Journal of Complex Networks Advance Access published October 13, 2016

Journal of Complex Networks (2016) **0000**, 1–15
doi: 10.1093/comnet/cnw027

Switching edges to randomize networks: what goes wrong and how to fix it

C. J. Carstens[†,‡] and K. J. Horadam
School of Mathematical and Geospatial Sciences, RMIT University, 124 La Trobe Street, Melbourne VIC 3000, Australia and [‡]*Present address: Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Science Park 105, 1098 XG Amsterdam, The Netherlands*
[†]Corresponding author. Email: c.j.carstens@uva.nl

Edited by: Ali Pinar

[Received on 19 February 2016; accepted on 6 September 2016]

The switching model is a well-known random network model that randomizes a network while keeping its degree sequence fixed. The idea behind the switching model is simple: a network is randomized by repeatedly rewiring pairs of edges. In this paper we demonstrate that despite its simple description, and in part due to it, much can go wrong when implementing the switching model. Specifically, we show that the model needs to be implemented carefully, to avoid biased sampling. We propose a precise definition of the switching model which guides its implementation. Furthermore, we argue that we should refer to the switching model *with respect to* a specific network class, and in fact define a *family* of switching models. This formalizes previous use of the switching model to randomize networks from numerous network classes. We show that the properties of these models depend on the network class, and in particular that their stationary distributions differ. Hence it is important to take into account which class of networks is being randomized. We derive conditions, and where possible adjust the models, such that sampling is unbiased for the switching model with respect to eight common network classes. These unbiased null-models are important, since common network analysis techniques such as motif finding and community detection rely on them.

Keywords: random networks; Markov chain; switching model; unbiased sampling; fixed degree sequence.

1. Introduction

The randomization of networks with fixed degree sequence is used extensively throughout network science [1–6]. It has turned out to be hard to generate provably unbiased samples [1, 7–10]. There are two types of approach to this problem: building a network from scratch as in the configuration model [4–6] or randomizing a network by making small changes to it as in the switching model [1, 9, 11–14]. We focus on the latter Markov chain approach to network randomization.

In this paper we point out that the simple and flexible nature of the switching algorithm comes at a price. First, in order to obtain unbiased samples with the switching model, a specific Markov chain has to be implemented [1, 9, 12]. Second, even though the switching model has been used to randomize different classes of networks, it is typically referred to as *the* switching model [15]. We argue that really we should be talking about the *family* of switching models. We demonstrate that it is important to distinguish the switching model with respect to different classes of networks, because ignoring this distinction has led to biased sampling. Third, we identify mistakes in its implementation that are easily overlooked but introduce bias in sampling.

We use Markov chain terminology to propose a simple and flexible, yet unambiguous definition of the family of switching models. Furthermore, using a well-known theorem about finite Markov chains

(Theorem 2.1), we analyse it with respect to eight distinct network classes. In particular we derive the conditions for which this family of switching models converges to a stationary distribution, as well as their stationary distributions. When the stationary distribution is not the uniform distribution we propose small changes to the particular switching model to ensure unbiased sampling.

Throughout this paper, we illustrate how the theoretical properties of the switching model relate to its implementation. We show that there are many subtleties in the implementation of switching models that may cause biased sampling. For instance, we show that an implementation error in the software package MFinder [16] has previously produced biased samples for undirected networks. Ultimately, to ensure that the results of experimental studies are correct, it is of great importance that correct algorithms are available or easily implemented based on theoretical papers.

An important question that we do not address in this paper is the number of steps required to reach the stationary distribution of the switching models. Instead, we focus on how to ensure this distribution is uniform. However, taking too few steps in the algorithm will also introduce sampling bias. There are several theoretical results but only for special network classes [17–19]. Furthermore, the theoretical limits are too large to be practical [20]. We refer to [21] for a recent discussion on how to bridge the gap between theoretical limits and practical run-times.

This paper is organized as follows. In Section 2 we give a brief overview on Markov chains. We then discuss two Markov chains corresponding to the switching model that have different sampling distributions. We propose a simple and precise definition of the switching model that directs its implementation.

In Section 3 we generalize the definition of the switching model to any subset of the class of directed multigraphs. We analyse the theoretical properties of the switching model with respect to four common subsets and derive the corresponding sampling distributions. We also propose simple alterations to ensure sampling is unbiased.

In Section 4 we propose a similar definition of the switching model with respect to any subset of the class of undirected multigraphs. We show that the switching model with respect to undirected networks has its own subtleties, and how, when these are not taken into account, sampling can be very biased.

In Section 5 we present selected proofs of the theoretical properties of the switching models as described in Sections 3 and 4. The remainder appears in the Supplementary data.

We conclude by summarizing our results and making recommendations in Section 6.

2. Different Markov chains corresponding to the switching model

In order to derive the sampling distribution of the switching model, we use the well-known fact that it corresponds to a Markov chain [1, 12, 13, 22]. That is, the switching model generates a stochastic sequence of networks where the probability of a network occurring at some point in the sequence only depends on its immediate predecessor. This stochastic sequence starts with a network G, followed by networks obtained from repeatedly applying switches.

We will use the following facts and terminology about finite Markov chains. A finite Markov chain is uniquely defined by its finite state space $\{G_i\}$ and the transition probabilities p_{ij} between each pair of states G_i and G_j. The state graph of a Markov chain is a graph with vertices that correspond to its states and edges that correspond to non-zero transition probabilities. A Markov chain is *irreducible* if its state graph is strongly connected. A Markov chains is called *aperiodic* if for all states x, the greatest common divisor of the length of walks starting and ending at x equals 1. A finite Markov chain converges to a stationary distribution under the following conditions (see, e.g. [23, Theorem 7.10]).

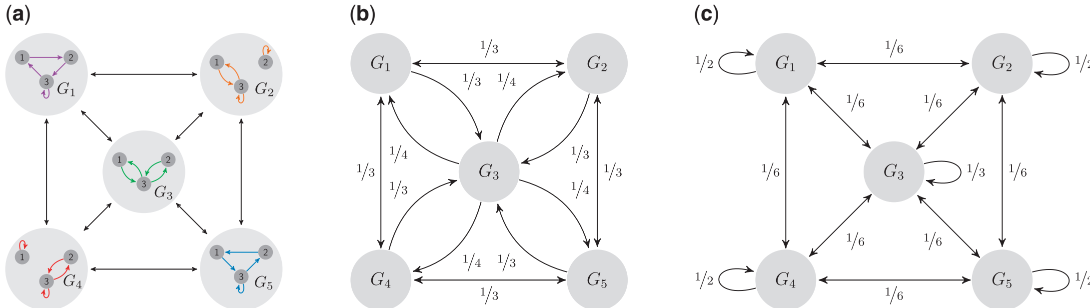

FIG. 1. (a) The five directed networks with degree sequences $k^{\text{in}} = (1, 1, 2)$ and $k^{\text{out}} = (1, 1, 2)$. Networks are connected by an edge if there is a switch that transforms one into the other. For instance there is an edge from network G_1 to network G_2 since replacing edges $(1, 2)$ and $(2, 3)$ in network G_1 by edges $(1, 3)$ and $(2, 2)$ results in network G_2. (b) The transition probabilities corresponding to the interpretation of the switching model excluding repeated states. (c) The transition probabilities corresponding to the interpretation of the switching model including repeated states.

THEOREM 2.1 A finite irreducible and aperiodic Markov chain converges to a *unique* stationary distribution. If there exists a probability distribution π on its state space such that the *detailed balance equations*, $\pi_i p_{ij} = \pi_j p_{ji}$, are satisfied for all i, j, then π is this unique stationary distribution.

This theorem implies that a finite, irreducible and aperiodic Markov chain converges to the uniform distribution if $p_{ij} = p_{ji}$ for all i, j.

We now focus on the randomization of a single class of networks: directed networks. In this paper, directed networks may include self-loops; directed networks without self-loops are referred to as *simple* directed networks. The switching model for these networks was introduced within network science as follows: 'Let G be a directed network, randomly select two edges (x, y) and (u, v) and replace them by (x, v) and (u, y). In case one or both of these new edges already exist in G, abort this step and select a new pair of edges. Repeat this until the network is sufficiently randomized.' [11, 16]

This definition is ambiguous with respect to the abortion of a switch: it is unclear whether aborting a switch corresponds to repeating a network in the corresponding Markov chain, or not. This seemingly innocent distinction makes the difference between biased and unbiased sampling, as has been discussed a number of times in the literature [1, 9, 12]. The following two examples show that two different ways of addressing this issue, in terms of repeating states [1, 12] and in terms of fixing the number of switches [9], are in fact the same.

EXAMPLE 2.2 Let G_1 be the directed network illustrated in Fig. 1(a). Its in-degree sequence k^{in} equals $(1, 1, 2)$ and its out-degree sequence k^{out} equals $(1, 1, 2)$. There are exactly five directed networks with these degree sequences, as shown in Fig. 1(a).

The network G_1 has four edges, hence there are six pairs of edges that can potentially be switched. Only three of these edge pairs result in different directed networks when switched: $(1, 2)$ and $(2, 3)$ result in network G_2, $(1, 2)$ and $(3, 3)$ result in network G_3, and $(1, 2)$ and $(3, 1)$ result in network G_4.

In the first interpretation of the switching model, e.g. [1], the Markov chain changes state in every step. To do so it selects a neighbour of the current state uniformly at random. Hence the transition probability p_{ij}, between states that differ by a switch, equals $1/k(G_i)$ with $k(G_i)$ the degree of G_i. For instance, the transition probability from state G_1 to G_2 equals $1/3$ because G_1 has three neighbours. The Markov chain

(a)

```
for(i in 1:N){
  edgePair = getRandomEdgePair(G)
  if(switchIsAllowed(G, edgePair)){
    G = switch(G, edgePair)
    i = i+1
  }
}
```

(b)

```
for(i in 1:N){
  edgePair = getRandomEdgePair(G)
  if(switchIsAllowed(G, edgePair)){
    G = switch(G, edgePair)
  }
  i = i+1
}
```

FIG. 2. Pseudo-code for two implementations of the switching model. (a) In this implementation, the index i is only incremented when an allowed switch is made. Hence each of the N states G_i differs by a switch from its immediate predecessor G_{i-1}. This implementation corresponds to the Markov chain without repeated states in Example 2.2 (first interpretation) and samples with bias. (b) In this implementation, the index i is incremented regardless of whether an allowed switch is found. Instead of fixing the number of switches, here the number of switching attempts is fixed. This implementation corresponds to the Markov chain with repeated states in Example 2.2 (second interpretation) and hence samples without bias.

for directed networks with degree sequences $(1, 1, 2)$ has transition matrix P

$$P = \begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 & 0 \end{pmatrix}, \quad \lim_{N\to\infty} P^N = \begin{pmatrix} 3/16 & 3/16 & 1/4 & 3/16 & 3/16 \\ 3/16 & 3/16 & 1/4 & 3/16 & 3/16 \\ 3/16 & 3/16 & 1/4 & 3/16 & 3/16 \\ 3/16 & 3/16 & 1/4 & 3/16 & 3/16 \\ 3/16 & 3/16 & 1/4 & 3/16 & 3/16 \end{pmatrix}.$$

By taking the limit of P^N, we find that this Markov chain converges to the distribution $\pi = (3/16, 3/16, 1/4, 3/16, 3/16)$. Another way to arrive at the same conclusion is by using Theorem 2.1 and checking that π satisfies the detailed balance equations. Thus sampling using this Markov chain is not uniform: it is more likely to sample network G_3 than the other networks.

In the second interpretation, e.g. [24], the Markov chain contains repeated states. At each step, a pair of edges of G_i is randomly selected; if switching these edges results in another simple directed network, then the resulting network is the next state, else G_i is repeated. The transition probability from network G_1 to G_2 now equals $1/6$ because only one out of the six edge pairs of G_1 results in network G_2 when rewired. The probability of repeating G_1 corresponds to the probability of selecting an edge pair that results in the same network, i.e. $(3, 1)$ and $(3, 3)$, or that gets rejected, e.g. $(2, 3)$ and $(3, 1)$. For G_1 this probability equals $1/2$. In this case, the transition probabilities are those illustrated in Fig. 1(c). Notice that $p_{ij} = p_{ji}$ for all states i, j. The Markov chain furthermore is irreducible and aperiodic and hence converges to the uniform distribution.

This example shows that for a specific network, the switching model samples without bias when repeated states are included, and with bias when they are excluded. In fact it is well-known that the switching model with repeated states converges to the uniform distribution for *all* directed networks [1, 12].

Our next example shows that the two interpretations from Example 2.2 correspond to fixing the number of switches or the number of attempted switches in the switching model. This alternative point of view was discussed in [9]. However, we think it is important to point out that this is just a different way of describing the same problem.

EXAMPLE 2.3 Figure 2 shows pseudo-code of two distinct implementations of the switching model. The implementations only differ in the placement of a bracket. However, this small difference causes algorithm (a) to sample with bias whereas algorithm (b) samples without bias for directed networks.

The first implementation (Fig. 2(a)) fixes the number of switches N made by the switching method, since the index variable i is only incremented when a switch is made. As such it does not include repeated states and hence results in biased sampling as in the first interpretation discussed in Example 2.2.

In the second implementation (Fig. 2(b)) N switches are *attempted* and i is incremented regardless of whether a switch is accepted or rejected. This corresponds to the inclusion of repeated states and hence to the second interpretation of the switching model discussed in Example 2.2.

We finish this section by proposing a definition of the switching model with respect to directed networks, designed to remove this ambiguity. To do so, we first define a directed switch.

DEFINITION 2.4 Let G_i and G_j be directed networks. There exists a *directed switch* from network $G_i = (V_i, E_i)$ to $G_j = (V_j, E_j)$ if and only if $V_j = V_i$, $E_i \neq E_j$ and there are two edges (x, y) and (u, v) in E_i such that $E_j = (E_i \backslash \{(x, y), (u, v)\}) \cup \{(x, v), (u, y)\}$.

We now define the switching model *with respect to directed networks* as follows.

DEFINITION 2.5 Let G be a finite directed network. The *switching model for G with respect to directed networks* is defined by a Markov chain starting at G. The states of this Markov chain are all directed networks that have the same degree sequences as G. If there exists a directed switch from G_i to G_j then the transition probability p_{ij} equals the probability of randomly selecting an edge pair in G_i that corresponds to this switch.

The resulting transition probabilities can be expressed in terms of the number of edge pairs, $M = \frac{m(m-1)}{2}$, with m the total number of edges of G, and $k(G_i)$ the degree of state G_i in the state graph for G.

$$p_{ij} = \begin{cases} \frac{1}{M} & \text{if there exists a directed switch between } G_i \text{ and } G_j, \\ 1 - \frac{k(G_i)}{M} & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases}.$$

Notice that this definition includes repeated states and leaves no ambiguity about its implementation. Hence it generates unbiased samples for directed networks. In the next section we will show that even though this definition is easily generalized to randomize different network classes, this does not imply that these generalized switching models sample without bias.

3. Randomizing different classes of directed networks

The switching model offers a flexible approach to network randomization. It can easily be altered to sample from different network classes by changing which switches are rejected or accepted. For instance, by rejecting switches that introduce self-loops we can randomize simple directed networks (i.e. directed networks without self-loops).

However, care must be taken when altering the switching model. We show that several changes will introduce bias in the sampling distribution.

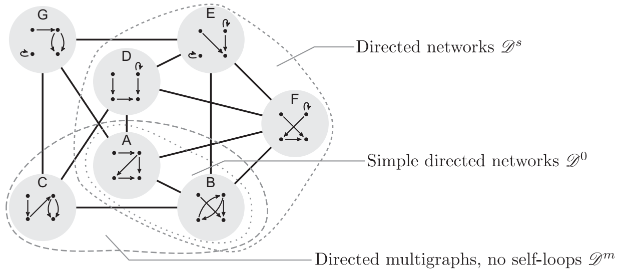

FIG. 3. The state graph of the switching model with respect to $\mathscr{D}$ for network A. Each state has a self-loop, which we have not depicted for visual clarity. The state graphs with respect to classes $\mathscr{D}^0$, $\mathscr{D}^s$ and $\mathscr{D}^m$ are subgraphs as indicated. For instance, the state graph of the switching model with respect to $\mathscr{D}^s$ is the induced subgraph on states A, B and $D - F$.

We start this section by generalizing the definition of the switching model to different classes of directed networks. We then analyse under which conditions the switching model with respect to four common network classes converges to a stationary distribution and, where possible, introduce acceptance probabilities to ensure unbiased sampling.

DEFINITION 3.1 Let $\mathscr{D}$ be the set of all finite directed multigraphs and let $\mathscr{G} \subset \mathscr{D}$ be a subset. Let G be a network in the set $\mathscr{G}$. The *switching model for G with respect to* $\mathscr{G}$ is defined by a Markov chain starting at G. The states of this Markov chain are all the networks in $\mathscr{G}$ that have the same degree sequences as G. Let G_i and G_j be such states, then if there exists a directed switch between G_i and G_j the transition probability p_{ij} equals the probability of selecting an edge pair corresponding to this directed switch.

In this general setting, $\mathscr{G}$ can be any subset of the class of directed multigraphs. We focus on four common classes of networks: simple directed networks ($\mathscr{D}^0$), directed networks, i.e. which may have self-loops, ($\mathscr{D}^s$), and directed multigraphs with and without self-loops ($\mathscr{D}$ and $\mathscr{D}^m$). These network classes are related by the following inclusions: $\mathscr{D}^0 \subset \mathscr{D}^s \subset \mathscr{D}$ and $\mathscr{D}^0 \subset \mathscr{D}^m \subset \mathscr{D}$. The state graphs of these models reflect the same inclusions as illustrated in Fig. 3.

As discussed in Section 2, the switching model with respect to $\mathscr{D}^s$ samples without bias.[1] This was proven using Theorem 2.1.

For certain networks and network classes, the three sufficient conditions in Theorem 2.1 are not satisfied. Figure 4 shows different situations in which one or more of these conditions are no longer true. In these situations the Markov chain either no longer converges to a stationary distribution or converges to a distribution other than the uniform distribution.

Luckily, these problems are either so rare that they can safely be ignored, or they can easily be fixed. We now discuss problems related to each of the three properties in Theorem 2.1 in some more detail.

3.1 *Irreducibility*

The irreducibility for the switching model with respect to $\mathscr{D}^s$ was proven in 1963 by Ryser [25, Theorem 3.1]. Rao *et al.* [12] later gave an alternative proof. Furthermore they showed that for the class of simple directed networks (i.e. $\mathscr{D}^0$), the Markov chain of the switching model is not always irreducible. The

[1] Under mild conditions, see Table 1.

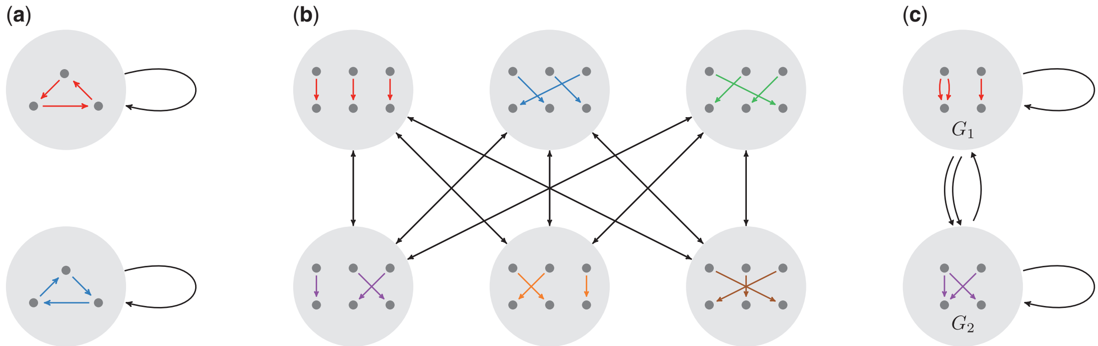

FIG. 4. (a) The switching model with respect to classes $\mathscr{D}^0$ and $\mathscr{D}^m$ is reducible for the directed three cycle. There are only two realizations of this network without self-loops. Since all switches introduce self-loops, the switching model with respect to $\mathscr{D}^0$ and $\mathscr{D}^m$ never leaves the state it starts in and hence is reducible since the state graph is disconnected. For the same network, the switching model with respect to $\mathscr{D}^s$ and $\mathscr{D}$ is irreducible. (b) For any of these networks, the Markov chain of the switching model with respect to all four classes is periodic, with period 2. (c) For multigraphs, transition probabilities are no longer necessarily symmetric. In this example the probability of selecting an edge pair in G_1 that results in G_2 is twice as big as the probability of selecting the reverse switch. This is due to the fact that there are two edges between vertex 1 and vertex 2. In this situation $p_{12} = 2p_{21}$.

networks for which this switching model has a reducible Markov chain were classified in [24]. The problem for these networks is that the orientation of a directed three-cycle cannot be reversed (Fig. 4(a)). The same problem arises for class $\mathscr{D}^m$. There are two solutions to this problem: (1) check if a network has reducible Markov chain [24], and if so, use a pre-sampling step [26], or (2) include *triangle reorientations* [12, hexagonal move; 13, triangle swap; 24, three-cycle reorientation]. On the other hand, the Markov chain of the switching model with respect to $\mathscr{D}$ is always irreducible, as we prove in Lemma 5.1.

3.2 *Aperiodicity*

In practice, the switching model with respect to classes $\mathscr{D}^0$, $\mathscr{D}^s$, $\mathscr{D}^m$ and $\mathscr{D}$ has an aperiodic Markov chain for all networks of interest. Lemma 5.2 proves that for classes $\mathscr{D}^0$ and $\mathscr{D}^m$ it suffices for a network to contain at least one vertex with total degree 2 or more. In the Supplementary data we prove that for classes $\mathscr{D}^s$ and $\mathscr{D}$ it suffices for a network to contain at least one vertex with in-degree 2 or more or out-degree 2 or more.

3.3 *Detailed balance equations*

As illustrated in Fig. 4(c), the transition probabilities of the switching model with respect to multigraphs ($\mathscr{D}^m$ and $\mathscr{D}$) are no longer symmetric (i.e. $p_{ij} \neq p_{ji}$). In the Supplementary data we derive both the transition probabilities and the stationary distribution for the switching model with respect to classes $\mathscr{D}^m$ and $\mathscr{D}$.

The switching model with respect to these classes produces biased samples. However, in Section 5 we show that these switching models can be forced to sample without bias by introducing acceptance probabilities [22, 27].

Table 1 summarizes the properties of the switching model with respect to the four network classes discussed in this section. Under mild conditions, and for some classes after a small adjustment, the switching model can be used to produce unbiased samples for all four classes: $\mathscr{D}^0$, $\mathscr{D}^s$, $\mathscr{D}^m$ and $\mathscr{D}$.

TABLE 1 *Properties of the Markov chains corresponding to switching methods with respect to four classes of directed networks.*

Network classes			Properties of the switching model			
Class	Multiple	Self-loops	Irreducible	Aperiodic	$p_{ij} = p_{ji}$	Uniform
$\mathscr{D}^0$	No	No	No [12, 24]	Yes†	Yes	Yes‡
$\mathscr{D}^s$	No	Yes	Yes [12, 24, 25]	Yes§	Yes	Yes
$\mathscr{D}^m$	Yes	No	No	Yes†	No	Yes‡,¶
$\mathscr{D}$	Yes	Yes	Yes	Yes§	No	Yes¶

†If the network contains at least one vertex with total degree two or more.
‡For most networks this Markov chain is irreducible [24], if not the bias in sampling can be removed by using a pre-sampling step[26].
§If the network contains at least one vertex with in-degree two or more or out-degree two or more.
¶If acceptance probabilities are used.

4. Randomizing undirected networks

The switching model has also been used to randomize undirected networks [15]. The randomization of undirected networks has its own subtleties, and again, the switching model needs to be implemented with care to avoid biased sampling. We point out that MFinder, the popular motif finding software accompanying [15, 16], up until recently[2] produced biased samples of undirected networks. Results obtained using MFinder for undirected networks should hence be checked for correctness.

To formulate the switching model with respect to undirected networks we first define an undirected switch.

DEFINITION 4.1 There exists a *switch* from the undirected network $G_i = (V_i, E_i)$ to the undirected network $G_j = (V_j, E_j)$ if and only if $V_j = V_i$, $E_i \neq E_j$ and there are two edges $\{x, y\}$ and $\{u, v\}$ in E_i such that either $E_j = (E_i \backslash \{\{x, y\}, \{u, v\}\}) \cup \{\{x, v\}, \{u, y\}\}$ (*switch 1*) or $E_j = (E_i \backslash \{\{x, y\}, \{u, v\}\}) \cup \{\{x, u\}, \{y, v\}\}$ (*switch 2*) (see Fig. 5(a)).

In a sense, this definition of a switch is overspecified. We could restrict to switch 1 only, since switch 2 is just switch 1 where one edge is labelled in reverse order. However, the reason that we *do mention* both switches is that it is important to realize that selecting an edge pair does not correspond to selecting a switch, there are *two* potential switches for each pair of edges. This is important for two reasons. Firstly, when implementing the switching algorithm for undirected networks, it is most likely that edges are stored with vertices in fixed order. Example 4.3 shows how implementing just switch 1 results in biased sampling. In fact, this is exactly what caused MFinder to produce biased samples of undirected networks. Secondly, to derive the transition probabilities for the Markov chains, we need to take into account that switch 1 and switch 2 may result in different networks.

We now define the switching model with respect to any subset of undirected multigraphs.

[2] After we contacted the authors of MFinder the software was fixed and updated in May 2015.

TABLE 2 *Properties of the Markov chains corresponding to switching methods with respect to four classes of undirected networks*

Network classes			Properties of the switching model			
Class	Multiple	Self-loops	Irreducible	Aperiodic	$p_{ij} = p_{ji}$	Uniform
$\mathscr{U}^0$	No	No	Yes [14, 28, 29]	Yes	Yes	Yes
$\mathscr{U}^s$	No	Yes	No	Yes	No	No†
$\mathscr{U}^m$	Yes	No	Yes [30, 31]	Yes	No	Yes‡
$\mathscr{U}$	Yes	Yes	Yes [30]	Yes	No	Yes‡

†We suspect that this Markov chain is only reducible for a small class of networks. Furthermore we believe that a triangle move may be introduced to ensure irreducibility; however we have no proof of this (see Supplementary data for more details).
‡If acceptance probabilities are used.

DEFINITION 4.2 Let $\mathscr{U}$ be the set of all finite undirected multigraphs and let $\mathscr{G} \subset \mathscr{U}$ be a subset. Let G be a network in the set $\mathscr{G}$. The *switching model for G with respect to* $\mathscr{G}$ is defined by a Markov chain starting at G. The states of this Markov chain are all the networks in $\mathscr{G}$ that have the same degree sequence as G. If G_i, G_j are such networks and there exists a switch between G_i and G_j then the transition probability p_{ij} is the probability of selecting an edge pair and picking switch 1 or switch 2 (with equal probability), corresponding to this switch.

We have summarized the properties of the switching model with respect to simple undirected networks ($\mathscr{U}^0$), undirected networks ($\mathscr{U}^s$), undirected multigraphs without self-loops ($\mathscr{U}^m$) and undirected multigraphs ($\mathscr{U}$) in Table 2. Lemmas 5.3 and 5.4 show that the switching model with respect to $\mathscr{U}^m$ converges to a distribution different from the uniform distribution. Section 5 furthermore shows how acceptance probabilities [13, 22] may be introduced to enforce unbiased sampling. The idea of acceptance probabilities is simple: instead of accepting all allowed switches, accept each switch with a given probability. If chosen well, this ensures $p_{ij} = p_{ji}$ for all i, j and hence when the Markov chain is irreducible and aperiodic, it converges to the uniform distribution. The Supplementary data contain proofs for all statements without citations, as well as the derivations of the stationary distribution for classes $\mathscr{U}^0$ and $\mathscr{U}$.

The example below illustrates the impact of neglecting to implement both switches 1 and 2 for the switching model with respect to undirected networks.

EXAMPLE 4.3 Let G_1 be the undirected network with six vertices $v_1, \ldots, v_6$ and six edges $\{v_1, v_3\}$, $\{v_2, v_4\}$, $\{v_3, v_4\}$, $\{v_3, v_5\}$, $\{v_4, v_6\}$ and $\{v_5, v_6\}$ as illustrated in Fig. 5(b). There are 17 simple undirected networks with the same degree sequence as G_1 (see Fig. 1 Supplementary data).

We first look at a worst-case scenario, where the edges are stored in such a way that switch 1 is not allowed for any of the edge pairs: (v_1, v_3), (v_4, v_2), (v_4, v_3), (v_4, v_6), (v_5, v_3), (v_5, v_6) (Fig. 5(c)). For any edge pair, applying switch 1 either does not alter the network (when the edges have the same source or the same target) or introduces a self-loop or a multiple edge. Any 'random' network generated by this implementation will be the network G_1 itself.

We now show that even when this implementation of the switching model is able to sample from the full set of 17 networks, the sample may still be biased. This is the case when the edge list is stored

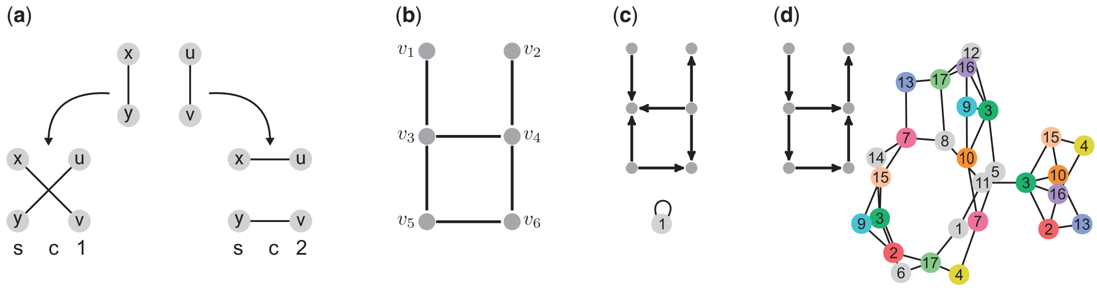

FIG. 5. (a) Illustration of switch 1 and switch 2 corresponding to the edge pair $\{x, y\}$ and $\{u, v\}$. (b) An undirected network G_1. (c) A directed network corresponding to storing G_1 with vertex order $\{(v_1, v_3), (v_4, v_2), (v_4, v_3), (v_4, v_6), (v_5, v_3), (v_5, v_6)\}$ and the state graph of the switching model with just switch 1 implemented. (d) A directed network corresponding to storing G_1 with vertex order $\{(v_1, v_3), (v_4, v_2), (v_3, v_4), (v_6, v_4), (v_3, v_5), (v_5, v_6)\}$ and the state graph of the switching model with just switch 1 implemented. In this state graph, many of the 17 undirected networks correspond to several states as indicated by labels. Each state also has a self-loop, which we have not depicted for visual clarity.

as $(v_1, v_3), (v_4, v_2), (v_3, v_4), (v_6, v_4), (v_3, v_5), (v_5, v_6)$. Several networks appear multiple times in the state graph corresponding to this Markov chain (Fig. 5(d)). The reason for this is that these networks are not uniquely represented by an ordered edge list. For instance the following two edge lists with distinct vertex ordering, correspond to a single undirected network (G_2 in Fig. 1 Supplementary data).

$$(v_1, v_6), (v_4, v_5), (v_5, v_3), (v_3, v_4), (v_3, v_2), (v_6, v_4), \tag{4.1}$$

$$(v_1, v_6), (v_4, v_3), (v_5, v_4), (v_3, v_5), (v_3, v_2), (v_6, v_4). \tag{4.2}$$

In this situation, all non-zero transition probabilities between distinct states are equal to $1/15$, the probability of selecting the edge pair corresponding to the switch (switch 1) between them, and hence $p_{ij} = p_{ji}$ for all i and j. The state graph is also irreducible and aperiodic and hence this Markov chain converges to the uniform distribution. This implies that each undirected networks is sampled proportionally to the number of times it appears in the state graph, and hence sampling is biased. For instance, the probability of sampling state G_1 is three times less than the probability of sampling state G_3 (see Fig. 5(d) and Fig. 1 Supplementary data).

5. Selected proofs

In this section we present mathematical proofs of some of the properties of the switching models discussed in Sections 2–4.

Our first proof uses the symmetric edge set difference of graphs [32]. Let $G = (V, E)$ and $G' = (V', E')$ be two graphs. The *symmetric edge set difference* of G and G' is defined as $E \Delta E' := (E \setminus E') \cup (E' \setminus E)$. We write $G \Delta G'$ for the graph with vertices $V \cup V'$ and edge set $E \Delta E'$.

The symmetric edge set difference of two simple directed graphs with equal degree sequences has some nice 'Eulerian' properties [28]. Here we use the property that it has even cardinality (see Corollary 2 in the Supplementary data).

LEMMA 5.1 Let $G = (V, E) \in \mathscr{D}$. The Markov chain of the switching model for G with respect to $\mathscr{D}$ is irreducible.

Proof. Irreducibility of a Markov chain is equivalent to the corresponding state graph being strongly connected. Since the directed switch is a symmetric move, it is enough to show that the state graph of the switching model with respect to $\mathscr{D}$ is connected. That is, for any network $G' = (V, E') \in \mathscr{D}$ with degree sequences equal to those of G, we need to show that there exists a sequence of graphs $G = G_0, \ldots, G_k = G'$ such that $G_i \in \mathscr{D}$ and there exists a directed switch between G_i and G_{i+1} for all $i \in (0, \ldots, k-1)$. This can be proven in terms of the edge set difference by showing $|E_i \Delta E_{i+1}| = 4$ for all $i \in (0, \ldots, k-1)$ (see [28]). The edge set difference $E\Delta E' = E\backslash E' \cup E'\backslash E$ has even cardinality, $|E\Delta E'| = 2\kappa$.

We prove this lemma by induction on κ. Let $\kappa > 2$ and let $(v_1, v_2) \in E\backslash E'$, then there exist edges (v_1, v_3) and (v_4, v_2) in $E\backslash E'$ with $v_3 \neq v_2$ and $v_4 \neq v_1$ since G and G' have equal degree sequences. Let G^* be the graph with vertices V and edges $E^* = (E'\backslash\{(v_1, v_3), (v_4, v_2)\}) \cup \{(v_1, v_2), (v_4, v_3)\}$ then $G^* \in \mathscr{D}$, $|E^*\Delta E'| = 4$ and $|E^*\Delta E| \leq 2\kappa - 2$. □

LEMMA 5.2 Let $G \in \mathscr{D}^0$ or $\mathscr{D}^m$. The Markov chain of the switching model for G with respect to $\mathscr{D}^0$ or to $\mathscr{D}^m$ is aperiodic if and only if G contains a vertex v with total degree at least 2.

Proof. We first show that the Markov chain is trivially aperiodic when G contains a vertex v of total degree at least 2, by showing that each state has a non-zero probability of being repeated (i.e. $p_{ii} > 0$ for all i). Let G_i be a state, then v has degree at least 2, since all states have the same degree sequence as G. If v has an incoming and an outgoing edge then switching these edges is not allowed, since it would result in a network that contains a self-loop at v. Hence $p_{ii} > 0$. If instead, v has either two incoming or two outgoing edges, then switching these edge does not change the network and hence $p_{ii} > 0$.

To prove the reverse claim, we use proof by contrapositive: we show that if G does not contain a vertex v with total degree at least 2, then the Markov chain of the switching model with respect to $\mathscr{D}^0$ and $\mathscr{D}^m$ is periodic.

If G does not contain a vertex v with total degree at least 2 then G is a disjoint union of single vertices and single edges. We ignore the single vertices since the switching model leaves these invariant. Thus the interesting part of G is a collection of single edges $\{(s_1, t_1), (s_2, t_2), \ldots, (s_m, t_m)\}$. We can represent G as the ordered tuple $T = (t_1, t_2, \ldots, t_m)$. The set of simple directed networks with the same degree sequences corresponds to all permutations of T. For instance, if $(t_{i_1}, t_{i_2}, \ldots, t_{i_m})$ is a permutation of T, then the network with edge set $\{(s_1, t_{i_1}), (s_2, t_{i_2}), \ldots, (s_m, t_{i_m})\}$ is simple directed and has the same degree sequence as G.

A directed switch corresponds to a transposition of two elements in $(t_1, t_2, \ldots, t_m)$. The identity is an even permutation and can thus only be obtained as the composition of an even number of transpositions. This precisely means that any sequence of networks in the Markov chain starting and ending at a network G_i has to be of even length. The chain is periodic with period 2 (Fig. 4(b) illustrates the $m = 3$ case). □

We next derive the stationary distribution for the switching model with respect to $\mathscr{U}^m$. To do so, we first derive the transition probabilities for this switching model.

LEMMA 5.3 Let G_i and G_j be two networks in $\mathscr{U}^m$ with equal degree sequence. The transition probability p_{ij} from G_i to G_j is given by

$$p_{ij} = \begin{cases} \frac{A^i_{xy}A^i_{uv}}{2M} & \text{if } E_j = E_i \backslash \{\{x,y\},\{u,v\}\} \cup \{\{x,v\},\{u,y\}\}, \\ & \text{or } E_j = E_i \backslash \{\{x,y\},\{u,v\}\} \cup \{\{x,u\},\{y,v\}\}, \\ 1 - \sum_{k,k\neq i} p_{ik} & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases}$$

where $[A^i_{kl}]$ is the weighted adjacency matrix of the network G_i and M is the number of edge pairs $m(m-1)/2$ in G_i.

Proof. If there exists a switch between G_i and G_j, then without loss of generality we may assume that $E_j = E_i \backslash \{\{x,y\},\{u,v\}\} \cup \{\{x,v\},\{u,y\}\}$. Notice that all four vertices are distinct since neither network contains self-loops and the networks are different. The transition probability p_{ij} is the probability of selecting the edge pair $\{x,y\}$ and $\{u,v\}$ in G_i and selecting switch 1. It is not hard to see that there are $A^i_{xy}A^i_{uv}$ distinct edge pairs $\{x,y\}$ and $\{u,v\}$. Hence $p_{ij} = A^i_{xy}A^i_{uv}/2M$. □

LEMMA 5.4 Let $G \in \mathscr{U}^m$ be an undirected multigraph without self-loops. For any $G_i \in \mathscr{U}^m$ with the same degree sequence as G, let

$$\beta_i = \frac{1}{\prod_{k<l} A^i_{kl}!},$$

where $[A^i_{kl}]$ is the weighted adjacency matrix corresponding to G_i. Then $\pi_i = \beta_i/\sum_r \beta_r$ is the stationary distribution for the switching model for G with respect to $\mathscr{U}^m$.

Proof. Using the transition probabilities from Lemma 5.3 we show that π_i satisfies the detailed balance equations. Let $G_i = (V, E_i) \in \mathscr{U}^m$ and $G_j = (V, E_j) \in \mathscr{U}^m$ with the same degree sequence as G. Without loss of generality we assume that $E_j = E_i \backslash \{\{x,y\},\{u,v\}\} \cup \{\{x,v\},\{u,y\}\}$. Define $A^i_{\{rs\}}$ to be A^i_{rs}. Hence $A^i_{\{rs\}}$ is also equal to A^i_{sr}, since the network is undirected. Let,

$$K_i = \frac{1}{2M} \frac{1}{\sum_r \beta_r} \frac{1}{\prod_{k<l,\{k,l\}\notin E_i \Delta E_j} A^i_{\{kl\}}!}$$

and notice that K_i equals K_j. Then,

$$\begin{aligned} \pi_i p_{ij} &= \frac{\beta_i}{\sum_r \beta_r} \frac{A^i_{\{xy\}}A^i_{\{uv\}}}{2M} \\ &= K_i \frac{A^i_{\{xy\}}A^i_{\{uv\}}}{A^i_{\{xy\}}!A^i_{\{uv\}}!A^i_{\{xv\}}!A^i_{\{uy\}}!} \\ &= K_i \frac{1}{(A^i_{\{xy\}}-1)!(A^i_{\{uv\}}-1)!A^i_{\{xv\}}!A^i_{\{uy\}}!} \end{aligned}$$

$$= K_i \frac{1}{A^j_{\{xy\}}!A^j_{\{uv\}}!(A^j_{\{xv\}}-1)!(A^j_{\{uy\}}-1)!}$$
$$= K_j \frac{A^j_{\{xv\}}A^j_{\{uy\}}}{A^j_{\{xy\}}!A^j_{\{uv\}}!A^j_{\{xv\}}!A^j_{\{uy\}}!}$$
$$= \pi_j p_{ji}. \qquad \square$$

We can introduce *acceptance probabilities* to force the switching model with respect to $\mathscr{U}^m$ to converge to the uniform distribution. Let the *adjusted switching model with respect to* $\mathscr{U}^m$ have acceptance probabilities $a_{ij} = 1/A^i_{xy}A^i_{uv}$ when G_i and G_j differ by $E_j = E_i \backslash \{(x,y),(u,v)\} \cup \{(x,v),(u,y)\}$ and $a_{ij} = 1$ otherwise. These acceptance probabilities ensure $p_{ij} = p_{ji}$ for all states i and j and hence the adjusted switching model with respect to $\mathscr{U}^m$ samples without bias. In the Supplementary data we introduce the adjusted switching model with respect to classes $\mathscr{D}^m$, $\mathscr{D}$ and $\mathscr{U}$.

6. Conclusion

In this paper we propose a precise definition of the switching model with respect to any subset of directed or undirected multigraphs. This definition is intended to guide practitioners to correctly implement the switching model. Furthermore, we show that it is necessary to define the switching model with respect to a subset, since its properties differ depending on the subset.

We demonstrated that in implementing the switching model much can potentially go wrong and cause sampling to be biased. However, by using Theorem 2.1, most of these issues can be resolved.

For eight common network classes, we analysed under which conditions the switching model converges to a stationary distribution and derived this distribution. When necessary, we introduced acceptance probabilities, such that the resulting adjusted switching model converges to the uniform distribution.

By carefully analysing the properties of the switching model with respect to undirected networks, we found and resolved an error in the well-known software package MFinder.

One of the strengths of the switching model is its simplicity: it is a simple procedure of edge swaps that randomizes networks while fixing their degree sequence(s). When treated carefully, the switching model can be used to draw uniform samples from a variety of network classes. This makes the switching model an attractive candidate for a null-model.

As already noted, in this paper we have ignored one crucial question about switching models: how many switches do we need to attempt, in order to obtain a truly random network? Or, in other words, how many steps does the corresponding Markov chain need to take to reach its stationary distribution. This question is generally hard to answer [21, 33].

A recently proposed approach, the Expand and Contract method [34], may provide a solution by combining the fast but less flexible configuration model (which samples from $\mathscr{D}$ and $\mathscr{U}$) with the (adjusted) switching model with respect to $\mathscr{D}$ and $\mathscr{U}$. This would be an interesting direction for further research.

Funding

This work was supported by the Australian Department of Defence under Research Agreement 4500743680.

Supplementary data

Supplementary data are available at *COMNET* online.

References

1. Artzy-Randrup, Y. & Stone, L. (2005) Generating uniformly distributed random networks. *Phys. Rev. E*, **72**, 056708.
2. Carstens, C. J. (2014) A uniform random graph model for directed acyclic networks and its effect on motif-finding. *J. Complex Netw.*, **2**, 419–430.
3. Carstens, C. J. (2015) Proof of uniform sampling of binary matrices with fixed row sums and column sums for the fast curveball algorithm. *Phys. Rev. E*, **91**, 042812.
4. Molloy, M. & Reed, B. (1995) A critical point for random graphs with a given degree sequence. *Random Struct. Algorithms*, **6**, 161–180.
5. Newman, M. E. J., Strogatz, S. H. & Watts, D. J. (2001) Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, **64**, 026118.
6. Snijders, T. A. B. (1991) Enumeration and simulation methods for 0âŁ"1 matrices with given marginals. *Psychometrika*, **56**, 397–417.
7. King, O. D. (2004) Comment on "subgraphs in random networks". *Phys. Rev. E*, **70**, 058101.
8. Klein-Hennig, H. & Hartmann, A. K. (2012) Bias in generation of random graphs. *Phys. Rev. E*, **85**, 026101.
9. Miklós, I. & Podani, J. (2004) Randomization of presence-absence matrices: comments and new algorithms. *Ecology*, **85**, 86–92.
10. Milo, R., Kashtan, N., Itzkovitz, S., Newman, M. E. J. & Alon, U. (2003) On the uniform generation of random graphs with prescribed degree sequences. arXiv preprint arXiv:cond-mat/0312028.
11. Maslov, S. & Sneppen, K. (2002) Specificity and stability in topology of protein networks. *Science*, **296**, 910–913.
12. Rao, A. R., Jana, R. & Bandyopadhyay, S. (1996) A Markov chain Monte Carlo method for generating random (0, 1)-matrices with given marginals. *Sankhya Indian J. Stat. Ser. A*, **58**, 225–242.
13. Roberts, E. S. & Coolen, A. C. C. (2012) Unbiased degree-preserving randomization of directed binary networks. *Phys. Rev. E*, **85**, 046103.
14. Taylor, R. (1981) Constrained switchings in graphs. *Combinatorial Mathematics VIII*. Berlin: Springer, pp. 314–336.
15. Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S. Ayzenshtat, I., Sheffer, M. & Alon, U. (2004) Superfamilies of evolved and designed networks. *Science*, **303**, 1538–1542.
16. Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D. & Alon, U. (2002) Network motifs: simple building blocks of complex networks. *Science*, **298**, 824–827.
17. Greenhill, C. (2011) A polynomial bound on the mixing time of a Markov chain for sampling regular directed graphs. *Electron. J. Combin.*, **18**, P234.
18. Kannan, R., Tetali, P. & Vempala, S. (1999) Simple Markov-chain algorithms for generating bipartite graphs and tournaments. *Random Struct. Algorithms*, **14**, 293-308.
19. Miklós, I., Erdös, P. L. & Soukup, L. (2013) Towards random uniform sampling of bipartite graphs with given degree sequence. *Electron. J. Combin.*, **20**, P16.
20. Rechner, S. & Berger, A. (2016) *Marathon*: an open source software library for the analysis of markov-chain monte carlo algorithms. *PLoS One*, **11**, e0147935.
21. Ray, J., Pinar, A. & Seshadhri, C. (2014) A stopping criterion for Markov chains when generating independent random graphs. *J. Complex Netw.*, **3**, 204–220.
22. Coolen, A. C. C., De Martino, A. & Annibale, A. (2009) Constrained Markovian dynamics of random graphs. *J. Stat. Phys.*, **136**, 1035–1067.
23. Mitzenmacher, M. & Upfal, E. (2005) *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. New York: Cambridge University Press.
24. Berger, A. & Müller-Hannemann, M. (2010) Uniform sampling of digraphs with a fixed degree sequence. *Graph Theoretic Concepts in Computer Science*. Lecture Notes in Computer Science. Berlin: Springer, pp. 220–231.

25. Ryser, H. J. (1963) *Combinatorial mathematics*. Carus Mathematical Monographs. The Mathematical Association of America, New York, NY: Wiley.

26. Berger, A. (2011) Directed degree sequences. *Ph.D. Thesis*, Martin-Luther University Halle-Wittenberg. Retrieved from http://wcms.itz.uni-halle.de/download.php?down=22851&elem=1624638 (accessed on September 21, 2016).

27. Roberts, E. S., Annibale, A. & Coolen, A. C. C. (2014) Controlled Markovian dynamics of graphs: unbiased generation of random graphs with prescribed topological properties. *Nonlinear Maps and Their Applications*. New York: Springer, pp. 25–34.

28. Berger, A. & Müller-Hannemann, M. (2009) Uniform sampling of undirected and directed graphs with a fixed degree sequence. arXiv preprint arXiv:0912.0685.

29. Eggleton, R. B. & Holton, D. A. (1981) Simple and multigraphic realizations of degree sequences. *Combinatorial Mathematics VIII* (K. L. McAvaney ed.). Berlin: Springer, pp. 155–172.

30. Eggleton, R. B. & Holton, D. A. (1979) The graph of type $(0,\infty,\infty)$ realizations of a graphic sequence. *Combinatorial Mathematics VI* (C. Grácio, D. Fournier-Prunaret, T. Ueta & Y. Nishio eds), New York: Springer, pp. 41–54.

31. Hakimi, S. (1962) On realizability of a set of integers as degrees of the vertices of a linear graph. I. *J. Soc. Ind. Appl. Math.*, **10**, 496–506.

32. Bunke, H., Dickinson, P. J., Kraetzl, M. & Wallis, W. D. (2007) *A Graph-Theoretic Approach to Enterprise Network Dynamics*. Basel: Birkhauser.

33. Ray, J., Pinar, A. & Seshadhri, C. (2012) Are we there yet? when to stop a Markov chain while generating random graphs. *Algorithms and Models for the Web Graph* (A. Bonato & J. Janssen eds), vol. 7323. Lecture Notes in Computer Science. Berlin: Springer, pp. 153–164.

34. Zhao, J. (2013) Expand and contract: sampling graphs with given degrees and other combinatorial families. arXiv preprint arXiv:1308.6627.