Viral Marketing On Configuration Model

Bartlomiej Blaszczyszyn, Kumar Gaurav

Viral Marketing On Configuration Model

Bartłomiej Błaszczyszyn*and Kumar Gaurav†

September 23, 2013

Abstract

We consider propagation of influence on a Configuration Model, where each vertex can be influenced by any of its neighbours but in its turn, it can only influence a random subset of its neighbours. Our (enhanced) model is described by the total degree of the typical vertex, representing the total number of its neighbours and the transmitter degree, representing the number of neighbours it is able to influence. We give a condition involving the joint distribution of these two degrees, which if satisfied would allow with high probability the influence to reach a non-negligible fraction of the vertices, called a *big (influenced) component*, provided that the source vertex is chosen from a set of *good pioneers*. We show that asymptotically the big component is essentially the same, regardless of the good pioneer we choose, and we explicitly evaluate the asymptotic relative size of this component. Finally, under some additional technical assumption we calculate the relative size of the set of good pioneers. The main technical tool employed is the "fluid limit" analysis of the joint exploration of the configuration model and the propagation of the influence up to the time when a big influenced component is completed. This method was introduced in Janson & Luczak (2008) to study the giant component of the configuration model. Using this approach we study also a reverse dynamic, which traces all the possible sources of influence of a given vertex, and which by a new "duality" relation allows to characterise the set of good pioneers.

Keywords: *enhanced Configuration Model, influence propagation, backtracking, duality, big component*

*Inria/ENS, 23 av. d'Italie 75214 Paris, France; Bartek.Blaszczyszyn@ens.fr
†UPMC/Inria, 23 av. d'Italie 75214 Paris, France; Kumar.Gaurav@inria.fr

1 Introduction

The desire for understanding the mechanics of complex networks [1, 13], describing a wide range of systems in nature and society, motivated many applied and theoretical investigations of the last two decades. A motivation for our work can come from the phenomenon of viral marketing in social networks: A person after getting acquainted with an advertisement (or a news article or a Gangnam style video, for that matter) through one of his “friends”, may decide to share it with some (not necessarily all) of his friends, who will, in turn, pass it along to some of their friends, and so on. The campaign is successful if starting from a relatively small number of initially targeted persons, the influence (or information) can spread as an epidemic “infecting” a non-negligible fraction of the population.

Enhanced Configuration Model Traditionally, social networks have been modeled as random graphs [8, 14], where the vertices denote the individuals and edges connect individuals who know one another. The Configuration Model is considered as a useful approximation in this matter, and we assume it for our study of the viral marketing. It is a random (multi-)graph, whose vertices have prescribed degrees, realized by half-edges emanating from them and uniformly pair-wise matched to each other to create edges. In order to model a selective character of the influence propagation (each vertex can be influenced by any of its neighbours but in its turn, it can only influence a subset of its neighbours), we enhance the original Configuration Model by considering *two types of half-edges*. *Transmitter half edges* of a given vertex represent links through which this vertex will influence (pass the information once it has it) to its neighbours. Its *receiver half-edges* represent links through which this vertex will not propagate the information to its neighbours. The neighbours receive the information both through their transmitter and receiver half-edges matched to a transmitter half edge of the information sender. The two types of half-edges are not distinguished during the uniform pair-wise matching of all half-edges, but only to trace the propagation of information. Assuming the usual consistency conditions for the numbers of transmitter and receiver half-edges, the Enhanced Configuration Model is asymptotically (when the number of vertices n goes to infinity) described by the vector of two, not necessarily independent, integer valued random variables, representing the *transmitter* and *receiver degree* of the typical vertex. Equivalently, we can consider the *total vertex degree*, representing the total number of friends of a person and its transmitter degree, representing the number of friends he/she is able to influence.

Results We consider the advertisement campaign started from some initial target (source vertex) and following the aforementioned dynamic on a realization of the Enhanced Configuration Model of the total number of vertices n. The results are formulated with high probability (whp), i.e. with probability approaching one as $n \to \infty$.

First, we give a condition involving the total degree and the transmitter degree distributions of the Enhanced Configuration Model, which if satisfied, would allow whp the advertisement campaign to reach a non-negligible fraction ($O(n)$) of the population, called a *big (influenced) component*, provided that the initial target is chosen from a set of *good pioneers*. Further in this case, we show that asymptotically the *big component is essentially the same* regardless of the good pioneer chosen, and we explicitly evaluate the asymptotic size of this component relative to n. The essential uniqueness of the big component means that the subsets of influenced vertices reached from two different good pioneers differ by at most $o(n)$ vertices whp. Finally, under some additional technical assumption we calculate the relative size of the set of good pioneers.

Methodology A standard technique for the analysis of diffusion of information on the Configuration Model consists in simultaneous exploration of the model and the propagation of the influence. We adopt this technique and, more precisely, the approach proposed in [11] for the study of the giant component of the (classical) Configuration Model. In this approach, instead of the branching process approximating the early stages of the graph exploration, one uses a "fluid limit" analysis of the process up to the time when the exploration of the big component is completed. We tailor this method to our specific dynamic of influence propagation and calculate the relative size of the big influenced component, as well as prove its essential uniqueness.

A fundamental difference with respect to the study of the giant component of the classical model stems from the directional character of our propagation dynamic. Precisely, the edges matching a transmitter and a receiver half-edge can relay the influence from the transmitter half-edge to the receiver one, but not the other way around. This means that the good pioneers do not need to belong to the big (influenced) component, and vice versa. In this context, we introduce a *reverse dynamic*, in which a message (think of an "acknowledgement") can be sent in the reversed direction on every edge (from an arbitrary half-edge to the receiver one), which traces all the possible sources of influence of a given vertex. This reversed dynamic can be studied using the same approach as the original one. In particular, one can establish the essential uniqueness of the big component of the reversed process as well as calculate its relative size. Interestingly, this relative size coincides with the probability of the non-extinction of the branching process approximating the initial phase of the original exploration process, whence the hypothesis that the big component of the reverse process coincides with the set of good pioneers. We prove this conjecture under some additional (technical) assumption. We believe the method of introducing a reverse process to derive results for the original one has not been seen in a related context in the existing literature.

Related Work The propagation of influence through a network has been previously studied in various contexts. The Configuration Model has formed the base for an increasing number of influence propagation studies, of which one relevant to the phenomenon of viral networking in social networks is discussed in [2] and [12], where a vertex in the network gets influenced only if a certain proportion of its neighbours have already been influenced. This interesting propagation dynamic is further studied by introducing *cliques* in Configuration Model to observe the impact of clustering on the size of the population influenced (see [5],[6]). This dynamic is a kind of *pull model* where influence propagation depends on whether a vertex decides to receive the influence from its neighbours. We study a *push model*, where the influence propagation depends on whether a vertex decides to transmit the influence. A propagation dynamic where every influenced node, at all times, keeps choosing one of its neighbours uniformly at random and transmits the message to it is studied on a d-regular graph in [9]. This dynamic is close in its spirit to the one we considered in this paper, however the process stops when *all* nodes receive the message, and this stopping time is studied in the paper. The same dynamic but restricted to some (possibly random) maximal number of transmissions allowed for each vertex is considered in [4] on a complete graph. This can be thought as a special case of our dynamic (although we study it on a different underlying graph) if we assume that the transmitter and receiver degrees correspond to the number of collected and non-collected coupons, respectively, in the classical coupon collector problem with the number of coupons being the vertex degree and the number of trials being the number of allowed transmissions. In a more applied context, a rudimentary special case of our dynamic of influence propagation has actually been studied on real-world networks like *flixster* and *flickr* (see [10]).

Paper organization The remaining part of this paper is organized as follows. In the next section we describe our model and formulate the results. In Sections 3 and 4 we analyze, respectively, the original and reversed dynamic of influence propagation. The relations between the two dynamics are explored in Section 5.

2 Notation and Results

Given a degree sequence $(d_i^{(n)})_1^n$ for n vertices labelled 1 to n, Configuration Model, denoted $G^*(n,(d_i)_1^n)$, is a random multigraph obtained by giving d_i half-edges to each vertex i and then uniformly matching pair-wise the set of half-edges. Conditioning the Configuration Model to be simple, we obtain a uniform random graph with the given degree sequence, denoted by $G(n,(d_i)_1^n)$. Since it is convenient to work with the Configuration Model, we will prove all our results for the Configuration Model and the corresponding results for the uniform random graph can be obtained by passing through a standard conditioning procedure (see, for example, [14]).

Further, in our model, we represent the degree, d_i, of each vertex i as the sum of two (not necessarily independent) degrees: transmitter degree, $d_i^{(t)}$ and receiver degree, $d_i^{(r)}$.

We will asssume the following set of consistency conditions for our enhanced Configuration Model, which are analogous to those assumed for Configuration Model in [11].

Condition 2.1. *For each n, $\mathbf{d}^{(n)} = (d_i)_1^n$,is a sequence of non-negative integers such that $\sum_{i=1}^n d_i := 2m$ is even and for each i, $d_i = d_i^{(r)} + d_i^{(t)}$. For $k \in \mathbb{N}$, let $u_{k,l} = |\{i : d_i^{(r)} = k, d_i^{(t)} = l\}|$, and $D_n^{(r)}$ and $D_n^{(t)}$ be the receiver and transmitter degrees respectively of a uniformly chosen vertex in our model, i.e., $\mathbb{P}(D_n^{(r)} = k, D_n^{(t)} = l) = u_{k,l}/n$. Let $D^{(r)}$ and $D^{(t)}$ be two random variables taking value in non-negative integers with joint probability distribution $(p_{v,w})_{(v,w)\in\mathbb{N}^2}$, and $D := D^{(r)} + D^{(t)}$. Then the following hold.*

(i) $\frac{u_{k,l}}{n} \to p_{k,l}$ *for all* $(k,l) \in \mathbb{N}^2$.

(ii) $\mathbb{E}[D] = \mathbb{E}[D^{(r)} + D^{(t)}] = \sum_{k,l}(k+l)p_{k,l} \in (0,\infty)$. *Let* $\lambda_r = \mathbb{E}[D^{(r)}]$, $\lambda_t = \mathbb{E}[D^{(t)}]$ *and* $\lambda = \lambda_r + \lambda_t$.

(iii) $\sum_{i=1}^n (d_i)^2 = O(n)$.

(iv) $\mathbb{P}(D = 1) > 0$.

Let $g(x,y) := \mathbb{E}[x^{D^{(r)}} y^{D^{(t)}}]$ be the joint probability generating function of $(p_{v,w})_{(v,w)\in\mathbb{N}^2}$. Further let

$$h(x) := x \left.\frac{\partial g(x,y)}{\partial y}\right|_{y=x} = \mathbb{E}[D^{(t)} x^D], \tag{1}$$

and

$$H(x) := \lambda x^2 - \lambda_r x - h(x). \tag{2}$$

If two neighbouring vertices x and y are connected via the pairing of a transmitter half-edge of x with any half-edge of y, then x has the ability to directly influence y. More generally, for any two vertices x and y in the graph and $k \geq 1$, if there exists a set of vertices $x_0 = x, x_1, \ldots.., x_{k-1}, x_k = y$ such that $\forall i : 1 \leq i \leq k$, x_{i-1} has the ability to directly influence x_i, we say that x has the ability to influence y and denote it by $x \to y$; in other words, y can be influenced starting from the initial source x. Let $C(x)$ be the set of vertices of $G(n, (d_i)_1^n)$ which are influenced starting from an initial source of influence, x, until the process stops, i.e.,

$$C(x) = \{y \in v(G(n, (d_i)_1^n)) : x \to y\}, \tag{3}$$

where $v(G(n, (d_i)_1^n))$ denotes the set of all the vertices of $G(n, (d_i)_1^n)$. We use $|.|$ to denote the number of elements in a set here, although at other times we also use the symbol to denote the absolute value, which would be clear from the context. We have the following theorems for the forward influence propagation process.

Theorem 2.2. *Suppose that Condition 2.1 holds and consider the random graph $G(n, (d_i)_1^n)$, letting $n \to \infty$.*

If $\mathbb{E}[D^{(t)}D] > \mathbb{E}[D^{(t)} + D]$, then there is a unique $\xi \in (0,1)$ such that $H(\xi) = 0$ and there exists at least one x_n in $G(n, (d_i)_1^n)$ such that

$$\frac{|C(x_n)|}{n} \xrightarrow{p} 1 - g(\xi, \xi) > 0. \tag{4}$$

We denote $C(x_n)$ constructed in the proof of Theorem 2.2 by C^*. For every $\epsilon > 0$, let

$$\mathbb{C}^s(\epsilon) := \{x \in v(G(n, (d_i)_1^n)) : |C(x)|/n < \epsilon\}$$

and

$$\mathbb{C}^L(\epsilon) := \{x \in v(G(n, (d_i)_1^n)) : |C(x) \triangle C^*|/n < \epsilon\},$$

where $\triangle$ denotes the symmetric difference.

Theorem 2.3. *Under assumptions of Theorem 2.2, we have that*

$$\forall \epsilon, \quad \frac{|\mathbb{C}^s(\epsilon)| + \left|\mathbb{C}^L(\epsilon)\right|}{n} \xrightarrow{p} 1. \tag{5}$$

Informally, the above theorem says that asymptotically ($n \to \infty$) and under assumptions of Theorem 2.2, there is essentially one and only one big (i.e., of size $O(n)$) graph component that can possibly be influenced starting with propagation from a given vertex in the graph. What this theorem doesn't tell, however, is the relative size of the set of vertices which are indeed able to reach this big component (we call them *pioneers*) to the set of vertices which are able to reach only a component of size $o(n)$, and this is the question we turn to next.

Our analysis technique to obtain the above results involves the simultaneous exploration of the Configuration Model and the propagation of influence. Another commonly used method to explore the components of Configuration Model is to make the branching process approximation in the initial stages of the exploration process. Although we won't explicitly follow this path in this paper, an heuristic analysis of the branching process approximation of our propagation model provides some important insights about the size of the set of pioneers.

We will need the following fundamental result on branching processes (see, for example, [7]).

Fact 2.4 (Survival vs. Extiction)**.** *For the Galton-Watson branching process whose progeny distribution is given by a random variable Z, the extinction probability p_{ext} is given by the smallest solution in $[0, 1]$ of*

$$x = \mathbb{E}(x^Z). \tag{6}$$

In particular, the following regimes can happen:

(i) Subcritical regime: If $\mathbb{E}[Z] < 1$, then $p_{ext} = 1$.

(ii) Critical regime: If $\mathbb{E}[Z] = 1$ and Z is not deterministic, then $p_{ext} = 1$.

(iii) Supercritical regime: If $\mathbb{E}[Z] > 1$, then $p_{ext} < 1$.

Now coming to the approximation, if we start the exploration with a uniformly chosen vertex i, then the number of its neighbours that it does not influence and those that it does, denoted by the random vector $(D_i^{(r)}, D_i^{(t)})$, will have a joint distribution $(p_{v,w})$. But since the probability of getting influenced is proportional to the degree, the number of neighbours of a first-generation vertex excluding its parent (the vertex which influenced it) won't follow this joint distribution. Their joint distribution as well the joint distribution in the subsequent generations, denoted by $(\widetilde{D}^{(r)}, \widetilde{D}^{(t)})$, is given by

$$\widetilde{p}_{v,w} = \frac{(v+1)\, p_{v+1,w} + (w+1)\, p_{v,w+1}}{\lambda}. \tag{7}$$

Note that Condition 2.1(iv) implies that $\mathbb{P}(\widetilde{D}^{(t)} = 0) > 0$, and therefore, from Fact 2.4, this branching process gets extinct a.s. unless,

$$\begin{aligned}
&\mathbb{E}\left[\widetilde{D}^{(t)}\right] > 1; \\
\text{equivalently,}\quad &\sum_{v,w} w\widetilde{p}_{v,w} > 1, \\
&\sum_{v,w} \frac{w\,(v+1)\, p_{v+1,w} + w\,(w+1)\, p_{v,w+1}}{\lambda} > 1, \\
&\mathbb{E}\left[D^{(r)} D^{(t)}\right] + \mathbb{E}\left[D^{(t)}\left(D^{(t)} - 1\right)\right] > \mathbb{E}\left[D\right], \\
&\mathbb{E}\left[D D^{(t)}\right] > \mathbb{E}\left[D + D^{(t)}\right].
\end{aligned}$$

This condition for non-extinction of branching process remarkably agrees with the condition in Theorem 2.2 which determines the possibility of influencing a non-negligible proportion of population.

Further from Fact 2.4, if this condition is satisfied, the extinction probability of the branching process which diverges from the first-generation vertex, $\widetilde{p}_{ext}$, is given by the smallest $x \in (0,1)$ which satisfies

$$\begin{aligned}
&\mathbb{E}\left[x^{\widetilde{D}^{(t)}}\right] = x; \\
\text{equivalently,}\quad &\sum_{v,w} \frac{x^w\,(v+1)\, p_{v+1,w} + (w+1)\, x^w p_{v,w+1}}{\lambda} = x, \\
&\mathbb{E}\left[D^{(r)} x^{D^{(t)}}\right] + \mathbb{E}\left[D^{(t)} x^{D^{(t)}-1}\right] = x\mathbb{E}\left[D\right], \\
&\mathbb{E}\left[D\right] x^2 - \mathbb{E}\left[D^{(t)} x^{D^{(t)}}\right] - x\mathbb{E}\left[D^{(r)} x^{D^{(t)}}\right] = 0.
\end{aligned} \tag{8}$$

Note that 0 is excluded as a solution since $\mathbb{P}(\widetilde{D}^{(t)} = 0) > 0$.

Finally, the extinction probability of the branching process starting from the root, p_{ext}, is given by

$$p_{ext} = \mathbb{E}\left[\left(\widetilde{p}_{ext}\right)^{D^{(t)}}\right]. \tag{9}$$

Since the root is uniformly chosen, we would expect the proportion of the vertices which can influence a non-negligible proportion to be roughly $1 - p_{ext} = 1 - \mathbb{E}\left[(\widetilde{p}_{ext})^{D^{(t)}}\right]$. Indeed, we confirm this result using a more rigorous analysis involving the introduction and study of a reverse influence propagation which essentially traces all the possible sources of influence of a given vertex. This method of introducing a reverse process (in a way, dual to the original process)to derive results for the original process has not been seen in a related context in the existing literature to the best of our knowledge, although the analysis of this dual process uses the familiar tools used for the original process.

Let $\overline{g}(x) := \mathbb{E}[x^{D^{(t)}}]$, $\overline{h}(x) := \mathbb{E}[D^{(t)}x^{D^{(t)}}] + x\mathbb{E}[D^{(r)}x^{D^{(t)}}]$ and

$$\overline{H}(x) := \mathbb{E}[D]x^2 - \overline{h}(x) = \lambda x^2 - \overline{h}(x). \tag{10}$$

Let $\overline{C}(y)$ be the set of vertices of $G(n,(d_i)_1^n)$ starting from which y can be influenced, i.e., $\overline{C}(y) := \{x \in v(G(n,(d_i)_1^n)) : x \to y\}$. We have the following theorems for the dual backward propagation process.

Theorem 2.5. *Under assumptions of Theorem 2.2, there is a unique $\overline{\xi} \in (0,1)$ such that $\overline{H}(\overline{\xi}) = 0$ and there exists at least one y_n in $G^*(n,(d_i)_1^n)$ such that*

$$\frac{\left|\overline{C}(y_n)\right|}{n} \xrightarrow{p} 1 - \overline{g}(\overline{\xi}) > 0. \tag{11}$$

Remark that $\overline{H}(x) = 0$ is the same as equation (8) and therefore $\overline{\xi} \equiv \widetilde{p}_{ext}$ and $1 - \overline{g}(\overline{\xi}) \equiv p_{ext}$ from the branching process approximation.

We denote $\overline{C}(y_n)$ constructed in the proof of Theorem 2.5 by $\overline{C}^*$. For every $\epsilon > 0$, let

$$\overline{\mathbb{C}}^s(\epsilon) := \left\{y \in v(G(n,(d_i)_1^n)) : \left|\overline{C}(y)\right|/n < \epsilon\right\},$$

and

$$\overline{\mathbb{C}}^L(\epsilon) := \left\{y \in v(G(n,(d_i)_1^n)) : \left|\overline{C}(y) \triangle \overline{C}^*\right|/n < \epsilon\right\}.$$

Theorem 2.6. *Under assumptions of Theorem 2.2,*

$$\forall \epsilon, \quad \frac{\left|\overline{\mathbb{C}}^s(\epsilon)\right| + \left|\overline{\mathbb{C}}^L(\epsilon)\right|}{n} \xrightarrow{p} 1. \tag{12}$$

Informally, the above theorem says that asymptotically ($n \to \infty$) and under assumptions of Theorem 2.2, there is essentially one and only one big *source* component in the graph, to which a given vertex can possibly trace back while tracing all the possible sources of its influence.

Finally, we have the following theorem which establishes the duality relation between the two processes.

Theorem 2.7. *Under assumptions of Theorem 2.2, for any $\epsilon > 0$ and $n \to \infty$,*

$$n^{-1}|\overline{\mathbb{C}}^L(\epsilon)|\left|n^{-1}|\overline{C}^*| - n^{-1}|\mathbb{C}^L(\epsilon)|\right| \le \alpha\epsilon + R_n(\epsilon), \tag{13}$$

where $\alpha > 0$ and $R_n(\epsilon) \xrightarrow{p} 0$.

The theorem leads to the following fundamental result of this paper, where it all comes together and we are able to essentially identify, under one additional assumption apart from those in Theorem 2.2, the set of pioneers with the one big source component that we discovered above. In particular, this gives us the relative size (w.r.t. n) of the set of pioneers since we know the relative size of the source component.

Corollary 2.8. *Under assumptions of Theorem 2.2, for any $\epsilon > 0$ and $n \to \infty$, if there exists $a > 0$ such that $n^{-1}|\mathbb{C}^L(\epsilon)| > a$ whp, then*

$$n^{-1}|\mathbb{C}^L(\epsilon) \triangle \overline{C}^*| \le \alpha'\epsilon + R'_n(\epsilon), \tag{14}$$

where $\alpha' > 0$ and $R'_n(\epsilon) \xrightarrow{p} 0$.

Remark 2.9. In particular, if $\mathbb{E}[D^{(t)}(D^{(t)} - 2)] > 0$, then the Configuration Model with the degree sequence $(d_i^{(t)})_1^n$ will have a giant component $C^{(t)}$ whp. In this case, whp $n^{-1}|\mathbb{C}^L(\epsilon)| \ge n^{-1}|C^{(t)}| > a$ for some $a > 0$, and thus the condition in the above corollary is satisfied.

Future Work There is a strong indication that in Corollary 2.8, we do not need the lower bound on $n^{-1}|\mathbb{C}^L(\epsilon)|$ for (14) to hold. One possible approach to prove this would be to make rigorous the branching process approximation heuristically illustrated in the previous section to provide insight (see [3], where the branching process approximation is used to find the largest component of Erdös-Rényi graph). This approach could give not only the required lower bound on $n^{-1}|\mathbb{C}^L(\epsilon)|$ in Corollary 2.8, but even the desired approximation of $n^{-1}|\mathbb{C}^L(\epsilon)|$ which we otherwise obtain by the identification of $\mathbb{C}^L(\epsilon)$ with $\overline{C}^*$ in Corollary 2.8. But even in that case, the introduction of the dual process which leads to the identification of $\mathbb{C}^L(\epsilon)$ with $\overline{C}^*$ is useful since this would provide us with important additional information regarding the structure of $\mathbb{C}^L(\epsilon)$, which we have not explored in this paper.

We also believe that the sufficient condition on the total and the transmitter degree distribution ($\mathbb{E}[D^{(t)}D] > \mathbb{E}[D^{(t)} + D]$) in Theorem 2.2 for influence propagation to go viral, is necessary as well.

3 Analysis of the Original Forward-Propagation Process

The following analysis is similar to the one presented in [11] and wherever the proofs of analogous lemmas, theorems etc. don't have any new point of note, we refer the reader to [11] without giving the proofs.

Throughout the construction and propagation process, we keep track of what we call *active transmitter* half-edges. To begin with, all the vertices and the attached half-edges are *sleeping* but once influenced, a vertex and its half-edges become *active*. Both sleeping and active half-edges at any time constitute what we call *living* half-edges and when two half-edges are matched to reveal an edge along which the flow of influence has occurred, the half-edges are pronounced *dead*. Half-edges are further classified according to their ability or inability to transmit information as *transmitters* and *receivers* respectively. We initially give all the half-edges i.i.d. random maximal lifetimes with distribution given by $\tau \sim \exp(1)$, then go through the following algorithm.

C1 If there is no active half-edge (as in the beginning), select a sleeping vertex and declare it active, along with all its half-edges. For definiteness, we choose the vertex uniformly at random among all sleeping vertices. If there is no sleeping vertex left, the process stops.

C2 Pick an active transmitter half-edge and kill it.

C3 Wait until the next living half-edge dies (spontaneously, due to the expiration of its exponential life-time). This is joined to the one killed in previous step to form an edge of the graph along which information has been transmitted. If the vertex it belongs to is sleeping, we change its status to active, along with all of its half-edges. Repeat from the first step.

Every time C1 is performed, we choose a vertex and trace the flow of influence from here onwards. Just before C1 is performed again, when the number of active transmitter half-edges goes to 0, we've explored the extent of the graph component that the chosen vertex can influence, that had not been previously influenced.

Let $S_T(t)$, $S_R(t)$, $A_T(t)$ and $A_R(t)$ represent the number of sleeping transmitter, sleeping receiver, active transmitter and active receiver half-edges, respectively, at time t. Therefore, $R(t) := A_R(t) + S_R(t)$ and $L(t) := A_T(t) + A_R(t) + S_T(t) + S_R(t) = A_T(t) + S_T(t) + R(t)$ denotes the number of receiver and living half-edges, respectively, at time t.

For definiteness, we will take them all to be right-continuous, which along with C1 entails that $L(0) = 2m - 1$. Subsequently, whenever a living half-edge dies spontaneously, C3 is performed, immediately followed by C2. As such, $L(t)$ is decreased by 2 every time a living half-edge dies spontaneously, up until the last living one die and the process terminates. Also remark that all the receiver half-edges, both sleeping and active, continue to die spontaneously.

The following consequences of *Glivenko-Cantelli* theorem are analogous to those given in [11] and we state them without proof.

Lemma 3.1. *As $n \to \infty$,*

$$\sup_{t \geq 0} \left| n^{-1} L(t) - \lambda e^{-2t} \right| \xrightarrow{p} 0. \tag{15}$$

Lemma 3.2. *As $n \to \infty$,*

$$\sup_{t \geq 0} \left| n^{-1} R(t) - \lambda_r e^{-t} \right| \xrightarrow{p} 0. \tag{16}$$

Let $V_{k,l}(t)$ be the number of sleeping vertices at time t which started with receiver and transmitter degrees k and l respectively . Clearly,

$$S_T(t) = \sum_{k,l} l V_{k,l}(t). \tag{17}$$

Among the three steps, only C1 is responsible for premature death (before the expiration of exponential life-time) of sleeping vertices. We first ignore its effect by letting $\widetilde{V}_{k,l}(t)$ be the number of vertices with receiver and transmitter degrees k and l respectively, such that all their half-edges would die spontaneously (without the aid of C1) after time t. Correspondingly, let $\widetilde{S}_T(t) = \sum_{k,l} l \widetilde{V}_{k,l}(t)$.

Then,

Lemma 3.3. *As $n \to \infty$,*

$$\sup_{t\geq 0}\left|n^{-1}\widetilde{V}_{k,l}(t) - p_{k,l}e^{-(k+l)t}\right| \xrightarrow{p} 0. \tag{18}$$

for all $(k,l) \in \mathbb{N}^2$, and

$$\sup_{t\geq 0}\left|n^{-1}\sum_{k,l}\widetilde{V}_{k,l}(t) - g(e^{-t},e^{-t})\right| \xrightarrow{p} 0. \tag{19}$$

$$\sup_{t\geq 0}\left|n^{-1}\widetilde{S}_T(t) - h(e^{-t})\right| \xrightarrow{p} 0. \tag{20}$$

Proof. Again, (18) follows from *Glivenko-Cantelli* theorem. To prove (20), note that by Condition 2.1(iii), $D_n = D_n^{(r)} + D_n^{(t)}$ are uniformly integrable, i.e., for every $\epsilon > 0$ there exists $K < \infty$ such that for all n,

$$\mathbb{E}(D_n; D_n > K) = \sum_{(k,l;k+l>K)}(k+l)\frac{u_{k,l}}{n} < \epsilon. \tag{21}$$

This, by Fatou's inequality, further implies that

$$\sum_{(k,l;k+l>K)}(k+l)p_{k,l} < \epsilon. \tag{22}$$

Thus, by (18), we have whp,

$$\begin{aligned}
\sup_{t\geq 0}\left|n^{-1}\widetilde{S}_T(t) - h(e^{-t})\right| &= \sup_{t\geq 0}\left|\sum_{k,l} l(n^{-1}\widetilde{V}_{k,l}(t) - p_{k,l}e^{-(k+l)t})\right| \\
&\leq \sum_{(k,l;k+l\leq K)} l\sup_{t\geq 0}\left|(n^{-1}\widetilde{V}_{k,l}(t) - p_{k,l}e^{-(k+l)t})\right| + \\
&\quad \sum_{(k,l;k+l>K)} l(\frac{u_{k,l}}{n} + p_{k,l}) \\
&\leq \epsilon + \epsilon + \epsilon,
\end{aligned}$$

which proves (20). A similar argument also proves (19). ∎

Lemma 3.4. *If $d_{max} := \max_i d_i$ is the maximum degree of $G^*(n,(d_i)_1^n)$, then*

$$0 \leq \widetilde{S}_T(t) - S_T(t) < \sup_{0\leq s\leq t}(\widetilde{S}_T(s) + R(s) - L(s)) + d_{max}. \tag{23}$$

Proof. Clearly, $V_{k,l}(t) \leq \widetilde{V}_{k,l}(t)$, and thus $S_T(t) \leq \widetilde{S}_T(t)$. Therefore, we have that $\widetilde{S}_T(t) - S_T(t) \geq 0$ and the difference increases only when C1 is performed. Suppose that happens at time t and a sleeping vertex of degree $j > 0$ gets activated, then C2 applies immediately and we have $A_T(t) \leq j - 1 < d_{max}$, and consequently,

$$\begin{aligned}\widetilde{S}_T(t) - S_T(t) &= \widetilde{S}_T(t) - (L(t) - R(t) - A_T(t)) \\ &< \widetilde{S}_T(t) + R(t) - L(t) + d_{max}.\end{aligned}$$

Since $\widetilde{S}_T(t) - S_T(t)$ does not change in the intervals during which C1 is not performed, $\widetilde{S}_T(t) - S_T(t) \leq \widetilde{S}_T(s) - S_T(s)$, where s is the last time before t that C1 was performed. The lemma follows. □

Let

$$\widetilde{A}_T(t) := L(t) - R(t) - \widetilde{S}_T(t) = A_T(t) - (\widetilde{S}_T(t) - S_T(t)). \tag{24}$$

Then, Lemma 3.4 can be rewritten as

$$\widetilde{A}_T(t) \leq A_T(t) < \widetilde{A}_T(t) - \inf_{s \leq t} \widetilde{A}_T(s) + d_{max}. \tag{25}$$

Also, by Lemmas 3.1, 3.2 and 3.3 and (2),

$$\sup_{t \geq 0} \left| n^{-1}\widetilde{A}_T(t) - H(e^{-t}) \right| \xrightarrow{p} 0. \tag{26}$$

Lemma 3.5. *Suppose that Condition 2.1 holds and let $H(x)$ be given by (2).*

(i) If $\mathbb{E}[D^{(t)}D] > \mathbb{E}[D^{(t)} + D]$, then there is a unique $\xi \in (0,1)$, such that $H(\xi) = 0$; moreover, $H(x) < 0$ for $x \in (0,\xi)$ and $H(x) > 0$ for $x \in (\xi, 1)$.

(ii) If $\mathbb{E}[D^{(t)}D] \leq \mathbb{E}[D^{(t)} + D]$, then $H(x) < 0$ for $x \in (0,1)$.

Proof. Remark that $H(0) = H(1) = 0$ and $H'(1) = 2\mathbb{E}[D] - \mathbb{E}[D^{(r)}] - \mathbb{E}[D^{(t)}D] = \mathbb{E}[D + D^{(t)}] - \mathbb{E}[D^{(t)}D]$. Furthermore we define $\phi(x) := H(x)/x = \lambda x - \lambda_r - \sum_{k,l} l p_{k,l} x^{k+l-1}$, which is a concave function on $(0,1]$, in fact, strictly concave unless $p_{k,l} = 0$ whenever $k + l \geq 3$ and $l \geq 1$, in which case $H'(1) = p_{0,1} + p_{1,1} + \sum_{k\geq 1} k p_{k,0} \geq p_{0,1} + p_{1,0} = \mathbb{P}(D = 1) > 0$, by Condition 2.1(iv).

In case (ii), we thus have ϕ concave and $\phi'(1) = H'(1) - H(1) \geq 0$, with either the concavity or the above inequality strict, and thus $\phi'(x) > 0$ for all $x \in (0,1)$, whence $\phi(x) < \phi(1) = 0$ for $x \in (0,1)$.

In case (i), $H'(1) < 0$, and thus $H(x) > 0$ for x close to 1. Further,

$$\begin{aligned} H'(0) &= -\lambda_r - \sum_{\{(k,l):k+l=1\}} l p_{k,l} \\ &= -\lambda_r - p_{0,1} \\ &\leq -p_{1,0} - p_{0,1} < 0 \end{aligned}$$

by Condition 2.1(iv), which implies that $H(x) < 0$ for x close to 0. Hence there is at least one $\xi \in (0,1)$ with $H(\xi) = 0$. Now, since $H(x)/x$ is strictly concave and also $\phi(1) = H(1) = 0$, there is at most one such ξ. This proves the result. □

Proof of Theorem 2.2. Let ξ be the zero of H given by Lemma 3.5(i) and let $\tau := -\ln \xi$. Then, by Lemma 3.5, $H(e^{-t}) > 0$ for $0 < t < \tau$, and thus $\inf_{t \leq \tau} H(e^{-t}) = 0$. Consequently, (26) implies

$$n^{-1} \inf_{t \leq \tau} \widetilde{A}_T(t) = n^{-1} \inf_{t \leq \tau} \widetilde{A}_T(t) - \inf_{t \leq \tau} H(e^{-t}) \xrightarrow{p} 0. \tag{27}$$

Further, by Condition 2.1(iii), $d_{max} = O(n^{1/2})$, and thus $n^{-1} d_{max} \to 0$. Consequently, by (25) and (27)

$$\sup_{t \leq \tau} n^{-1} \left| A_T(t) - \widetilde{A}_T(t) \right| = \sup_{t \leq \tau} n^{-1} \left| \widetilde{S}_T(t) - S_T(t) \right| \xrightarrow{p} 0, \tag{28}$$

and thus, by (26),

$$\sup_{t \geq 0} \left| n^{-1} A_T(t) - H(e^{-t}) \right| \xrightarrow{p} 0. \tag{29}$$

Let $0 < \epsilon < \tau/2$. Since $H(e^{-t}) > 0$ on the compact interval $[\epsilon, \tau - \epsilon]$, (29) implies that whp $A_T(t)$ remains positive on $[\epsilon, \tau - \epsilon]$, and thus C1 is not performed during this interval.

On the other hand, again by Lemma 3.5(i), $H(e^{-\tau-\epsilon}) < 0$ and (26) implies $n^{-1} \widetilde{A}_T(\tau + \epsilon) \xrightarrow{p} H(e^{-\tau-\epsilon})$, while $A_T(\tau + \epsilon) \geq 0$. Thus, with $\delta := \left| H(e^{-\tau-\epsilon}) \right| / 2 > 0$, whp

$$\widetilde{S}_T(\tau + \epsilon) - S_T(\tau + \epsilon) = A_T(\tau + \epsilon) - \widetilde{A}_T(\tau + \epsilon) \geq -\widetilde{A}_T(\tau + \epsilon) > n\delta, \tag{30}$$

while (28) implies that $\widetilde{S}_T(\tau) - S_T(\tau) < n\delta$ whp. Consequently, whp $\widetilde{S}_T(\tau + \epsilon) - S_T(\tau + \epsilon) > \widetilde{S}_T(\tau) - S_T(\tau)$, so C1 is performed between τ and $\tau + \epsilon$.

Let T_1 be the last time that C1 is performed before $\tau/2$, let x_n be the sleeping vertex declared active at this point of time and let T_2 be the next time C1 is performed. We have shown that for any $\epsilon > 0$, whp $0 \le T_1 \le \epsilon$ and $\tau - \epsilon \le T_2 \le \tau + \epsilon$; in other words, $T_1 \xrightarrow{p} 0$ and $T_2 \xrightarrow{p} \tau$.

We next use the following lemma.

Lemma 3.6. *Let T_1^* and T_2^* be two (random) times when C1 are performed, with $T_1^* \le T_2^*$, and assume that $T_1^* \xrightarrow{p} t_1$ and $T_2^* \xrightarrow{p} t_2$ where $0 \le t_1 \le t_2 \le \tau$. If C is the union of all the vertices informed between T_1^* and T_2^*, then*

$$|C|/n \xrightarrow{p} g(e^{-t_1}, e^{-t_1}) - g(e^{-t_2}, e^{-t_2}). \tag{31}$$

Proof. For all $t \ge 0$, we have

$$\sum_{i,j} (\widetilde{V}_{i,j}(t) - V_{i,j}(t)) \le \sum_{i,j} j(\widetilde{V}_{i,j}(t) - V_{i,j}(t)) = \widetilde{S}_T(t) - S_T(t).$$

Thus,

$$\begin{aligned} |C| &= \sum (V_{k,l}(T_1^*-) - V_{k,l}(T_2^*-)) = \sum (\widetilde{V}_{k,l}(T_1^*-) - \widetilde{V}_{k,l}(T_2^*-)) + o_p(n) \\ &= ng(e^{-T_1^*}, e^{-T_1^*}) - ng(e^{-T_2^*}, e^{-T_2^*}) + o_p(n). \end{aligned}$$

□

Let C' be the set of vertices informed up till T_1 and C'' be the set of vertices informed between T_1 and T_2. Then, by Lemma 3.6, we have that

$$\frac{|C'|}{n} \xrightarrow{p} 0 \tag{32}$$

and

$$\frac{|C''|}{n} \xrightarrow{p} g(1,1) - g(e^{-\tau}, e^{-\tau}) = 1 - g(e^{-\tau}, e^{-\tau}). \tag{33}$$

Evidently, $C'' \subset C(x_n)$. Note that $C(x_n) = \{y \in v(G^*(n, (d_i)_1^n)) : x_n \to y\}$. It is clear that if $x_n \to y$, then $y \notin (C' \cup C'')^c$. Therefore, we have that $C(x_n) \subset C' \cup C''$, which implies that

$$|C''| \le |C(x_n)| \le |C'| + |C''|, \tag{34}$$

and thus, from (32) and (33),

$$\frac{|C(x_n)|}{n} \xrightarrow{p} 1 - g(e^{-\tau}, e^{-\tau}), \tag{35}$$

which completes the proof of Theorem 2.2.

□

Proof of Theorem 2.3. We continue from where we left in the proof of previous theorem, with the following Lemmas. Assumptions of Theorem 2.2 continue to hold for what follows in this section.

Lemma 3.7. *$\forall \epsilon > 0$, let*

$$\mathbb{A}(\epsilon) := \left\{ y \in v(G^*(n, (d_i)_1^n)) : \frac{|C(y)|}{n} \geq \epsilon \text{ and } \left| \frac{|C(y)|}{n} - (1 - g(\xi, \xi)) \right| \geq \epsilon \right\}.$$

Then,

$$\forall \epsilon, \quad \frac{|\mathbb{A}(\epsilon)|}{n} \xrightarrow{p} 0. \tag{36}$$

Proof. Suppose the converse is true. Then, there exists $\delta > 0$, $\delta' > 0$ and a sequence $(n_k)_{k>0}$ such that

$$\forall k, \quad \mathbb{P}\left(\frac{|\mathbb{A}(\epsilon)|}{n_k} > \delta \right) > \delta'. \tag{37}$$

Since the vertex initially informed to start the transmission process, say a, is uniformly chosen, we have

$$\forall n_k, \quad \mathbb{P}(a \in \mathbb{A}(\epsilon)) > \delta\delta' \tag{38}$$

and thus,

$$\forall k, \quad \mathbb{P}\left(\frac{|C'|}{n_k} \geq \epsilon \text{ or } \left| \frac{|C''|}{n_k} - (1 - g(\xi, \xi)) \right| \geq \epsilon \right) > \delta\delta', \tag{39}$$

which contradicts (33). □

Lemma 3.8. *For every $\epsilon > 0$, let*

$$\mathbb{B}(\epsilon) := \left\{ y \in C' \cup C'' : |C(y)|/n \geq \epsilon \text{ and } |C(y) \triangle C^*|/n \geq \epsilon \right\}. \tag{40}$$

Then,

$$\forall \epsilon, \quad \frac{|\mathbb{B}(\epsilon)|}{n} \xrightarrow{p} 0. \tag{41}$$

Proof. Recall that for any three sets A, B and C, we have that $A \triangle B \subset (A \triangle C) \cup (B \triangle C)$. Therefore, for any $y \in C' \cup C''$, we have that

$$C(y) \triangle C^* \subset \left[C(y) \triangle (C' \cup C'') \right] \cup \left[C^* \triangle (C' \cup C'') \right]. \tag{42}$$

But recall that $C^* \subset C' \cup C''$ and by a similar argument, for every $y \in C' \cup C''$, $C(y) \subset C' \cup C''$. Thus,

$$C(y) \bigtriangleup C^* \subset \left[(C' \cup C'') \setminus C(y)\right] \cup \left[(C' \cup C'') \setminus C^*\right]. \tag{43}$$

Hence, if $|C(y) \bigtriangleup C^*|/n \geq \epsilon$, then either $|(C' \cup C'') \setminus C(y)|/n \geq \epsilon/2$ or $|(C' \cup C'') \setminus C^*|/n \geq \epsilon/2$. Consequently,

$$\begin{aligned}\mathbb{B}(\epsilon) \subset &\left\{y \in v(G^*(n,(d_i)_1^n)) : \epsilon \leq |C(y)|/n \leq \left|(C' \cup C'')\right|/n - \epsilon/2\right\} \\ &\cup \left\{y \in v(G^*(n,(d_i)_1^n)) : \left|(C' \cup C'') \setminus C^*\right|/n \geq \epsilon/2\right\}.\end{aligned}$$

Letting $e1 := |\{y \in v(G^*(n,(d_i)_1^n)) : \epsilon \leq |C(y)|/n \leq |(C' \cup C'')|/n - \epsilon/2\}|/n$ and $E2 := \{|(C' \cup C'') \setminus C^*|/n \geq \epsilon/2\}$, we have

$$\mathbb{B}(\epsilon)/n \leq e1 + \mathbf{1}_{E2}. \tag{44}$$

Now, $e1 \xrightarrow{p} 0$ by (33) and Lemma 3.7, while $\mathbf{1}_{E2} \xrightarrow{p} 0$ because $\mathbb{P}(E2) \to 0$ by (32), (33) and (34). This concludes the proof. □

Lemma 3.9. *Let T_3 be the first time after T_2 that C1 is performed and let z_n be the sleeping vertex activated at this moment. If C''' is the set of vertices informed between T_2 and T_3, then*

$$\frac{|C'''|}{n} \xrightarrow{p} 0. \tag{45}$$

Proof. Since $\widetilde{S}_T(t) - S_T(t)$ increases by at most $d_{max} = o_p(n)$ each time C1 is performed, we obtain that

$$\sup_{t \leq T_3}(\widetilde{S}_T(t) - S_T(t)) \leq \sup_{t \leq T_2}(\widetilde{S}_T(t) - S_T(t)) + d_{max} = o_p(n). \tag{46}$$

Comparing this to (30) we see that for every $\epsilon > 0$, whp $\tau + \epsilon > T_3$. Since also $T_3 > T_2 \xrightarrow{p} \tau$, it follows that $T_3 \xrightarrow{p} \tau$. This in combination with Lemma 3.6 yields that

$$\frac{|C'''|}{n} \xrightarrow{p} 0.$$

□

Lemma 3.10. *For every $\epsilon > 0$, let*

$$\mathbb{C}(\epsilon) := \left\{z \in \left(C' \cup C''\right)^c : |C(z)|/n \geq \epsilon \text{ and } |C(z) \bigtriangleup C^*|/n \geq \epsilon\right\}. \tag{47}$$

Then, we have that

$$\forall \epsilon, \quad \frac{|\mathbb{C}(\epsilon)|}{n} \xrightarrow{p} 0. \tag{48}$$

Proof. We start by remarking that by Lemma 3.7, it is sufficient to prove that

$$\frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n} \xrightarrow{p} 0. \tag{49}$$

Now assume that there exist $\delta, \delta' > 0$ and a sequence $(n_k)_{k>0}$ such that

$$\forall k, \quad \mathbb{P}\left(\frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n_k} > \delta\right) > \delta'. \tag{50}$$

Let

$$\mathcal{E}_1 := \{\text{Configuration Model completely revealed}\},$$

$$\mathcal{E}_2 := \{\text{Influence propagation revealed upto } C''\}$$

and $\mathcal{E}_3 := \mathcal{E}_1 \cap \mathcal{E}_2$.Then, denoting be z_{n_k} the vertex awakened by C1 at time T_2, we have that

$$\begin{aligned}
&\mathbb{P}\left(z_{n_k} \in \mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon) | \mathcal{E}_3\right) \\
&\geq \frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n_k - |C' \cup C''|} \mathbf{1}\left(\frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n_k} > \delta\right) \\
&\geq \frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n_k} \mathbf{1}\left(\frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n_k} > \delta\right) \\
&\geq \delta \mathbf{1}\left(\frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n_k} > \delta\right).
\end{aligned}$$

Taking expectations, we have

$$\mathbb{P}\left(z_{n_k} \in \mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)\right) \geq \delta\delta'. \tag{51}$$

But this leads to contradiction. Indeed, we have that

$$C(z_n) \vartriangle C^* \subset \left[C(z_n) \vartriangle (C' \cup C'' \cup C''')\right] \cup \left[C^* \vartriangle (C' \cup C'' \cup C''')\right]. \tag{52}$$

Again recall that $C^* \subset C' \cup C'' \cup C'''$ and by a similar argument, $C(z_n) \subset C' \cup C'' \cup C'''$ so that

$$C(z_n) \vartriangle C^* \subset \left[(C' \cup C'' \cup C''') \setminus C(z_n)\right] \cup \left[(C' \cup C'' \cup C''') \setminus C^*\right]. \tag{53}$$

Hence, if $|C(z_n) \vartriangle C^*| / n \geq \epsilon$, then either

$$\begin{gathered}
\left|(C' \cup C'' \cup C''') \setminus C(z_n)\right| / n \geq \epsilon/2 \\
\text{equivalently, } |C(z_n)| / n \leq \left|(C' \cup C'' \cup C''')\right| / n - \epsilon/2,
\end{gathered}$$

or,

$$\left|(C' \cup C'' \cup C''') \setminus C^*\right| / n \geq \epsilon/2.$$

Let

$$E3 := \left\{ |C(z_n)| / n \leq \left|(C' \cup C'' \cup C''')\right| / n - \epsilon/2 \right\}$$

and

$$E4 := \left\{ \left|(C' \cup C'' \cup C''') \setminus C^*\right| / n \geq \epsilon/2 \right\}.$$

Now assume that $z_n \in \mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)$. This implies that either $E4$ holds or $\left\{1 - g(\xi, \xi) - \epsilon \leq \frac{|C(z_n)|}{n} \leq 1 - g(\xi, \xi) + \epsilon\right\} \cap E3$ holds. But thanks to (32), (33) and Lemma 3.9, neither of these two events hold with asymptotically positive probability.

This completes the proof. □

Finally, Lemma 3.8 and Lemma 3.10 allow us to conclude that

$$\forall \epsilon, \quad \frac{|\mathbb{C}^s(\epsilon)| + \left|\mathbb{C}^L(\epsilon)\right|}{n} \xrightarrow{p} 1. \tag{54}$$

□

4 Analysis of the Dual Back-Propagation Process

Now we introduce the algorithm to trace the possible sources of influence of a randomly chosen vertex. We borrow the terminology from the previous section, only in this case we put a *bar* over the label to indicate that we're talking about the dual process. The analysis also proceeds along the same lines as that of the original process, and we do not give the proof when it differs from the analogous proof in the previous section only by notation.

As before, we initially give all the half-edges i.i.d. random maximal lifetimes with distribution $\overline{\tau} \sim \exp(1)$ and then go through the following algorithm.

$\overline{\text{C}}1$ If there is no active half-edge (as in the beginning), select a sleeping vertex and declare it active, along with all its half-edges. For definiteness, we choose the vertex uniformly at random among all sleeping vertices. If there is no sleeping vertex left, the process stops.

$\overline{\text{C}}2$ Pick an active half-edge and kill it.

$\overline{\mathrm{C}}3$ Wait until the next transmitter half-edge dies (spontaneously). This is joined to the one killed in previous step to form an edge of the graph. If the vertex it belongs to is sleeping, we change its status to active, along with all of its half-edges. Repeat from the first step.

Again, as before, $\overline{L}(0) = 2m - 1$ and we have the following consequences of *Glivenko-Cantelli* theorem.

Lemma 4.1. *As* $n \to \infty$,

$$\sup_{t\geq 0} \left| n^{-1}\overline{L}(t) - \lambda e^{-2t} \right| \xrightarrow{p} 0. \tag{55}$$

Let $\overline{V}_{k,l}(t)$ be the number of sleeping vertices at time t which had receiver and transmitter degrees k and l respectively at time 0. It is easy to see that

$$\overline{S}(t) = \sum_{k,l} (ke^{-t} + l)\overline{V}_{k,l}(t). \tag{56}$$

Let $\widetilde{\overline{V}}_{k,l}(t)$ be the corresponding number if the impact of $\overline{\mathrm{C}}1$ on sleeping vertices is ignored. Correspondingly, let $\widetilde{\overline{S}}(t) = \sum_{k,l}(ke^{-t} + l)\widetilde{\overline{V}}_{k,l}(t)$.

Then,

Lemma 4.2. *As* $n \to \infty$,

$$\sup_{t\geq 0} \left| n^{-1}\widetilde{\overline{V}}_{k,l}(t) - p_{k,l}e^{-lt}) \right| \xrightarrow{p} 0. \tag{57}$$

for all $(k,l) \in \mathbb{N}^2$*, and*

$$\sup_{t\geq 0} \left| n^{-1} \sum_{k,l} \widetilde{\overline{V}}_{k,l}(t) - \overline{g}(e^{-t}) \right| \xrightarrow{p} 0. \tag{58}$$

$$\sup_{t\geq 0} \left| n^{-1}\widetilde{\overline{S}}(t) - \overline{h}(e^{-t}) \right| \xrightarrow{p} 0. \tag{59}$$

Proof. Again, (57) follows from *Glivenko-Cantelli* theorem.

To prove (59), note that by (3) of Condition(2.1), $D_n = D_n^{(r)} + D_n^{(t)}$ are uniformly integrable, i.e., for every $\epsilon > 0$ there exists $K < \infty$ such that for all n,

$$\mathbb{E}(D_n; D_n > K) = \sum_{(k,l;k+l>K)} (k+l)\frac{u_{k,l}}{n} < \epsilon. \tag{60}$$

This, by Fatou's inequality, further implies that

$$\sum_{(k,l;k+l>K)} (k+l)p_{k,l} < \epsilon. \tag{61}$$

Thus, by (57), we have whp,

$$\begin{aligned}
\sup_{t\geq 0}\left|n^{-1}\widetilde{\overline{S}}(t) - \overline{h}(e^{-t})\right| &= \sup_{t\geq 0}\left|\sum_{k,l}(ke^{-t}+l)(n^{-1}\widetilde{\overline{V}}_{k,l}(t) - p_{k,l}e^{-lt})\right| \\
&\leq \sum_{(k,l;k+l\leq K)} (k+l)\sup_{t\geq 0}\left|(n^{-1}\widetilde{\overline{V}}_{k,l}(t) - p_{k,l}e^{-lt})\right| + \\
&\quad \sum_{(k,l;k+l>K)} (k+l)(\frac{u_{k,l}}{n} + p_{k,l}) \\
&\leq \epsilon + \epsilon + \epsilon,
\end{aligned}$$

which proves (59). A similar argument also proves (58). $\square$

Lemma 4.3. *If $d_{max} := \max_i d_i$ is the maximum degree of $G^*(n,(d_i)_1^n)$, then*

$$0 \leq \widetilde{\overline{S}}(t) - \overline{S}(t) < \sup_{0\leq s\leq t}(\widetilde{\overline{S}}(s) - L(s)) + d_{max}. \tag{62}$$

Proof. Clearly, $\overline{V}_{k,l}(t) \leq \widetilde{\overline{V}}_{k,l}(t)$, and thus $\overline{S}(t) \leq \widetilde{\overline{S}}(t)$. Therefore, we have that $\widetilde{\overline{S}}(t) - \overline{S}(t) \geq 0$ and the difference increases only when $\overline{\mathrm{C}}1$ is performed. Suppose that happens at time t and a sleeping vertex of degree $j>0$ gets activated, then $\overline{\mathrm{C}}2$ applies immediately and we have $\overline{A}(t) \leq j-1 < d_{max}$, and consequently,

$$\begin{aligned}
\widetilde{\overline{S}}(t) - \overline{S}(t) &= \widetilde{\overline{S}}(t) - (\overline{L}(t) - \overline{A}(t)) \\
&< \widetilde{\overline{S}}(t) - \overline{L}(t) + d_{max}.
\end{aligned}$$

Since $\widetilde{\overline{S}}(t) - \overline{S}(t)$ does not change in the intervals during which $\overline{\mathrm{C}}1$ is not performed, $\widetilde{\overline{S}}(t) - \overline{S}(t) \leq \widetilde{\overline{S}}(s) - \overline{S}(s)$, where s is the last time before t that $\overline{\mathrm{C}}1$ was performed. The lemma follows. $\square$

Let

$$\widetilde{\overline{A}}(t) := \overline{L}(t) - \widetilde{\overline{S}}(t) = \overline{A}(t) - (\widetilde{\overline{S}}(t) - \overline{S}(t)). \tag{63}$$

Then, Lemma 4.3 can be rewritten as

$$\widetilde{\overline{A}}(t) \leq \overline{A}(t) < \widetilde{\overline{A}}(t) - \inf_{s \leq t} \widetilde{\overline{A}}(s) + d_{max}. \tag{64}$$

Also, by Lemmas 4.1 and 4.2 and (10),

$$\sup_{t \geq 0} \left| n^{-1} \widetilde{\overline{A}}(t) - \overline{H}(e^{-t}) \right| \xrightarrow{p} 0. \tag{65}$$

Lemma 4.4. *Suppose that Condition 2.1 holds and let $\overline{H}(x)$ be given by (10).*

(i) If $\mathbb{E}[D^{(t)}D] > \mathbb{E}[D^{(t)} + D]$, then there is a unique $\overline{\xi} \in (0,1)$, such that $\overline{H}(\overline{\xi}) = 0$; moreover, $\overline{H}(x) < 0$ for $x \in (0, \overline{\xi})$ and $\overline{H}(x) > 0$ for $x \in (\overline{\xi}, 1)$.

(ii) If $\mathbb{E}[D^{(t)}D] \leq \mathbb{E}[D^{(t)} + D]$, then $\overline{H}(x) < 0$ for $x \in (0,1)$.

Proof. Remark that $\overline{H}(0) = \overline{H}(1) = 0$ and $\overline{H}'(1) = 2\mathbb{E}[D] - \mathbb{E}[(D^{(t)})^2] - \mathbb{E}[(D^{(r)})] - \mathbb{E}[D^{(r)}D^{(t)}] = \mathbb{E}[D + D^{(t)}] - \mathbb{E}[D^{(t)}D]$. Furthermore we define $\overline{\phi}(x) := \overline{H}(x)/x = \lambda x - \sum_{k,l} l p_{k,l} x^{l-1} - \sum_{k,l} k p_{k,l} x^l$, which is a concave function on $(0,1]$, in fact, strictly concave unless $p_{k,l} = 0$ whenever $l > 2$, or $l = 2$ and $k \geq 1$, in which case $\overline{H}'(1) = \sum_{k \geq 0} p_{k,1} + \sum_{k \geq 0} k p_{k,0} \geq p_{1,0} + p_{0,1} > 0$ by Condition 2.1(iv).

In case (ii), we thus have $\overline{\phi}$ concave and $\overline{\phi}'(1) = \overline{H}'(1) - \overline{H}(1) \geq 0$, with either the concavity or the above inequality strict, and thus $\overline{\phi}'(x) > 0$ for all $x \in (0,1)$, whence $\overline{\phi}(x) < \overline{\phi}(1) = 0$ for $x \in (0,1)$.

In case (i), $\overline{H}'(1) < 0$, and thus $\overline{H}(x) > 0$ for x close to 1. Further, in case (i),

$$\overline{H}'(0) = -\sum_k p_{k,1} - \sum_k k p_{k,0} \leq -p_{1,0} - p_{0,1} < 0 \tag{66}$$

by Condition 2.1(iv), which implies that $\overline{H}(x) < 0$ for x close to 0. Hence there is at least one $\overline{\xi} \in (0,1)$ with $\overline{H}(\overline{\xi}) = 0$. Now, since $\overline{H}(x)/x$ is strictly concave and also $\overline{H}(1) = 0$, there is at most one such $\overline{\xi}$. This proves the result.

□

Proof of Theorem 2.5. Let $\overline{\xi}$ be the zero of $\overline{H}$ given by Lemma 4.4(i) and let $\overline{\tau} := -\ln \overline{\xi}$. Then, by Lemma 4.4, $\overline{H}(e^{-t}) > 0$ for $0 < t < \overline{\tau}$, and thus $\inf_{t \leq \overline{\tau}} \overline{H}(e^{-t}) = 0$. Consequently, (65) implies

$$n^{-1} \inf_{t \leq \overline{\tau}} \widetilde{\overline{A}}(t) = n^{-1} \inf_{t \leq \overline{\tau}} \widetilde{\overline{A}}(t) - \inf_{t \leq \overline{\tau}} \overline{H}(e^{-t}) \xrightarrow{p} 0. \tag{67}$$

Further, by Condition 2.1(iii), $d_{max} = O(n^{1/2})$, and thus $n^{-1}d_{max} \to 0$. Consequently, by (64) and (67)

$$\sup_{t\le\overline{\tau}} n^{-1}\left|\overline{A}(t) - \widetilde{\overline{A}}(t)\right| = \sup_{t\le\overline{\tau}} n^{-1}\left|\widetilde{\overline{S}}(t) - \overline{S}(t)\right| \xrightarrow{p} 0 \tag{68}$$

and thus, by (65),

$$\sup_{t\ge 0}\left|n^{-1}\overline{A}(t) - \overline{H}(e^{-t})\right| \xrightarrow{p} 0. \tag{69}$$

Let $0 < \epsilon < \overline{\tau}/2$. Since $\overline{H}(e^{-t}) > 0$ on the compact interval $[\epsilon, \overline{\tau}-\epsilon]$, (69) implies that whp $\overline{A}(t)$ remains positive on $[\epsilon, \overline{\tau}-\epsilon]$, and thus $\overline{\mathrm{C}}1$ is not performed during this interval.

On the other hand, again by Lemma 4.4(i), $\overline{H}(e^{-\overline{\tau}-\epsilon}) < 0$ and (65) implies $n^{-1}\widetilde{\overline{A}}(\overline{\tau}+\epsilon) \xrightarrow{p} \overline{H}(e^{-\overline{\tau}-\epsilon})$, while $\overline{A}(t)(\overline{\tau}+\epsilon) \ge 0$. Thus, with $\delta := \left|\overline{H}(e^{-\overline{\tau}-\epsilon})\right|/2 > 0$, whp

$$\widetilde{\overline{S}}(\overline{\tau}+\epsilon) - \overline{S}(\overline{\tau}+\epsilon) = \overline{A}(t)(\overline{\tau}+\epsilon) - \widetilde{\overline{A}}(\overline{\tau}+\epsilon) \ge -\widetilde{\overline{A}}(\overline{\tau}+\epsilon) > n\delta, \tag{70}$$

while (68) implies that $\widetilde{\overline{S}}(\overline{\tau}) - \overline{S}(\overline{\tau}) < n\delta$ whp. Consequently, whp $\widetilde{\overline{S}}(\overline{\tau}+\epsilon) - \overline{S}(\overline{\tau}+\epsilon) > \widetilde{\overline{S}}(\overline{\tau}) - \overline{S}(\overline{\tau})$, so $\overline{\mathrm{C}}1$ is performed between $\overline{\tau}$ and $\overline{\tau}+\epsilon$.

Let $\overline{T}_1$ be the last time that $\overline{\mathrm{C}}1$ is performed before $\overline{\tau}/2$, let y_n be the sleeping vertex declared active at this point of time and let $\overline{T}_2$ be the next time $\overline{\mathrm{C}}1$ is performed. We have shown that for any $\epsilon > 0$, whp $0 \le \overline{T}_1 \le \epsilon$ and $\overline{\tau} - \epsilon \le \overline{T}_2 \le \overline{\tau}+\epsilon$; in other words, $\overline{T}_1 \xrightarrow{p} 0$ and $\overline{T}_2 \xrightarrow{p} \overline{\tau}$.

We next use the following lamma.

Lemma 4.5. *Let $\overline{T}_1^*$ and $\overline{T}_2^*$ be two (random) times when $\overline{C}1$ are performed, with $\overline{T}_1^* \le \overline{T}_2^*$, and assume that $\overline{T}_1^* \xrightarrow{p} t_1$ and $\overline{T}_2^* \xrightarrow{p} t_2$ where $0 \le t_1 \le t_2 \le \overline{\tau}$. If $\overline{C}$ is the union of all the informer vertices reached between $\overline{T}_1^*$ and $\overline{T}_2^*$, then*

$$\left|\overline{C}\right|/n \xrightarrow{p} \overline{g}(e^{-t_1}) - \overline{g}(e^{-t_2}). \tag{71}$$

Proof. For all $t \ge 0$, we have

$$\sum_{i,j}(\widetilde{\overline{V}}_{i,j}(t) - \overline{V}_{i,j}(t)) \le \sum_{i,j} j(\widetilde{\overline{V}}_{i,j}(t) - \overline{V}_{i,j}(t)) = \widetilde{\overline{S}}(t) - \overline{S}(t).$$

Thus,

$$\begin{aligned}\left|\overline{C}\right| &= \sum(\overline{V}_{k,l}(\overline{T}_1^*-) - \overline{V}_{k,l}(\overline{T}_2^*-)) = \sum(\widetilde{\overline{V}}_{k,l}(\overline{T}_1^*-) - \widetilde{\overline{V}}_{k,l}(\overline{T}_2^*-)) + o_p(n)\\ &= n\overline{g}(e^{-\overline{T}_1^*}) - n\overline{g}(e^{-\overline{T}_2^*}) + o_p(n).\end{aligned}$$

□

Let $\overline{C}'$ be the set of possible influence sources traced up till $\overline{T}_1$ and $\overline{C}''$ be the set of those traced between $\overline{T}_1$ and $\overline{T}_2$. Then, by Lemma 4.5, we have that

$$\frac{\left|\overline{C}'\right|}{n} \xrightarrow{p} 0 \tag{72}$$

and

$$\frac{\left|\overline{C}''\right|}{n} \xrightarrow{p} \overline{g}(1) - \overline{g}(e^{-\overline{\tau}}) = 1 - \overline{g}(e^{-\overline{\tau}}). \tag{73}$$

Evidently, $\overline{C}'' \subset \overline{C}(y_n)$ and $\overline{C}(y_n) \subset \overline{C}' \cup \overline{C}''$, therefore

$$\left|\overline{C}''\right| \leq \left|\overline{C}(y_n)\right| \leq \left|\overline{C}'\right| + \left|\overline{C}''\right| \tag{74}$$

and thus, from (72) and (73),

$$\frac{\left|\overline{C}(y_n)\right|}{n} \xrightarrow{p} 1 - \overline{g}(e^{-\overline{\tau}}), \tag{75}$$

which completes the proof.

□

Proof of Theorem 2.6. As in the previous section, we have the following set of Lemmmas, which we state without proof since the only change is notational. As before, assumptions of Theorem 2.2 continue to hold.

Lemma 4.6. *$\forall \epsilon > 0$, let*

$$\overline{\mathbb{A}}(\epsilon) := \left\{ x \in v(G^*(n, (d_i)_1^n)) : \frac{\left|\overline{C}(x)\right|}{n} \geq \epsilon \text{ and } \left|\frac{\left|\overline{C}(x)\right|}{n} - (1 - \overline{g}(\overline{\xi}))\right| \geq \epsilon \right\}.$$

Then,

$$\forall \epsilon, \quad \frac{\left|\overline{\mathbb{A}}(\epsilon)\right|}{n} \xrightarrow{p} 0. \tag{76}$$

Lemma 4.7. *For every $\epsilon > 0$, let*

$$\overline{\mathbb{B}}(\epsilon) := \left\{ x \in \overline{C}' \cup \overline{C}'' : \left|\overline{C}(x)\right| / n \geq \epsilon \text{ and } \left|\overline{C}(x) \vartriangle \overline{C}^*\right| / n \geq \epsilon \right\}. \tag{77}$$

Then,

$$\forall \epsilon, \quad \frac{\left|\overline{\mathbb{B}}(\epsilon)\right|}{n} \xrightarrow{p} 0. \tag{78}$$

Lemma 4.8. *Let $\overline{T}_3$ be the first time after $\overline{T}_2$ that $\overline{C}1$ is performed and let w_n be the sleeping vertex activated at this moment. If $\overline{C}'''$ is the set of informer vertices reached between $\overline{T}_2$ and $\overline{T}_3$, then*

$$\frac{\left|\overline{C}'''\right|}{n} \xrightarrow{p} 0. \tag{79}$$

Lemma 4.9. *For every $\epsilon > 0$, let*

$$\overline{\mathbb{C}}(\epsilon) := \left\{ w \in \left(\overline{C}' \cup \overline{C}''\right)^c : \left|\overline{C}(w)\right|/n \geq \epsilon \text{ and } \left|\overline{C}(w) \vartriangle \overline{C}^*\right|/n \geq \epsilon \right\}. \tag{80}$$

Then, we have that

$$\forall \epsilon, \quad \frac{\left|\overline{\mathbb{C}}(\epsilon)\right|}{n} \xrightarrow{p} 0. \tag{81}$$

Finally, Lemma 4.7 and Lemma 4.9 allow us to conclude that

$$\forall \epsilon, \quad \frac{\left|\overline{\mathbb{C}}^s(\epsilon)\right| + \left|\overline{\mathbb{C}}^L(\epsilon)\right|}{n} \xrightarrow{p} 1. \tag{82}$$

□

5 Duality Relation

The forward and backward processes are linked through the tautology: $y \in C(x) \iff x \in \overline{C}(y)$. To prove the Theorem 2.7, we consider the double sum: $\sum_{x,y \in v(G(n,(d_i)_1^n))} \mathbf{1}(y \in C(x))$.

From here onwards, we abridge $v(G(n,(d_i)_1^n))$ to $v(G)$. Assumptions of Theorem 2.2 continue to hold throughout this section. We start with the following Proposition.

Proposition 5.1. *We have,*

$$\mathbf{A}_n := \left| n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(y \in C(x)) - n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(x \in \overline{C}^*)\mathbf{1}(y \in C(x) \cap C^*) \right|$$
$$\xrightarrow{p} 0,$$

when $n \to \infty$.

Proof. The Proposition follows from the following two Lemmas.

Lemma 5.2. *For any $\epsilon > 0$ and $n \to \infty$,*

$$\left| n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(y \in C(x)) - n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(y \in C(x) \cap C^*) \right| \leq 2\epsilon + R_n^1(\epsilon),$$

where $R_n^1(\epsilon) \xrightarrow{p} 0$.

Proof. For $\epsilon > 0$, we have

$$\begin{aligned}
& \left| n^{-2} \sum_{x,y} \mathbf{1}(y \in C(x)) - n^{-2} \sum_{x,y} \mathbf{1}(y \in C(x) \cap C^*) \right| \\
\leq \quad & n^{-2} \sum_{x} \min(|C(x)|, |C(x) \vartriangle C^*|) \\
= \quad & n^{-2} \sum_{x \in \mathbb{C}^s(\epsilon)} \min\left(|C(x)|, |C(x) \vartriangle C^*|\right) \\
& + n^{-2} \sum_{x \in \mathbb{C}^L(\epsilon)} \min\left(|C(x)|, |C(x) \vartriangle C^*|\right) \\
& + n^{-2} \sum_{x \notin \mathbb{C}^s(\epsilon) \cup \mathbb{C}^L(\epsilon)} \min\left(|C(x)|, |C(x) \vartriangle C^*|\right) \\
\leq \quad & n^{-1} \sum_{x \in \mathbb{C}^s(\epsilon)} \epsilon + n^{-1} \sum_{x \in \mathbb{C}^L(\epsilon)} \epsilon + n^{-1} \sum_{x \notin \mathbb{C}^s(\epsilon) \cup \mathbb{C}^L(\epsilon)} 1 \\
\leq \quad & \epsilon + \epsilon + \left(1 - \frac{|\mathbb{C}^s(\epsilon)| + |\mathbb{C}^L(\epsilon)|}{n}\right).
\end{aligned}$$

Taking $R_n^1(\epsilon) := 1 - \frac{|\mathbb{C}^s(\epsilon)| + |\mathbb{C}^L(\epsilon)|}{n}$ and using Theorem 2.3, we conclude the proof. □

Lemma 5.3. *For any $\epsilon > 0$ and $n \to \infty$,*

$$\left| n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(y \in C(x) \cap C^*) - n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(x \in \overline{C}^*) \mathbf{1}(y \in C(x) \cap C^*) \right|$$
$$\leq 2\epsilon + R_n^2(\epsilon),$$

where $R_n^2(\epsilon) \xrightarrow{p} 0$.

Proof. Since $y \in C(x) \iff x \in \overline{C}(y)$, we have

$$\sum_{x,y \in v(G)} \mathbf{1}(y \in C(x) \cap C^*) = \sum_{x,y \in v(G)} \mathbf{1}(x \in \overline{C}(y)) \mathbf{1}(y \in C^*) \tag{83}$$

and

$$\sum_{x,y} \mathbf{1}(x \in \overline{C}^*)\mathbf{1}(y \in C(x) \cap C^*) = \sum_{x,y} \mathbf{1}(x \in \overline{C}(y) \cap \overline{C}^*)\mathbf{1}(y \in C^*). \quad (84)$$

Consequently,

$$\begin{aligned}
&\left| n^{-2} \sum_{x,y} \mathbf{1}(y \in C(x) \cap C^*) - n^{-2} \sum_{x,y} \mathbf{1}(x \in \overline{C}^*)\mathbf{1}(y \in C(x) \cap C^*) \right| \\
&\leq n^{-2} \sum_{y} \mathbf{1}(y \in C^*) \min\left(|\overline{C}(y)|, |\overline{C}(y) \vartriangle \overline{C}^*| \right) \\
&\leq n^{-2} \sum_{y} \min\left(|\overline{C}(y)|, |\overline{C}(y) \vartriangle \overline{C}^*| \right).
\end{aligned}$$

The result follows by the arguments similar to those in the proof of Lemma 5.2, with $R_n^2(\epsilon) := 1 - \frac{|\overline{\mathbb{C}}^s(\epsilon)| + |\overline{\mathbb{C}}^L(\epsilon)|}{n}$. □

□

Next, we have the following two Propositions, which lead to Theorem 2.7.

Proposition 5.4. *For any $\epsilon > 0$ and $n \to \infty$,*

$$\left| n^{-1}|\mathbb{C}^L(\epsilon)| - n^{-1}|\mathbb{C}^L(\epsilon) \cap \overline{C}^*| \right| \leq \alpha^1 \epsilon + R_n^3(\epsilon), \quad (85)$$

where $\alpha^1 > 0$ is a constant and $R_n^3(\epsilon) \xrightarrow{p} 0$. Analogously,

$$\left| n^{-1}|\overline{\mathbb{C}}^L(\epsilon)| - n^{-1}|\overline{\mathbb{C}}^L(\epsilon) \cap C^*| \right| \leq \alpha^2 \epsilon + R_n^4(\epsilon) \quad (86)$$

where $\alpha^2 > 0$ is a constant and $R_n^4(\epsilon) \xrightarrow{p} 0$.

Proof. Remark that

$$\begin{aligned}
\sum_{x,y \in v(G)} \mathbf{1}(y \in C(x)) = &\sum_{x \in v(G), y \in \overline{\mathbb{C}}^L(\epsilon)} \mathbf{1}(x \in \overline{C}(y)) + \sum_{x \in v(G), y \in \overline{\mathbb{C}}^s(\epsilon)} \mathbf{1}(x \in \overline{C}(y)) \\
&+ \sum_{x \in v(G), y \notin \overline{\mathbb{C}}^s(\epsilon) \cup \overline{\mathbb{C}}^L(\epsilon)} \mathbf{1}(x \in \overline{C}(y)).
\end{aligned}$$

Therefore, using the arguments similar to those in the proof of Lemma 5.2, we have

$$\left| n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(y \in C(x)) - n^{-2} |\overline{C}^*|.|\overline{\mathbb{C}}^L(\epsilon)| \right| \leq 2\epsilon + R_n^2(\epsilon). \qquad (87)$$

In the same way,

$$\left| n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(y \in C^*) \mathbf{1}(x \in \overline{C}(y) \cap \overline{C}^*) - n^{-2} |\overline{C}^*|.|\overline{\mathbb{C}}^L(\epsilon) \cap C^*| \right| \leq 2\epsilon + R_n^2(\epsilon).$$

From the above two equations and using Proposition 5.1, we have

$$\left| n^{-2} |\overline{C}^*|.|\overline{\mathbb{C}}^L(\epsilon)| - n^{-2} |\overline{C}^*|.|\overline{\mathbb{C}}^L(\epsilon) \cap C^*| \right| \leq 4\epsilon + 2R_n^2(\epsilon) + A_n.$$

Now using Theorem 2.2 and taking $\alpha^2 := \frac{5}{1-\overline{g}(\overline{\xi},\overline{\xi})}$ and $R_n^4(\epsilon) = \frac{3R_n^2(\epsilon)+2A_n}{1-\overline{g}(\overline{\xi},\overline{\xi})}$, we have the second part of the proposition. The proof of the first part is similar, with $\alpha^1 := \frac{5}{1-g(\xi,\xi)}$ and $R_n^3(\epsilon) = \frac{3R_n^1(\epsilon)+2A_n}{1-g(\xi,\xi)}$.

□

Proposition 5.5. *For any $\epsilon > 0$,*

$$\left| n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(y \in C(x)) - n^{-2} |C^* \cap \overline{\mathbb{C}}^L(\epsilon)|.|\mathbb{C}^L(\epsilon)| \right| \leq 3\epsilon + R_n^1(\epsilon) + R_n^2(\epsilon)$$

Proof. We can upper bound the double sum thus,

$$\begin{aligned} \sum_{x,y \in v(G)} \mathbf{1}(y \in C(x)) \leq & \sum_{x \in \mathbb{C}^L(\epsilon), y \in \overline{\mathbb{C}}^L(\epsilon)} \mathbf{1}(y \in C(x)) \\ & + \sum_{x \in \mathbb{C}^s(\epsilon), y \in v(G)} \mathbf{1}(y \in C(x)) \\ & + \sum_{x \in v(G), y \in \overline{\mathbb{C}}^s(\epsilon)} \mathbf{1}(y \in C(x)) \\ & + \sum_{x \notin \mathbb{C}^s(\epsilon) \cup \mathbb{C}^L(\epsilon), y \in v(G)} \mathbf{1}(y \in C(x)) \\ & + \sum_{x \in v(G), y \notin \mathbb{C}^s(\epsilon) \cup \mathbb{C}^L(\epsilon)} \mathbf{1}(y \in C(x)). \end{aligned}$$

The result follows, once again, by using the arguments similar to those in the proof of Lemma 5.2.

□

Now, from Proposition 5.5 and (86) and (87) from Proposition 5.4, we can conclude the proof of Theorem 2.7, with $\alpha := 5 + \alpha^1$ and $R_n(\epsilon) := R_n^1(\epsilon) + 2R_n^2(\epsilon) + R_n^4(\epsilon)$.

The Corollary 2.8 follows from Theorem 2.7 and Proposition 5.4.

Acknowledgements

We thank René Schott for introducing us to the influence propagation dynamic analysed in this paper through a pre-print of [4] and thus motivating this study. We also thank Marc Lelarge for his useful suggestions regarding the analytical tools for exploration on Configuration Model and pointing us to [11], which has heavily influenced our approach.

References

[1] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.

[2] Hamed Amini, Moez Draief, and Marc Lelarge. Marketing in a random network. *Proc. Network Control & Optimization*, LNCS 5425:17–25, 2009.

[3] Tom Britton, Svante Janson, and Anders Martin-Löf. Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. *Advances in Applied Probability*, 39(4):922–948, 2007.

[4] Francis Comets, Francois Delarue, and René Schott. Information Transmission under Random Emission Constraints. hal-00637304, 2011. http://hal.archives-ouvertes.fr/hal-00637304.

[5] Emilie Coupechoux and Marc Lelarge. How clustering affects epidemics in random networks. arXiv:1202.4974, 2012. http://arxiv.org/abs/1202.4974.

[6] Emilie Coupechoux and Marc Lelarge. Diffusion of innovations in random clustered networks with overlapping communities. arXiv:1303.4325, 2013. http://arxiv.org/abs/1303.4325v1.

[7] Moez Draief and Laurent Massoulié. *Epidemics and Rumors in Complex Networks*, volume 369 of *London Mathematical Society Lecture Notes*. Cambridge University Press, 2010.

[8] Richard Durrett. *Random graph dynamics*, volume 20. Cambridge university press, 2007.

[9] Nikolaos Fountoulakis and Konstantinos Panagiotou. Rumor spreading on random regular graphs and expanders. *Random Structures & Algorithms*, 43(2):201–220, 2013.

[10] Amit Goyal, Francesco Bonchi, and Laks V. S. Lakshmanan. A data-based approach to social influence maximization. *Proc. VLDB Endow.*, 5(1):73–84, September 2011.

[11] Svante Janson and Malwina J. Luczak. A new approach to the giant component problem. *Random Structures and Algorithms*, 34(2):197–216, 2008.

[12] Marc Lelarge. Diffusion and cascading behavior in random networks. *Games and Economic Behavior*, 75(2):752 – 775, 2012.

[13] Mark EJ Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.

[14] Remco Van Der Hofstad. *Random graphs and complex networks*. 2009. Available on http://www.win.tue.nl/rhofstad/NotesRGCN.pdf.