degrees, there will usually be a large number of unconnected directed stubs left over, after making all possible connections between directed stubs. See also Table 3 for more details on this.
(5) We want the graph to be simple, but the connection process may have left some stubs unconnected and may also have created self-loops and parallel edges. We make the graph simple by erasing some stubs and edges. We define the procedure in such a way that the connectivity of the graph is maintained:
    (a) Erase all unconnected stubs. There can be at most one unconnected undirected stub, while there may be a larger number of unconnected directed stubs as discussed above.
    (b) Erase all self-loops, both directed and undirected.
    (c) When there are parallel *identical* edges, erase all except one of them.
    (d) Erase all directed edges that are parallel to an undirected edge.
    (e) Erase each pair of reciprocal directed edges and add a single undirected edge instead. While this step decreases the number of directed edges, it also increases the number of undirected edges.

From the above description we see that there are two non-deterministic steps that affect the degrees of the vertices in the creation of the simple partially directed graph:

(1) Assigning degrees from the distribution F.
(2) Connecting the stubs uniformly at random. While this does not, in itself, modify the degrees of the vertices, it affects which stubs and edges that will be erased when making the graph simple.

This process results in a finite simple graph for which the degree distribution $F^{(n)}$, that was defined above, typically will *not* be identical to F since we may have erased edges and stubs. However, we later show that, with suitable restrictions on the distribution F, the distribution $F^{(n)}$, asymptotically approaches F.

2.3 Asymptotic Convergence of the Degree Distribution

The results in this section are inspired by, and to some degree follow [6]. The theorem establishes the asymptotic convergence of the degree distribution. We remind the reader that

F is the given degree distribution and that it is defined by $p_{\underline{d}}$. $F^{(n)}$, the resulting degree distribution for the simple graph of size n, is defined by $p_{\underline{d}}^{(n)} := \mathrm{E}\left[N_{\underline{d}}^{(n)}/n\right]$.

Theorem 1 *If F has finite mean for each component, so $\mu^{\leftarrow} < \infty$, $\mu^{\rightarrow} < \infty$, and $\mu^{\leftrightarrow} < \infty$, and also $\mu^{\leftarrow} = \mu^{\rightarrow}$ then, as $n \to \infty$*

(a) *$F^{(n)} \to F$ or, equivalently, $p_{\underline{d}}^{(n)} \to p_{\underline{d}}$ for all $\underline{d}$,*

(b) *$N_{\underline{d}}^{(n)}/n \xrightarrow{P} p_{\underline{d}}$, that is, the empirical distribution converges in probability to F.*

The proof, which is postponed to Sect. 4, follows the same line of reasoning as in [6], but with modifications to take into account the complications introduced by allowing both directed and undirected edges in the graph.

3 Examples of Partially Directed Graphs

Although Theorem 1 establishes the asymptotic convergence of the degree distribution, it remains to see how well this holds for finite graphs. In this section we investigate this by looking at a scale-free distribution, at a Poisson degree distribution and at an empirical network. In this paper, by scale-free distribution we mean a distribution with a power-law tail. Since we are working with a joint degree distribution, in addition to the distribution for each of the three stub types we also need to consider the possible dependence between the different types. Table 2 gives an overview of how the data for the plots were created.

We note that with three types of stubs many different types of correlations between the three degrees are possible for the scale-free and Poisson random graphs. In this paper we explore two such possibilities. To contrast the case where all three degrees are completely independent we show the case where all degrees of a node are identical, i.e. maximally dependent. When selecting the parameters for the distributions we can also choose in what way we want the distributions to match the degree distribution of the empirical graph. Both chosen distributions only have a single parameter and so we cannot match all properties of the empirical graph by adjusting this parameter. For the scale-free graph we focus on the slope of the distribution, while for the Poisson graph we focus on the mean degree. The choice of a scale-free distribution is motivated by the empirically observed phenomenon of degree distributions often having heavy tails of the power-law type found in scale-free distributions. Here we choose to model this heavy tail by using an approximation to the Zeta distribution, which is one variant of a scale-free distribution. In a more advanced model, degree distributions with more parameters could also be introduced to allow for making them more or less similar to the degree distribution of the empirical network.

Since Theorem 1 focuses on showing convergence to the correct degree distribution, studying the total variation distance, $d_{\mathrm{TV}}^{(n)}$ (defined in Sect. 3.1), is of interest (see e.g. [10]). We also study the number of erased edges as a function of the graph size. Finally, we study the size of the strongly connected giant component and the distribution of small components for a few different graphs based on the empirical data from LiveJournal. The dataset LiveJournal [9] is a directed graph created from the declaration of friends in a social internet community. The original graph contains self loops, but these have been removed in this analysis. The simple graph has a proportion of directed edges of about 0.4, so this is a good example of a graph where both directed and undirected edges play an important role. When sampling from this distribution to create the configuration model graph, the degrees of vertices from the original (partially directed) graph were drawn independently and uniformly at random,

Table 2 The table explains how the graphs were created

Degree distribution	Method	Independent	Dependent
Empirical	Empirical degree data from the dataset *soc-LiveJournal1* [9] was used. This is data from an on-line social site. Some characteristics can be found in Table 1. The mean degrees for in-degrees, out-degrees and undirected degrees are 3.6, 3.6 nd 10.6, respectively (not shown) When viewed as a directed graph and counting all stubs this gives a total mean degree of approximately 28.3. Here we count the undirected edges as two edges since it consists of an incoming and an outgoing edge, when viewed as an edge in a directed graph. Both the directed and the undirected edges have degree distributions that are *approximately* scale-free in the tail, with $\gamma_{\text{directed}} \approx 2.5$ and $\gamma_{\text{undirected}} \approx 3.5$ (not shown)	Each stub type is treated individually and independent samples are drawn, with replacement, for each vertex and each stub type	Independent samples of complete vertices are drawn, with replacement, from the pool of empirical vertices
Scale-free	The selected distribution function is $$F(k) = 1 - \frac{(k+d)^{-(\gamma-1)}}{d^{-(\gamma-1)}},$$ with $$d = (\zeta(\gamma) * (\gamma - 1))^{-\frac{1}{\gamma-1}},$$ where $\zeta(\gamma)$ is the Riemann zeta function. The tail of this distribution is asymptotically $p_k \propto k^{-\gamma}$. This specific distribution function was selected because of its scale-free property (it is an approximation to the Zeta distribution), while still being easy to simulate from using a discrete variant of the inverse transformation method [[11], see Sect. 11.2.1 and also Example 11.7]). For all simulations $\gamma = 2.5$, which is the coefficient for the directed edges in the empirical graph. This value gives finite expectation (approximately 2.7), but infinite variance. This is consistent with the assumptions in Theorem 1	For each vertex and each stub type an independent sample from the assigned distribution was drawn	For each vertex an independent sample from the assigned distribution was drawn and the same degree was assigned to all stubs for the vertex
Poisson	Degrees drawn from Poisson distribution with parameter 7, thus having mean degree 7. When treated as a directed graph and counting all stubs the total mean degree is 28, close to the value 28.3 for the empirical graph above	See above	See above

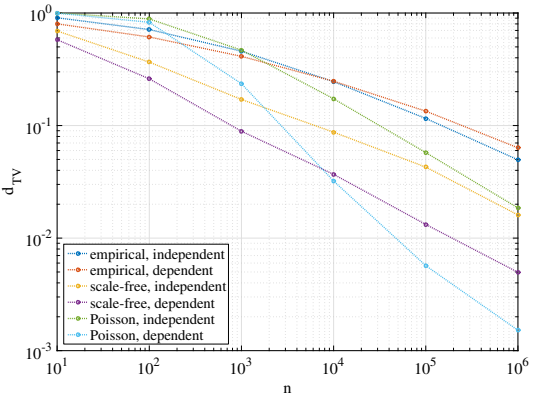

Fig. 1 The total variation distance versus graph size for three different degree distributions, with independent or dependent in-, out- and undirected degrees for the stubs. Each data point shows the average of 100 simulations. All *curves* decrease towards zero

with replacement. Thus the frequencies of the degrees found in the graph were used as the given distribution F and this distribution function is then compared with the distribution $F^{(n)}$ created by sampling from F, connecting the edges and making the graph simple.

3.1 Total Variation Distance

Theorem 1 states that $N_{\underline{d}}^{(n)}/n \xrightarrow{P} p_{\underline{d}}$ and thus we define the following version of the total variation distance:

$$d_{\mathrm{TV}}^{(n)} = \frac{1}{2}\sum_{\underline{d}} |p_{\underline{d}} - N_{\underline{d}}^{(n)}/n|, \tag{1}$$

where the $1/2$ is introduced so that $d_{\mathrm{TV}}^{(n)}$ can only take on values in the range $[0, 1]$. As $n \to \infty$ we expect to see that the total variation distance tends towards zero. When we generate the graphs according to the configuration model we replace $N_{\underline{d}}^{(n)}$ with the corresponding empirical sample from one realization of a random graph. We can then repeat this process with more samples of random graphs and plot this. The result is shown in Fig. 1, where we have also taken the average of the empirical total variation distance for 100 random graph samples.

In Fig. 1 we see that the total variation distance tends to decrease towards zero. The fastest decrease is for the Poisson graph, and the reason is that this distribution has a light tail when compared with the scale-free distribution. A closer look at the empirical graph reveals that the distributions for the directed and the undirected edges look much like a scale-free distribution. The in- and the out-degree have $\gamma \approx 2.5$ and the undirected degree has $\gamma \approx 3.5$ in the tail (not shown). Thus the tail for the empirical distribution is heavier than for the Poisson distribution and so we can expect a slower convergence for the empirical graph, at least initially. However, we have to remember that the empirical distribution is in fact finite, having a maximum degree. Thus, if we only consider very high degrees and large graphs then the Poisson graphs will exhibit higher maximum degrees than graphs based on the empirical degree distribution. For graphs up to 10^6 vertices this effect cannot yet be seen.

The slowest convergence can be observed for the scale-free distribution with $\gamma = 2.5$. For this distribution the variance is not finite and this reflects in the convergence being slower than for the other two distributions. Even slower convergence has been observed (not shown) for values of γ even closer to 2, e.g. try $\gamma = 2.1$. This is not surprising as the distribution then becomes more heavy-tailed. As γ becomes smaller, the number of erased edges increases

as an effect of an increased number of self loops and parallel edges. As an example we can consider the undirected edges only with γ approaching 2 from above. As this happens the probability that a single vertex dominates the total number of undirected edges in the graph gradually increases to become non negligible as γ reaches 2. This will result in a high probability of self loops for this vertex and also for parallel edges to other vertices. As these edges are erased during the simplification process, the degree distribution becomes less equal to the given degree distribution and the total variation distance shows slower convergence. If we continue even further, to $\gamma \leq 2$ the conditions used in the proof of Theorem 1 no longer hold, since the expectations are no longer finite, and thus we should not expect the total variation distance to converge to zero for these values of γ.

From the figure we also see that the *dependent* curve for the Poisson distribution is clearly lower than the *independent* curve. One explanation for this is that when the degrees for in-stubs and the out-stubs are identical for each vertex, as in the dependent graph (as defined in Table 2), the total number of in-stubs will be equal to the total number of out-stubs and thus no directed stubs will be erased for this reason. There may still be self-loops and parallel edges, but for the Poisson graph these are few compared to the number of stubs erased in the independent graph (as defined in Table 2) where there is a mismatch between the number of in-stubs and the number of out-stubs. For the empirical graph and for the scale-free graph the same phenomenon cannot be observed. One explanation to this is that the scale-free independent model is not necessarily dominated by the deletion of leftover directed edges. Instead the number of self-loops and parallel edges are of the same order of magnitude as the leftover directed edges (see Fig. 2). Thus the difference between the dependent and the independent curves for the total variation distance is much smaller for the scale-free graph and for the empirical graph.

Another answer to why the empirical graph does not show a big difference between the dependent and the independent curve can be that the dependent version of the empirical graph does not have the same type of complete dependence as for the scale-free or the Poisson graph. In the empirical dependent graph, degrees are assigned by sampling the degrees of vertices from the original empirical graph, and thus the number of in-stubs will in general not equal the number of out-stubs. Looking at Fig. 2 we see that the number of directed unconnected edges is almost the same for the independent version as for the dependent version of the empirical graph. Looking instead at the same plot for the Poisson graph we note that the deletion of directed unconnected stubs dominates the independent version of the graph, while there are no such erased stubs in the dependent version of the graph.

3.2 The Average Number of Erased Edges Per Vertex

The number of erased edges will depend on the degree distribution, on the graph size and will also be different each time a graph is created according to the configuration model. In Fig. 2 the average number of erased edges per vertex were plotted. Each point corresponds to the average of 100 simulations of random graphs according to the partially directed configuration model. The erased edges were classified as to the reason why they were erased as defined in the rules in Sect. 2.2.

For all plots, the graphs indicate that the average number of erased stubs or edges per vertex decreases with the size of the graph. Thus also the risk of any vertex having its degree affected by the deletion of a stub or an edge goes down and this indicates that the degree distribution $F^{(n)}$ converges to F asymptotically. The scale-free distribution is more difficult since for $\gamma \leq 2$ neither the variance nor the expectation exist. Here we have selected $\gamma = 2.5$ for the scale-free graph. This value gives finite expectation, but infinite variance.

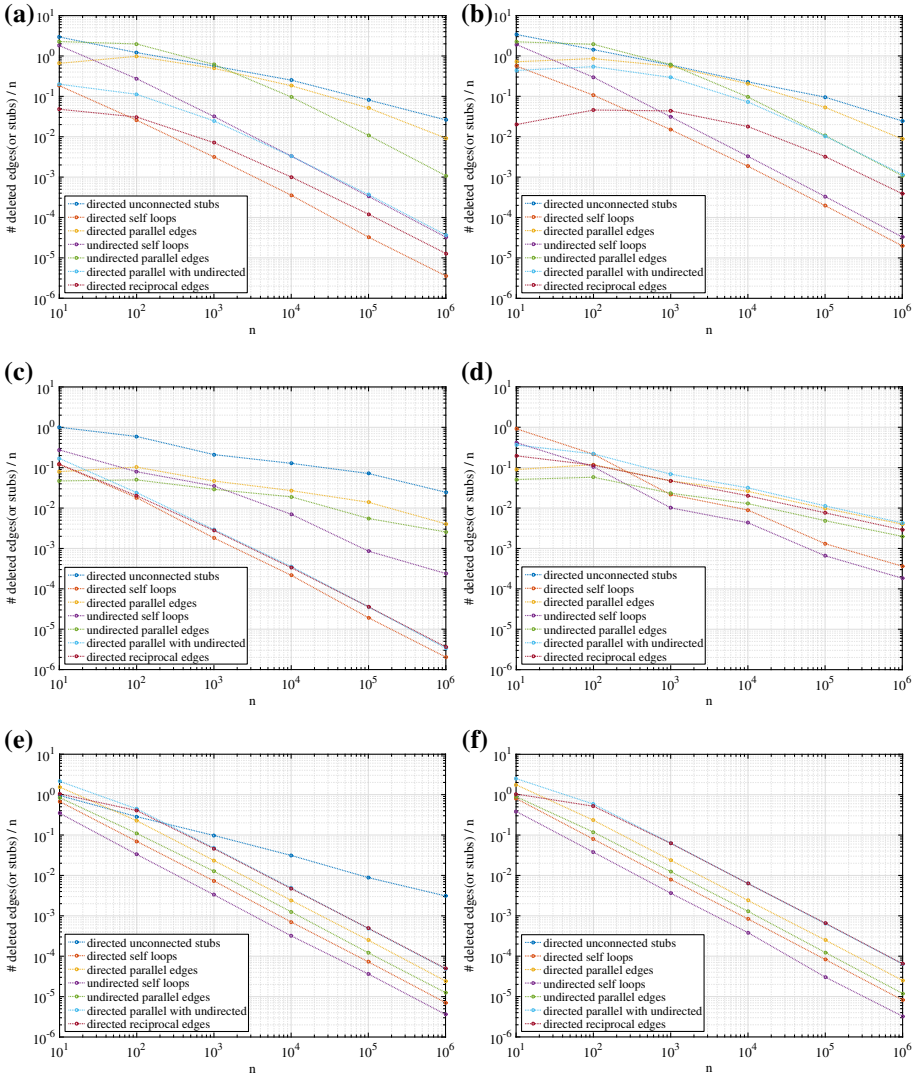

Fig. 2 Number of erased edges divided with the number of vertices for the scale-free configuration model with parameter $\gamma = 2.5$, for the Po(7) model and for the empirical configuration model. Each data point shows the average of 100 simulations. **a** Empirical, independent in-, out- and undirected degrees; **b** Empirical, dependent in-, out- and undirected degrees; **c** Scale-free, independent in-, out- and undirected degrees; **d** Scale-free, dependent in-, out- and undirected degrees; **e** Poisson, independent in-, out- and undirected degrees; **f** Poisson, dependent in-, out- and undirected degrees

Asymptotic results on the distribution of the number of self loops and parallel edges have been obtained for both undirected and directed graphs when both the expectation and the variance of the degree distribution are finite. For undirected graphs see [8, Sect. 7] and [12, Proposition 7.12], and for directed graphs see [1, Proposition 4.2]. In all of these cases the number of erased edges is asymptotically Poisson distributed, with parameters that depend on the first moments, the second moments and the covariances of the degree distribution.

For the partially directed graph the process of deleting edges also affects reciprocal directed edges and directed edges that are parallel with undirected edges. Expressions for the number of erased edges have been derived for these also. They are given in this paper without proof. All of these results can be found in Table 3.

Both the Poisson degree distribution and the empirical degree distribution have finite expectations and variances and the resulting plots in Fig. 2 for these are thus tightly connected to the asymptotic results for the number of erased edges in Table 3. A comparison with the simulations that Fig. 2 is based on shows that for the Poisson degree distribution we are approaching the asymptotic results for graphs of size 10^3–10^4 vertices, while for the empirical degree distribution a larger graph size is required. This is most notable for the directed parallel edges for which even the largest simulated graph shows a quite large deviation from the asymptotic results. According to the asymptotic results there should be about 1.4×10^4 parallel directed edges, while there are only about 0.9×10^4 parallel directed edges even in the largest simulated graph with 10^6 vertices. The reason for the slow convergence is the relatively heavy tail of the empirical degree distribution compared with the tail of the Poisson distribution. In the empirical graph the tail is heavier for the directed degrees than for the undirected degrees.

When the expectation of the degree distribution is finite, but the variance is infinite we expect the number of erased edges to grow with the size of the graph, however the details of this are not further explored in this paper.

As already briefly mentioned in Sect. 3.1, for the scale-free and for the Poisson dependent plots there are no erased directed *unconnected* stubs. This is due to the fact that when all nodes have equal in- and out-degree, then the total number of in-stubs will always equal the total number of out stubs exactly. Thus there will not be any directed stubs left over after the graph has been connected so no such stubs will be erased. For the empirical graph this is not the case since the dependent version of the graph is created by sampling from the empirical degrees of the vertices, and for these the number of in-stubs in general do not equal the number of out-stubs. In fact we note that the average number of erased directed stubs per vertex seem to be approximately equal for the dependent and the independent version of the empirical graph, possibly indicating a quite poor correlation between in-stubs and out-stubs in the original graph. This is not surprising, since the empirical graph has a large proportion of reciprocal directed edges and these have been assigned to undirected edges in the partially directed graph.

Another difference between the graphs is that for the scale-free dependent graph there are many more erased directed reciprocal edges, erased directed self loops and erased directed edges that are parallel with an undirected edge, compared with the independent scale-free graph. This can be explained by the heavy tail of the scale-free distribution. For instance, assume that some vertex has a very high degree. Since the degrees are dependent (equal, in this case), the risk is much higher that there will be self loops among the directed edges. Also, since the undirected degree will also be high for this vertex, the risk of having directed edges in parallel with the undirected edges also increases. Finally the chance of getting reciprocal directed edges also increases. This risk is high if there are many vertices with high degrees. In the dependent case if two vertices have many in-stubs both will also have many out-stubs, increasing the chance of parallel edges between these.

3.3 The Strongly Connected Components

Finally we study the strongly connected components in the original data from LiveJournal, compared with the configuration model based on partially directed stubs and also on directed

Table 3 Resulting asymptotic distribution and parameter for the number of erased edges or stubs when the expectations and the variances of the degree distribution are *finite*

Edge type	Distribution and parameter	Poisson degree dist.	Empirical degree dist.
Undirected self loops	$\text{Po}\left(\lambda = \frac{\text{E}[D^{\leftrightarrow}(D^{\leftrightarrow}-1)]}{2\mu^{\leftrightarrow}}\right)$	$\lambda = 3.5$	$\lambda \approx 32.9$
Undirected parallel edges	$\text{Po}\left(\lambda = \left(\frac{\text{E}[D^{\leftrightarrow}(D^{\leftrightarrow}-1)]}{2\mu^{\leftrightarrow}}\right)^2\right)$	$\lambda = 12.25$	$\lambda \approx 1082$
Directed self loops	$\text{Po}\left(\lambda = \frac{\text{E}[D^{\leftarrow}D^{\rightarrow}]}{\mu^{\leftarrow}}\right)$	$\lambda_{ind} = 7; \lambda_{dep} = 8$	$\lambda_{ind} \approx 3.6; \lambda_{dep} \approx 20.4$
Directed parallel edges	$\text{Po}\left(\lambda = \frac{\text{E}[D^{\leftarrow}(D^{\leftarrow}-1)]\text{E}[D^{\rightarrow}(D^{\rightarrow}-1)]}{2(\mu^{\leftarrow})^2}\right)$	$\lambda = 24.5$	$\lambda \approx 1.4 \times 10^4$
Directed reciprocal edges	$\text{Po}\left(\lambda = \left(\frac{\text{E}[D^{\leftarrow}D^{\rightarrow}]}{\mu^{\leftarrow}}\right)^2\right)$	$\lambda_{ind} = 49; \lambda_{dep} = 64$	$\lambda_{ind} \approx 12.6; \lambda_{dep} \approx 415$
Directed parallel with undirected	$\text{Po}\left(\lambda = \frac{\text{E}[D^{\leftarrow}D^{\leftrightarrow}]\text{E}[D^{\rightarrow}D^{\leftrightarrow}]}{\mu^{\leftarrow}\mu^{\leftrightarrow}}\right)$	$\lambda_{ind} = 49; \lambda_{dep} = 64$	$\lambda_{ind} \approx 37.6; \lambda_{dep} \approx 1239$
Directed unconnected stubs	Let $S^{(n)}_{\leftarrow}$ and $S^{(n)}_{\rightarrow}$ be the total number of in-stubs and out-stubs in a graph with n vertices. Then $W^{(n)} = S^{(n)}_{\leftarrow} - S^{(n)}_{\rightarrow}$ is the difference in the number of in-stubs and out-stubs. The number of erased stubs is then $\lvert W^{(n)}\rvert$, which has a folded normal distribution with mean $\mu = \sqrt{n\frac{2}{\pi}\sigma^2_{\Delta s}}$ and variance $\sigma^2 = n\left(1-\frac{2}{\pi}\right)\sigma^2_{\Delta s}$, where $\sigma^2_{\Delta s} = \text{E}\left[(D^{\leftarrow})^2\right] + \text{E}\left[(D^{\rightarrow})^2\right] - 2\text{E}\left[(D^{\leftarrow}D^{\rightarrow})\right]$	$\mu_{ind} \approx 3.0\sqrt{n}; \mu_{dep} = 0$	$\mu_{ind} \approx 28.5\sqrt{n}; \mu_{dep} \approx 27.1\sqrt{n}$

Here $\underline{D} = (D^{\leftarrow}, D^{\rightarrow}, D^{\leftrightarrow})$ is the degree of a randomly chosen vertex from the given degree distribution F. Also $\mu^{\leftarrow} = \text{E}[D^{\leftarrow}]$, $\mu^{\rightarrow} = \text{E}[D^{\rightarrow}]$ and $\mu^{\leftrightarrow} = \text{E}[D^{\leftrightarrow}]$. Note that $\mu^{\leftarrow} = \mu^{\rightarrow}$. The second column gives the distribution and the parameter for the number of erased edges of the specified type and columns three and four give values of the parameters for the independent and the dependent case as described in Table 2. When the parameters differ between the independent case and the dependent case, this has been indicated by specifying two different values for the parameter. For the number of erased directed stubs only the mean has been given in the table

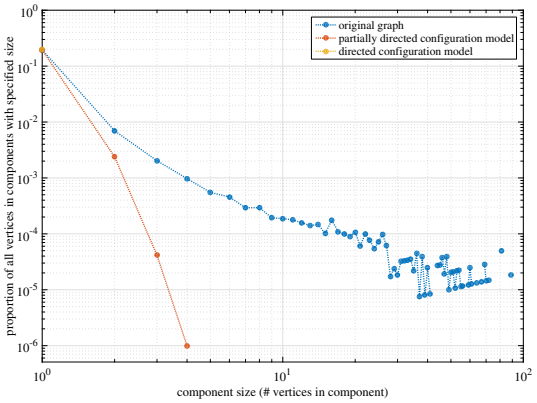

Fig. 3 The figure shows the proportion of all the vertices in the graph that belong to strongly connected components other than the largest component. Three plots are made: the first is for the *original empirical graph*, using the connectivity of the original dataset. The second one is for the configuration model for partially directed graphs with the same degree distribution as for the empirical *partially directed* graph. The third one is for the configuration model for directed graphs with the same degree distribution as the empirical graph, when viewed as a directed graph. The second and third graph are based on averages of ten simulations—results are similar for each simulation. Note that the third plot consists of only a single point, since for this plot all small components only consist of a single vertex. The total number of vertices in the graph is 4,847,571. The relative size of the largest component (not shown in the plot) is 0.7898 for the original graph, 0.8039 for the partially directed configuration model and 0.8026 for the directed configuration model (the last two based on averages of ten simulations, with the standard deviation being approximately 0.0002)

stubs. For any vertex i we define the out-component of vertex i as the set of all vertices that can be reached from vertex i by following the edges of the graph and respecting how they are directed. In the same way the in-component of vertex i is defined as the set of all vertices from which we can reach vertex i. The intersection of the out-component and the in-component defines the strongly connected component of vertex i. Any two vertices i, j where we can reach j from i and i from j, have the same strongly connected component. Thus the graph can be uniquely divided into a set of strongly connected components. Here we study the strongly connected components of the empirical graph and also of configuration model graphs created by using the degree sequence of the empirical graph as the given degree distribution. The largest component in the graph corresponds to the notion of a giant component, the size of which is proportional to the size of the graph. The size of the giant component for these simulations can be compared with theoretical results for a configuration model graph with given degree distribution (see [7, p. 5]). By plugging in the empirical degree distribution of the LiveJournal dataset, we get the theoretical size of the giant component to be 0.8040 for the partially directed graph, and 0.8028 for the directed graph. These values show a good match with the simulation data presented in Fig. 3.

It is not surprising that the largest component is largest in the configuration model for the partially directed graph. The original empirical graph is likely to have sub-communities that may connect only weakly to other communities, thus reducing the total size of the largest strongly connected component, but of course increasing the number of moderately sized strongly connected components. The directed graph lacks the undirected edges and thus the largest strongly connected component will not include vertices that are connected to it only via a directed edge (in one direction only). Thus its largest strongly connected component will be smaller than for the partially directed graph.

When looking at the variation in size among the medium sized components in Fig. 3, this is largest for the original empirical graph. For the configuration model on the directed graph *all* other components consist only of single vertices, while for the configuration model on the partially directed graph components of size 1–4 exist. The appearance of some larger small components for the partially directed graph is caused by the undirected edges, compared with only directed edges for the completely directed graph, as was already mentioned above.

4 Proofs

In this section we provide a proof of Theorem 1. The first part of the proof closely follows [6], with modifications for the joint distribution. In [6] the proof is for the undirected graph, and the addition of the directed edges makes things more complicated. There are mainly two things that need a more detailed treatment, the 3-dimensional degree distribution and the fact that combining undirected and directed edges in the same graph creates new reasons for why edges are erased, affecting the empirical degree distribution and thus also, possibly, the asymptotic behavior of it. The first part of the proof, that is similar to [6] has been moved to two lemmas (1 and 2) to make the part of the proof that is specific for the partially directed configuration model graph more accessible. A third lemma (3) that helps in the final part of the proof of Theorem 1 has also been included.

For Lemma 1, recall that $F^{(n)} = \{p_{\underline{d}}^{(n)}\}$, where $p_{\underline{d}}^{(n)} := \mathrm{E}\left[N_{\underline{d}}^{(n)}/n\right]$.

In the proof we will condition several probabilities and expectations on the degree of vertex one. To shorten the notation we define:

$$\mathrm{P}_{d_1}(\,\cdot\,) = \Pr\left(\,\cdot \mid \underline{D}_1=(d_1^{\leftarrow}, d_1^{\rightarrow}, d_1^{\leftrightarrow})\right) \tag{2}$$

and

$$\mathrm{E}_{d_1}(\,\cdot\,) = \mathrm{E}\left(\,\cdot \mid \underline{D}_1=(d_1^{\leftarrow}, d_1^{\rightarrow}, d_1^{\leftrightarrow})\right) \tag{3}$$

Lemma 1 *$N_{\underline{d}}^{(n)}/n \xrightarrow{P} p_{\underline{d}}$ implies $F^{(n)} \to F$ as $n \to \infty$.*

Proof (i) $N_{\underline{d}}^{(n)}/n \xrightarrow{P} p_{\underline{d}}$ and $0 \le N_{\underline{d}}^{(n)}/n \le 1$ imply $\mathrm{E}\left[N_{\underline{d}}^{(n)}/n\right] \to p_{\underline{d}}$, by bounded convergence [10, p. 180].

(ii) $\mathrm{E}\left[N_{\underline{d}}^{(n)}/n\right] = p_{\underline{d}}^{(n)}$ implies $p_{\underline{d}}^{(n)} \to p_{\underline{d}}\ \forall\, \underline{d}$.

(iii) Since (ii) is valid for any $\underline{d}$ we have $F^{(n)} \to F$ as $n \to \infty$.

□

In Lemma 2 we need a few definitions that are used both in the lemma and in its proof. Let $M_r^{(n)}$ be an indicator variable that shows if vertex r has had its degree modified in the process of creating a simple configuration model graph of size n. The total number of modified vertices can then be calculated by summing all of these and we define $M^{(n)} = \sum_{r=1}^{n} M_r^{(n)}$.

Lemma 2 *If $\Pr\left(M_r^{(n)}=0 \mid \underline{D}_r=(d_r^{\leftarrow}, d_r^{\rightarrow}, d_r^{\leftrightarrow})\right) \to 1\ \forall\, d_r^{\leftarrow}, d_r^{\rightarrow}, d_r^{\leftrightarrow}$ and for arbitrary r, then $N_{\underline{d}}^{(n)}/n \xrightarrow{P} p_{\underline{d}}$ as $n \to \infty$.*

Proof (i) Let $\widetilde{N}_{\underline{d}}^{(n)}$ be the number of vertices with degree $\underline{d}$ before any stub has been erased or added. By the law of large numbers we have that $\widetilde{N}_{\underline{d}}^{(n)}/n \xrightarrow{a.s.} p_{\underline{d}}$ as $n \to \infty$. Since

we want to show that $N_{\underline{d}}^{(n)}/n \xrightarrow{P} p_{\underline{d}}$ it is enough to show that $\left(\widetilde{N}_{\underline{d}}^{(n)} - N_{\underline{d}}^{(n)}\right)/n \xrightarrow{P} 0$ as $n \to \infty$.

(ii) We note that modifying the degree of a vertex affects not only the number of vertices with the original degree, but also the number of vertices with the new degree, thus $\widetilde{N}_{\underline{d}}^{(n)}$ can be less than $N_{\underline{d}}^{(n)}$. However, we can still be sure that $\left|\widetilde{N}_{\underline{d}}^{(n)} - N_{\underline{d}}^{(n)}\right| \leq M^{(n)}$. We wish to show that $M^{(n)}/n \xrightarrow{P} 0$, i.e. that $\Pr\left(\left|M^{(n)}/n\right| > \epsilon\right) \to 0$ as $n \to \infty,\ \forall \epsilon > 0$.

(iii) Using Markov's inequality and that $M^{(n)} \geq 0$ we get

$$\Pr\left(\left|M^{(n)}/n\right| > \epsilon\right) \leq \frac{\mathrm{E}\left[M^{(n)}/n\right]}{\epsilon}; \forall \epsilon > 0. \tag{4}$$

Thus it is enough to show that $\mathrm{E}\left[M^{(n)}/n\right] \to 0$.

(iv) The $\{M_r^{(n)}\}$ are identically distributed since the numbering of the vertices is arbitrary and so $\mathrm{E}[M^{(n)}/n] = \mathrm{E}\left[M_1^{(n)}\right] = \Pr\left(M_1^{(n)}=1\right)$, where vertex 1 has been chosen arbitrarily. We want to show that $\Pr\left(M_1^{(n)}=1\right) \to 0$. The proof could now continue by looking at how the creation of the simple graph can lead to a modification of the degree a vertex. However, there are several ways in which such a modification can occur, even for undirected graphs, and this is further complicated when looking at partially directed graphs. We can avoid this complication by instead studying the probability that a vertex is *saved* from modification. By looking at the actual creation process for the graph we can see that a vertex is saved from modification if, and only if, all stubs of the vertex connect to other unique vertices. Based on this observation we choose to show that $\Pr\left(M_1^{(n)}=0\right) \to 1$ as $n \to \infty$.

(v) Conditioning on the degree of vertex 1 gives

$$\Pr\left(M_1^{(n)}=0\right) = \sum_{d_1^{\leftarrow} d_1^{\rightarrow} d_1^{\leftrightarrow}} \mathrm{P}_{d_1}\left(M_1^{(n)}=0\right) \Pr\left(\underline{D}_1 = (d_1^{\leftarrow}, d_1^{\rightarrow}, d_1^{\leftrightarrow})\right) \tag{5}$$

Since we know

$$\sum_{d_1^{\leftarrow} d_1^{\rightarrow} d_1^{\leftrightarrow}} \Pr\left(\underline{D}_1 = (d_1^{\leftarrow}, d_1^{\rightarrow}, d_1^{\leftrightarrow})\right) = \sum_{d_1^{\leftarrow} d_1^{\rightarrow} d_1^{\leftrightarrow}} p_{d_1^{\leftarrow} d_1^{\rightarrow} d_1^{\leftrightarrow}} = 1 \tag{6}$$

it is enough to show that

$$\mathrm{P}_{d_1}\left(M_1^{(n)}=0\right) \to 1\ \forall\ d_1^{\leftarrow}, d_1^{\rightarrow}, d_1^{\leftrightarrow} \text{ as } n \to \infty. \tag{7}$$

□

Lemma 3 *Let $\{X_m\}$ be a sequence of non-negative random variables and let X be a non-negative random variable. Also let $0 \leq a < \infty$ be a real number.*

If $X_m \xrightarrow{D} X$ as $m \to \infty$, $\lim_{m\to\infty} \mathrm{E}[X_m] \leq a$ and $\mathrm{E}[X] = a$, then $\lim_{m\to\infty} \mathrm{E}[X_m] = a$.

Proof For non-negative random variables $\{Y_m\}$ Fatou's lemma states

$$\mathrm{E}\left[\liminf_{m\to\infty} Y_m\right] \leq \liminf_{m\to\infty} \mathrm{E}\left[Y_m\right]. \tag{8}$$

We apply Skorokhod's representation theorem and can thus define $\{Y_m\}$ and Y (all on the same probability space) to have the same distribution as $\{X_m\}$ and X, and $Y_m \xrightarrow{a.s.} Y$ as $m \to \infty$

Developing the left and right hand side of Fatou's Lemma now gives:

$$\text{LHS} = \mathrm{E}\left[\liminf_{m\to\infty} Y_m\right] = \mathrm{E}[Y] = \mathrm{E}[X] = a, \tag{9}$$

$$\text{RHS} = \liminf_{m\to\infty} \mathrm{E}[Y_m] \leq \lim_{n\to\infty} \mathrm{E}[Y_m] = \lim_{m\to\infty} \mathrm{E}[X_m] \leq a. \tag{10}$$

Thus

$$\lim_{m\to\infty} \mathrm{E}[X_m] = a \tag{11}$$

□

Now we are ready to prove the main theorem.

Proof (*of Theorem* 1)

(1) Lemma 1 shows that Theorem 1(b) implies (a).

(2) It remains to prove Theorem 1(b). Lemma 2 simplifies this process.
Let $M_1^{(n)}$ be the *indicator* variable for the event that a specific vertex (arbitrarily selected to be vertex 1) has had its degree *modified* when creating a simple configuration model graph of size n according to the procedure defined in Sect. 2.2. Also let the degree of vertex 1 be $\underline{D}_1 = \underline{d}_1 = (d_1^{\leftarrow}, d_1^{\rightarrow}, d_1^{\leftrightarrow})$. According to Lemma 2, in order to prove (b) it is sufficient to show that

$$\mathrm{P}_{d_1}\left(M_1^{(n)}{=}0\right) \to 1 \ \forall\, \underline{d}_1 \text{ as } n \to \infty \tag{12}$$

(3) Remembering that we do not allow self loops or parallel edges, $M_1^{(n)} = 0$ exactly when each stub from vertex 1 is saved. In total, vertex 1 has $d = d_1^{\leftarrow}+d_1^{\rightarrow}+d_1^{\leftrightarrow}$ stubs and these are all saved only when all of them successfully attach to other matching stubs, all from different vertices selected from vertices $\{2, \ldots, n\}$. In all other cases the degree of vertex 1 will surely be modified, giving no contribution to the probability of $M_1^{(n)}{=}0$. Now, if we knew the degrees of all the vertices, it would be easy to calculate the probability of $M_1^{(n)}{=}0$. We do this simply by considering *all* events where the stubs of vertex 1 connect to different vertices and then sum all the probabilities of these events. It is thus natural to continue the proof by conditioning on the degrees of vertices $\{2, \ldots, n\}$. Let the degrees of vertices $\{2, \ldots, n\}$ be $\mathbf{D}^{(n)} = \{\underline{D}_2, \ldots, \underline{D}_n\}$, where the $\underline{D}_r$ are i.i.d. from F. Then we want to study

$$\mathrm{P}_{d_1}\left(M_1^{(n)} = 0\right) = \mathrm{E}_{d_1}\left[\mathrm{P}_{d_1}\left(M_1^{(n)} = 0 \,|\, \mathbf{D}^{(n)}\right)\right]. \tag{13}$$

(4) We now look more closely at the conditional probability

$$\mathrm{P}_{d_1}\left(M_1^{(n)} = 0 \,|\, \mathbf{D}^{(n)} = \mathbf{d}^{(n)}\right), \tag{14}$$

where $\mathbf{D}^{(n)} = \mathbf{d}^{(n)} = \{\underline{d}_2, \ldots, \underline{d}_n\}$ is a specific outcome of the degrees of the vertices. From this we see that the total number of stubs of each type are $s^{(n)}_{\leftarrow} = \sum_{r=1}^{n} d_r^{\leftarrow}$, $s^{(n)}_{\rightarrow} = \sum_{r=1}^{n} d_r^{\rightarrow}$ and $s^{(n)}_{\leftrightarrow} = \sum_{r=1}^{n} d_r^{\leftrightarrow}$. We want to know where each stub of vertex 1 *attempts* to connect and define a set set of indices, $\mathbf{i} = \{i_1, \ldots, i_{d_1^{\leftarrow}}\}$, $\mathbf{j} = \{j_1, \ldots, j_{d_1^{\rightarrow}}\}$ and $\mathbf{k} = \{k_1, \ldots, k_{d_1^{\leftrightarrow}}\}$. Any set of values of these indices we call a *save-attempt*, indicating that we try to save all stubs of vertex 1 from being erased, by attempting to connect the stubs of vertex 1 to matching stubs from the vertices pointed to by these indices.

Given the degrees of all vertices we can calculate the probability of any such *save-attempt*. First some basic observations:

(a) If any one of the selected vertices does not have a matching stub the probability of the *save-attempt* is zero. As an example, assume that an in-stub attempts to connect to vertex 2, but vertex 2 does not have any out-stub at all. Then this event will have probability zero.
(b) As a consequence, for the *save-attempt* to have a probability larger than zero, *all* the vertices that the stubs of vertex 1 attempt to connect to must have matching stubs.

As an example, take a look at the *save-attempt* where each stub of vertex 1 tries to connect to the other vertices in order. The indices then take on the values $\{i_1 = 2, i_2 = 3, \ldots, k_{d_1^{\leftrightarrow}-1} = d, k_{d_1^{\leftrightarrow}} = d+1\}$. For now, we ignore the probability that there may not be enough matching stubs of vertices $\{2, \ldots, n\}$ to accommodate all the stubs of vertex 1. We do this now to make the main argument clearer, but we correct the equations for this special case later in the proof.
First we look at in-stub 1 from vertex 1. Since we are working with the configuration model, this stub has an equal chance of connecting to any of the matching stubs. Thus the probability that in-stub 1 from vertex 1 connects to any of the out-stubs from vertex 2 is

$$\frac{d_2^{\rightarrow}}{s^{(n)\rightarrow}}. \tag{15}$$

Once in-stub 1 of vertex 1 has connected to vertex 2 we continue with in-stub 2 of vertex 1. Once again the configuration model tells us that this stub has an equal chance of connecting to any of the remaining matching stubs. Thus the probability of it connecting to any of the out-stubs from vertex 3 is

$$\frac{d_3^{\rightarrow}}{s^{(n)\rightarrow} - 1}. \tag{16}$$

We can continue in the same way with the rest of the in-stubs, then the out-stubs and finally the undirected stubs of vertex 1. For the undirected stubs we note that we need to subtract 2 stubs every time we connect one stub, since the undirected stubs connect to other undirected stubs.
Now we can calculate the probability of this specific *save-attempt* and find that it is

$$\prod_{r=1}^{d_1^{\leftarrow}} \left(\frac{d_{i_r}^{\rightarrow}}{s^{(n)\rightarrow} - r + 1} \right) \prod_{r=1}^{d_1^{\rightarrow}} \left(\frac{d_{j_r}^{\leftarrow}}{s^{(n)\leftarrow} - r + 1} \right) \prod_{r=1}^{d_1^{\leftrightarrow}} \left(\frac{d_{k_r}^{\leftrightarrow}}{s^{(n)\leftrightarrow} - 2r + 1} \right) \tag{17}$$

In the expression we have ignored that we have already used $d_1^{\leftarrow}$ out-stubs when connecting the in-stubs of vertex 1. We correct for this in the final expressions given later in the proof.
Here we explicitly see that this expression is equal to zero iff any one of the degrees in the numerator is zero. Otherwise it will be positive, but always less than or equal to 1.
To shorten the expressions we will call each of the three parts of Eq. 17 $q_{\mathbf{i}}^{(n)\leftarrow}$, $q_{\mathbf{j}}^{(n)\rightarrow}$ and $q_{\mathbf{k}}^{(n)\leftrightarrow}$, respectively, where the arrow indicates what type of stub in vertex 1 we are dealing with.

Now we are ready to write down the expression for the conditional probability in Eq. 14 We need to sum Eq. 17 over *all* values of $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, such that *all* sub-indices are different - pointing to different vertices. We arrive at

$$\mathrm{P}_{d_1}\left(M_1^{(n)}{=}0 \,|\, \mathbf{D}^{(n)} = \mathbf{d}^{(n)}\right) = \sum_{\substack{\mathbf{i},\mathbf{j},\mathbf{k} \\ \text{all sub-indices different}}} q_{\mathbf{i}}^{\overset{(n)}{\leftarrow}} q_{\mathbf{j}}^{\overset{(n)}{\rightarrow}} q_{\mathbf{k}}^{\overset{(n)}{\leftrightarrow}}. \tag{18}$$

The number of terms in the sum will be $(n-1)(n-2)\cdots(n-d)$, which is simply the number of different ways in which we can select the d indices out of the $n-1$ possible vertices. Note that these combinations of indices include the ones we are interested in, where all stubs of vertex 1 are saved. Note also that the sum includes some combinations that we are not interested in, but all of these have probability zero and so it does not matter if we include them in the sum or not.

(5) We now need to deal with a few complications that will lead to corrections to $q_{\mathbf{i}}^{\overset{(n)}{\leftarrow}}$, $q_{\mathbf{j}}^{\overset{(n)}{\rightarrow}}$ and $q_{\mathbf{k}}^{\overset{(n)}{\leftrightarrow}}$.

(a) If the number of stubs of vertex 1 (d) is larger than the number of available vertices ($n-1$), then it is not possible to select all the sub-indices indices to be different. However, since d is fixed, this is always resolved as $n \to \infty$. In the following we will always assume that $n \geq d$.

(b) There may be a mismatch in the number of stubs. If the number of undirected stubs is odd, there will be one extra stub. Let $v^{(n)}$ be the number of such stubs. Clearly $v^{(n)}$ can only be 0 or 1.
In the same way the number of in-stubs may differ from the number of out-stubs. Let $w^{(n)} = s^{\overset{(n)}{\leftarrow}} - s^{\overset{(n)}{\rightarrow}}$, the difference between the number of in-stubs and the number of out-stubs. Clearly $w^{(n)}$ can be negative, zero or positive. If $v^{(n)}$ or $w^{(n)}$ are not zero then some stubs will remain unconnected.
In the following we will deal with both of these by imagining two extra *pools* of edges each of size $v^{(n)}$ and $|w^{(n)}|$, respectively. These pools behave just as any normal vertex and any stub has an equal probability to connect to any allowed stub, including these two pools. They are thus added to the denominators in Eq. 17

(c) As mentioned before, we have included some events that have probability zero in the sum. Although the numerator is always zero for these, in some cases the denominator may also become zero. This happens when there are not enough matching stubs to accommodate all the stubs of vertex 1. We deal with this by adding an extra indicator variable to the denominator so that it does not become zero, thus ensuring that these events do not contribute anything to the sum.

The corrected versions of $q_{\mathbf{i}}^{\overset{(n)}{\leftarrow}}$, $q_{\mathbf{j}}^{\overset{(n)}{\rightarrow}}$ and $q_{\mathbf{k}}^{\overset{(n)}{\leftrightarrow}}$ are thus

$$q_{\mathbf{i}}^{\overset{(n)}{\leftarrow}} = \prod_{r=1}^{d_1^{\leftarrow}} \frac{d_{i_r}^{\rightarrow}}{s^{\overset{(n)}{\rightarrow}} - r + 1 + w^{(n)} \mathbb{1}_{\{w^{(n)}>0\}} + d_1^{\leftarrow} \mathbb{1}_{\left\{d_1^{\leftarrow} > s^{\overset{(n)}{\rightarrow}}\right\}}} \tag{19}$$

$$q_{\mathbf{j}}^{\overset{(n)}{\rightarrow}} = \prod_{r=1}^{d_1^{\rightarrow}} \frac{d_{j_r}^{\leftarrow}}{s^{\overset{(n)}{\leftarrow}} - d_1^{\leftarrow} - r + 1 - w^{(n)} \mathbb{1}_{\{w^{(n)}<0\}} + (d_1^{\rightarrow} + d_1^{\leftarrow}) \mathbb{1}_{\left\{d_1^{\rightarrow} > s^{\overset{(n)}{\leftarrow}} - d_1^{\leftarrow}\right\}}} \tag{20}$$

$$q_{\mathbf{k}}^{(n)\leftrightarrow} = \prod_{r=1}^{d_1^{\leftrightarrow}} \frac{d_{k_r}^{\leftrightarrow}}{s^{(n)\leftrightarrow} - 2r + 1 + v^{(n)} + 2d_1^{\leftrightarrow}\mathbb{1}_{\left\{2d_1^{\leftrightarrow} > s^{(n)\leftrightarrow}\right\}}} \tag{21}$$

(6) To be able to obtain an expression for the probability in Eq. 14 we need to replace the degrees in Eq. 21 with their stochastic counterpart to obtain

$$\mathrm{P}_{d_1}\left(M_1^{(n)} = 0 \mid \mathbf{D}^{(n)}\right) = \sum_{\substack{\mathbf{i},\mathbf{j},\mathbf{k} \\ \text{all sub-indices different}}} Q_{\mathbf{i}}^{(n)\leftarrow} Q_{\mathbf{j}}^{(n)\rightarrow} Q_{\mathbf{k}}^{(n)\leftrightarrow}, \tag{22}$$

where

$$Q_{\mathbf{i}}^{(n)\leftarrow} = \prod_{r=1}^{d_1^{\leftarrow}} \frac{D_{i_r}^{\rightarrow}}{S^{(n)\rightarrow} - r + 1 + W^{(n)}\mathbb{1}_{\{W^{(n)}>0\}} + d_1^{\leftarrow}\mathbb{1}_{\left\{d_1^{\leftarrow} > S^{(n)\rightarrow}\right\}}} \tag{23}$$

$$Q_{\mathbf{j}}^{(n)\rightarrow} = \prod_{r=1}^{d_1^{\rightarrow}} \frac{D_{j_r}^{\leftarrow}}{S^{(n)\leftarrow} - d_1^{\leftarrow} - r + 1 - W^{(n)}\mathbb{1}_{\{W^{(n)}<0\}} + (d_1^{\rightarrow} + d_1^{\leftarrow})\mathbb{1}_{\left\{d_1^{\rightarrow} > S^{(n)\leftarrow} - d_1^{\leftarrow}\right\}}} \tag{24}$$

$$Q_{\mathbf{k}}^{(n)\leftrightarrow} = \prod_{r=1}^{d_1^{\leftrightarrow}} \frac{D_{k_r}^{\leftrightarrow}}{S^{(n)\leftrightarrow} - 2r + 1 + V^{(n)} + 2d_1^{\leftrightarrow}\mathbb{1}_{\left\{2d_1^{\leftrightarrow} > S^{(n)\leftrightarrow}\right\}}} \tag{25}$$

Here the uppercase variables are all the stochastic counterparts of the lowercase variables defined previously.

(7) Now we can continue with Eq. 13:

$$\begin{aligned}
\mathrm{P}_{d_1}\left(M_1^{(n)}=0\right) &= \mathrm{E}_{d_1}\left[\mathrm{P}_{d_1}\left(M_1^{(n)}=0 \mid \mathbf{D}^{(n)}\right)\right] && (26)\\
&= \mathrm{E}_{d_1}\left[\sum_{\substack{\mathbf{i},\mathbf{j},\mathbf{k} \\ \text{all sub-indices different}}} Q_{\mathbf{i}}^{(n)\leftarrow} Q_{\mathbf{j}}^{(n)\rightarrow} Q_{\mathbf{k}}^{(n)\leftrightarrow}\right] && (27)\\
&= \sum_{\substack{\mathbf{i},\mathbf{j},\mathbf{k} \\ \text{all sub-indices different}}} \mathrm{E}_{d_1}\left[Q_{\mathbf{i}}^{(n)\leftarrow} Q_{\mathbf{j}}^{(n)\rightarrow} Q_{\mathbf{k}}^{(n)\leftrightarrow}\right] && (28)\\
&= (n-1)(n-2)\cdots(n-d)\mathrm{E}_{d_1}\left[Q_{\mathbf{i}}^{(n)\leftarrow} Q_{\mathbf{j}}^{(n)\rightarrow} Q_{\mathbf{k}}^{(n)\leftrightarrow}\right] && (29)\\
&= \frac{(n-1)\cdots(n-d)}{n^d}\mathrm{E}_{d_1}\left[n^d Q_{\mathbf{i}}^{(n)\leftarrow} Q_{\mathbf{j}}^{(n)\rightarrow} Q_{\mathbf{k}}^{(n)\leftrightarrow}\right] && (30)\\
&= c^{(n)}\mathrm{E}_{d_1}\left[\left(n^{d_1^{\leftarrow}} Q_{\mathbf{i}}^{(n)\leftarrow}\right)\left(n^{d_1^{\rightarrow}} Q_{\mathbf{j}}^{(n)\rightarrow}\right)\left(n^{d_1^{\leftrightarrow}} Q_{\mathbf{k}}^{(n)\leftrightarrow}\right)\right]. && (31)
\end{aligned}$$

Note 1 The expectation and the summation can be interchanged since all terms are non-negative and since the summation does not depend on any random quantity (as mentioned before).

Note 2 Since vertex degrees are drawn independently at random, all expectation terms in the sum are identical and we simply take the number of terms times the expectation of one of the terms instead of the sum. The number of terms was already discussed above.
Note 3 $c^{(n)} = \frac{(n-1)\cdots(n-d)}{n^d}$.

(8) All that remains is to take the limit of Eq. 31. We start by studying the limit of what is inside the expectation. Rewriting the first term we get

$$n^{d_1^{\leftarrow}} Q_{\mathbf{i}}^{(n)\leftarrow} = n^{d_1^{\leftarrow}} \prod_{r=1}^{d_1^{\leftarrow}} \frac{D_{i_r}^{\rightarrow}}{S^{(n)\rightarrow} - r + 1 + W^{(n)} \mathbb{1}_{\{W^{(n)}>0\}} + d_1^{\leftarrow} \mathbb{1}_{\left\{d_1^{\leftarrow} > S^{(n)\rightarrow}\right\}}} \tag{32}$$

$$= \prod_{r=1}^{d_1^{\leftarrow}} \frac{D_{i_r}^{\rightarrow}}{\frac{S^{(n)\rightarrow}}{n} - \frac{r}{n} + \frac{1}{n} + \frac{W^{(n)}}{n} \mathbb{1}_{\{W^{(n)}>0\}} + \frac{d_1^{\leftarrow}}{n} \mathbb{1}_{\left\{d_1^{\leftarrow} > S^{(n)\rightarrow}\right\}}} \tag{33}$$

The remaining outgoing and undirected terms will be very similar, producing the additional terms $S^{(n)\leftarrow}/n$, $S^{(n)\leftrightarrow}/n$, $V^{(n)}/n$, $d_1^{\rightarrow}/n$ and $d_1^{\leftrightarrow}/n$ in the denominators.
Now note that, since $\mathrm{P}_{d_1}\left(M_1^{(n)}=0\right) \leq 1$ and $\lim_{n\to\infty} c^{(n)} = 1$, we have that

$$\lim_{n\to\infty} \mathrm{E}_{d_1}\left[\left(n^{d_1^{\leftarrow}} Q_{\mathbf{i}}^{(n)\leftarrow}\right)\left(n^{d_1^{\rightarrow}} Q_{\mathbf{j}}^{(n)\rightarrow}\right)\left(n^{d_1^{\leftrightarrow}} Q_{\mathbf{k}}^{(n)\leftrightarrow}\right)\right] \leq 1 \tag{34}$$

By the law of large numbers and using Slutsky's Theorem

$$\left(n^{d_1^{\leftarrow}} Q_{\mathbf{i}}^{(n)\leftarrow}\right)\left(n^{d_1^{\rightarrow}} Q_{\mathbf{j}}^{(n)\rightarrow}\right)\left(n^{d_1^{\leftrightarrow}} Q_{\mathbf{k}}^{(n)\leftrightarrow}\right) \xrightarrow{D} \left(\frac{\prod_{r=1}^{d_1^{\leftarrow}} D_{i_r}^{\rightarrow}}{(\mu^{\rightarrow})^{d_1^{\leftarrow}}}\right)\left(\frac{\prod_{r=1}^{d_1^{\rightarrow}} D_{j_r}^{\leftarrow}}{(\mu^{\rightarrow})^{d_1^{\rightarrow}}}\right)\left(\frac{\prod_{r=1}^{d_1^{\leftrightarrow}} D_{k_r}^{\leftrightarrow}}{(\mu^{\rightarrow})^{d_1^{\leftrightarrow}}}\right) \tag{35}$$

Here we used that $\lim_{n\to\infty} \frac{W^{(n)}}{n} = \lim_{n\to\infty} \frac{S^{(n)\leftarrow} - S^{(n)\rightarrow}}{n} = \mu^{\leftarrow} - \mu^{\rightarrow} = 0$, by assumption. $V^{(n)}$, which indicates if the number of undirected edges is odd, is at most 1.
Since all $D_{j_r}^{\leftarrow}$, $D_{i_r}^{\rightarrow}$ and $D_{k_r}^{\leftrightarrow}$ are independent by construction we also have

$$\mathrm{E}_{d_1}\left[\left(\frac{\prod_{r=1}^{d_1^{\leftarrow}} D_{i_r}^{\rightarrow}}{(\mu^{\rightarrow})^{d_1^{\leftarrow}}}\right)\left(\frac{\prod_{r=1}^{d_1^{\rightarrow}} D_{j_r}^{\leftarrow}}{(\mu^{\rightarrow})^{d_1^{\rightarrow}}}\right)\left(\frac{\prod_{r=1}^{d_1^{\leftrightarrow}} D_{k_r}^{\leftrightarrow}}{(\mu^{\rightarrow})^{d_1^{\leftrightarrow}}}\right)\right] = 1. \tag{36}$$

Now we use Lemma 3 and immediately conclude that

$$\lim_{n\to\infty} \mathrm{E}_{d_1}\left[\left(n^{d^{\leftarrow}} Q_{\mathbf{i}}^{(n)\leftarrow}\right)\left(n^{d^{\rightarrow}} Q_{\mathbf{j}}^{(n)\rightarrow}\right)\left(n^{d^{\leftrightarrow}} Q_{\mathbf{k}}^{(n)\leftrightarrow}\right)\right] = 1 \tag{37}$$

and thus also that

$$\lim_{n\to\infty} \mathrm{P}_{d_1}\left(M_1^{(n)}=0\right) = \lim_{n\to\infty} c^{(n)} \mathrm{E}_{d_1}\left[\left(n^{d_1^{\leftarrow}} Q_{\mathbf{i}}^{(n)\leftarrow}\right)\left(n^{d_1^{\rightarrow}} Q_{\mathbf{j}}^{(n)\rightarrow}\right)\left(n^{d_1^{\leftrightarrow}} Q_{\mathbf{k}}^{(n)\leftrightarrow}\right)\right] = 1. \tag{38}$$

This is what we wanted to show and so the proof is complete. □

5 Conclusions and Discussion

We have shown a way to create a partially directed configuration model graph from a given joint degree distribution. The graph is simple, and under specified conditions the degree distribution converges to the desired one. The only assumptions in the proof are that the degrees of different vertices are independent, that the expectation of the degree of each type of stub is finite and that the expectation of the degree for the in-stubs is equal to the expectation for the degree of the out-stubs. This means that the proof works also for undirected and for directed configuration model graphs, and also if the number of different types of stubs is increased to any finite number, as long as similar conditions as in this proof are fulfilled. The main idea of the proof is that a vertex is saved from modification if all of its stubs are connected to unique vertices. If the requirement for a simple graph is relaxed and self loops or parallel edges are allowed to remain in the graph, this only increases the chance of saving a vertex from having its degree modified and so is not a problem.

The main advantage of using a partially directed model to represent empirical networks, as opposed to using a completely directed or completely undirected model, is that the partially directed model preserves the proportion of undirected edges. This is important for networks where there is a significant proportion both of directed and of undirected edges, and where none of the different types of edges can be ignored. Examples of such graphs have been given in Table 1. The model also preserves any dependence between directed and undirected degrees present in the original empirical graph or the given degree distribution.

However, this model does not produce other structures that can often be found in empirical networks. E.g. it does not produce the same number of moderately sized strongly connected components that we see in the empirical networks. In this respect it does however perform slightly better than the configuration model on directed graphs. Possible improvements towards realism would be to see how e.g. triangles (of different types), different types of vertices and other heterogeneities could be included in the model.

Acknowledgments The authors would like to thank Pieter Trapman for discussions and for giving valuable input to the paper. K.S. was supported by the Swedish Research Council, Grant No. 2009-5759. T.B. is grateful to Riksbankens jubileumsfond (contract P12-0705:1) for financial support.

References

1. Chen, N., Olvera-Cravioto, M.: Directed random graphs with given degree distributions. Stoch. Syst. **3**(1), 147–186 (2013)
2. Meyers, L.A., Newman, M.E.J., Pourbohloul, B.: Predicting epidemics on directed contact networks. J. Theor. Biol. **240**, 400–418 (2006)
3. Malmros, J., Masuda, N., Britton, T.: Random Walks on Directed Networks: Inference and Respondent-Driven Sampling, Research Report 2013:7. Mathematical Statistics. Stockholm University, Stockholm (2013)
4. Molloy, M., Reed, B.: A critical point for random graphs with a given degree sequence. Random Struct. Algorithms **6**, 161–180 (1995)
5. Bollobás, B.: Random Graphs. Cambridge University Press, Cambridge (2001)
6. Britton, T., Deijfen, M., Martin-Löf, A.: Generating simple random graphs with prescribed degree distribution. J. Stat. Phys. **125**(6), 1377–1397 (2006)

7. Kenah, E., Robins, J.M.: Second look at the spread of epidemics on networks. Phys. Rev. **76**, 036113 (2007)
8. Janson, S.: The probability that a random multigraph is simple. Comb. Probab. Comput. **18**(1–2), 205–225 (2009)
9. Leskovec J., Krevl A.: SNAP Datasets, Stanford large network dataset collection. http://snap.stanford.edu/data (2014)
10. Grimmet, G., Stirzaker, D.: Probability and Random Processes, 3rd edn, pp. 126–127. Oxford University Press, Oxford (2009)
11. Ross, S.M.: Introduction to Probability Models, 11th edn. Academic Press, Burlington (2014)
12. van der Hofstad, R.: Lecture notes: random graphs and complex networks (2014). http://www.win.tue.nl/~rhofstad/NotesRGCN. Accessed 24 June 2015