

Flooding Time in edge-Markovian Dynamic Graphs*

Andrea E.F. Clementi[†] Claudio Macci[†] Angelo Monti[‡]
Francesco Pasquale^{†*} Riccardo Silvestri[‡]

May 30, 2008

Abstract

We introduce stochastic *time-dependency* in evolving graphs: starting from an arbitrary initial edge probability distribution, at every time step, every edge changes its state (existing or not) according to a two-state Markovian process with probabilities p (*edge birth-rate*) and q (*edge death-rate*). If an edge exists at time t then, at time $t + 1$, it dies with probability q . If instead the edge does not exist at time t , then it will come into existence at time $t + 1$ with probability p .

Such evolving graph model is a wide generalization of time-independent dynamic random graphs [6] and will be called *edge-Markovian dynamic graphs*.

We investigate the speed of information dissemination in such dynamic graphs. We provide nearly tight bounds (which in fact turn out to be tight for a wide range of probabilities p and q) on the completion time of the *flooding* mechanism aiming to broadcast a piece of information from a source node to all nodes. In particular, we provide: *i*) A tight characterization of the class of edge-Markovian dynamic graphs where flooding time is constant and, thus, it does not asymptotically depend on the initial probability distribution. *ii*) A tight characterization of the class of edge-Markovian dynamic graphs where flooding time does not asymptotically depend on the edge death-rate q .

*Partially supported by the EU under the EU/IST Project 15964 AEOLUS.

[†]Dipartimento di Matematica, Università di Roma “Tor Vergata”

[clementi,macci,pasquale]@mat.uniroma2.it

[‡]Dipartimento di Informatica, “La Sapienza” Università di Roma

[monti,silvestri]@di.uniroma1.it

1 Introduction

The network model. There is a growing interest in the study of graphs that evolve over time. Such dynamic structures arise from several areas such as computer networks, networks of users exchanging e-mail or instant messages, citation networks, hyperlinks networks, peer-to-peer networks, social networks (who-trust-whom, who-talks-to-whom, etc.), and many other more [7]. One important line of research is the development of models that well approximate some important specific features observed in real dynamic networks.

One of these observed features is the strong dependence between the existence (or the absence) of a link between two nodes at a given time step and the existence (or the absence) of the same link at the previous time step. Important examples of this behavior arise in faulty communication networks, mobile networks, peer-to-peer networks, and social networks [7, 6, 17, 11, 18, 19]¹.

We introduce and study stochastic *time-dependency* in evolving graphs: starting from an *arbitrary* initial edge probability distribution, at every time step, every edge changes its state (existing or not) according to a two-state Markovian process with probabilities $p(n)$ and $q(n)$ where n is the number of nodes. If an edge exists at time t then, at time $t + 1$, it dies with probability $q(n)$. If instead the edge does not exist at time t then it will come into existence at time $t + 1$ with probability $p(n)$.

For brevity's sake, functions $p(n)$ and $q(n)$ will be simply denoted as p and q , respectively². Formally, let $n \in \mathbb{N}$ and $0 \leq p, q \leq 1$, we say that the sequence of graphs (we use here notation $[n] = \{1, \dots, n\}$)

$$\mathcal{G} = \{G_t = ([n], E_t) : t \in \mathbb{N}\}$$

is *Markov* $\mathcal{G}(n, p, q, g)$ (in symbols $\mathcal{G} \sim \mathcal{G}(n, p, q, g)$) if

$$E_t = \left\{ e \in \binom{[n]}{2} : X_t(e) = 1 \right\}$$

where $\{X_t(e) : e \in \binom{[n]}{2}\}$ are independent Markov chains with transition matrix

$$M = \left(\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1-p & p \\ 1 & q & 1-q \end{array} \right)$$

and $g = g(n)$ is an arbitrary initial probability distribution over the set $\binom{[n]}{2}$ yielding E_0 . Notation $g = \emptyset$ will indicate the initial probability distribution assigning probability 0 to every edge (i.e. $E_0 = \emptyset$). A Markov $\mathcal{G}(n, p, q, g)$ sequence will be called *edge-Markovian dynamic graph*.

Observe that setting $q = 1 - p$ yields the (time-independent) *dynamic random graph* model [6] where links, at every time, are chosen independently at random. So, our dynamic graph model is a wide generalization of dynamic random graphs. Other time-independent dynamic random networks have been studied in the context of dynamic *gossip-based mechanisms* [12, 11, 18]. Very recently, a general Markovian (evolving) graph model has been introduced in [1]. The

¹Notice that, in some of these settings, there is an underlying physical network that supports the abstraction of point-to-point communication.

²Hence, any inequality $p \leq (\geq) b(n)$ means that $p(n)$ is eventually not larger (not smaller) than $b(n)$. The same holds for $q = q(n)$.

results in [1] concern covering and hitting time of random walks in specific, restricted cases of this dynamic model that are not related with ours.

The flooding mechanism. In modern network architectures, *global node outreach*, i.e. reaching all nodes from a source/initiator one, is a major challenge. Reaching all nodes from a given source/initiator node is typically required to disseminate or retrieve information: this task is performed via suitable protocols that aim to achieve low delay and message overhead. However, when the network topology is highly dynamic and unknown, (e.g. unstructured peer-to-peer networks, faulty/mobile networks, etc), it is very hard to design efficient protocols for that task and, as a result, the *flooding mechanism* is adopted [4, 9, 10, 14]. In the flooding mechanism, every *informed* node (i.e. any node that has the source message) always sends the source message to all its neighbors. So, the source is clearly informed since the beginning and any other node v gets informed at time step t if any of its neighbors in E_t is informed at time step $t - 1$.

The completion time of the flooding mechanism (or, simply, *flooding time*) is the first time step in which all nodes of the network are informed. Given any edge-Markovian dynamic graph $\mathcal{G} \sim \mathcal{G}(n, p, q, g)$, the random variable $T(\mathcal{G})$ is defined as the maximum flooding time in \mathcal{G} over all possible choices of source $s \in [n]$.

It is important to observe that flooding time in dynamic networks plays the same role of *diameter* in static networks. Flooding time and diameter represent “natural” lower bounds for broadcast protocols in dynamic networks and static ones, respectively. For this reason, flooding is often used in order to evaluate the relative efficiency of alternative protocols, especially in networks with unknown dynamic topology [4, 9, 17].

Our results. We study *flooding time* on edge-Markovian dynamic graphs for all possible functions p and q and provide asymptotic bounds that are tight or nearly tight.

- For any $0 \leq p, q \leq 1$ and for *any* initial probability distribution g , we prove that for $\mathcal{G} \sim \mathcal{G}(n, p, q, g)$, it holds *with high probability*³ (*w.h.p.*) that

$$T(\mathcal{G}) = O\left(\frac{\log n}{\log(1 + np)}\right) \quad (1)$$

- Then, we prove that for any $0 \leq p, q \leq 1$ and for initial probability distribution $g = \emptyset$, it holds *w.h.p.* that

$$T(\mathcal{G}) = \Omega\left(\frac{\log n}{np}\right) \quad (2)$$

- When $p \geq \log n/n$, $0 \leq q \leq 1$, and $g = \emptyset$, we prove that

$$T(\mathcal{G}) = \Omega\left(\frac{\log n}{\log(1 + np)}\right) \quad (3)$$

We emphasize that, when the edge-Markovian dynamic graph is “almost dense”, i.e. when $p \geq 1/n^\delta$ for some constant $0 < \delta < 1$ and arbitrary q , Bound (1) implies that flooding time is constant, so it does not asymptotically depend on the initial probability distribution g . Another important consequence of Bound (1) is that flooding time is logarithmic even for $p = o(\log n/n)$ (whatever is q), i.e. even below the connectivity threshold⁴ of (static) *random*

³Here in the sequel, the term *with high probability* means that the event holds with probability at least $1 - (1/n)^\alpha$, for some constant $\alpha > 0$.

⁴Notice that below this threshold, flooding time in static random graphs is clearly infinite.

graphs $\mathcal{G}_{n,p}$ [3]. In particular, flooding time is logarithmic even when the expected node degree is 1. By comparing this bound with flooding time (i.e. diameter) of static random graphs $\mathcal{G}_{n,p}$, it turns out that random dynamic changes significantly speed-up information dissemination even when such changes are not time-independent but follow an arbitrary Markov law.

For $p \geq \log n/n$, Bounds (1) and (3) are *tight* in the standard “computational-complexity” sense: the upper bound holds for *any* initial probability distribution while the lower bound is satisfied by *at least one* initial probability distribution and they are asymptotically tight. By comparing Bounds (1) and (2), we can state the same tightness for $p \leq 1/n$ since, in this case, $np = \Theta(\log(1 + np))$.

An important consequence of the above tight results is that the “death-rate” q might play an asymptotically-relevant role only when $1/n \leq p \leq \log n/n$. However, when $\log \log n/n \leq p \leq \log n/n$, we prove that w.h.p. $T(\mathcal{G}) = \Omega(\frac{\log n}{\log \log n})$ for $g = \emptyset$. By combining this lower bound with Bounds (1) and (2) we get that flooding time lies w.h.p. in the “small” ranges

- $c_1 \frac{\log n}{\log \log n} \leq T(\mathcal{G}) \leq c_2 \log n$
for $\frac{1}{n} \leq p \leq \frac{\log \log n}{n}$
- $c_1 \frac{\log n}{\log \log n} \leq T(\mathcal{G}) \leq c_2 \frac{\log n}{\log \log \log n}$
for $\frac{\log \log n}{n} \leq p \leq \frac{\log n}{n}$

where c_1 and c_2 are some positive constants. In the same range $1/n \leq p \leq \log n/n$, we have also investigated the special interesting case of *densifying* graphs [7, 13], i.e. the case $q = 0$, where an existing edge never dies. We get the following tight⁵ bounds that hold w.h.p.:

$$T(\mathcal{G}) = \Theta\left(\frac{\log n}{np}\right) \quad \text{for } p \leq \frac{\log \log n}{n} \quad (4)$$

$$T(\mathcal{G}) = \Theta\left(\frac{\log n}{\log \log n}\right) \quad \text{for } \frac{\log \log n}{n} \leq p \leq \frac{\log n}{n} \quad (5)$$

Rather interestingly, the second bound shows that flooding time in that range has a *plateau* behaviour: it is not a decreasing function in p .

Related works. A large number of previous works deals with the issue of flooding (and broadcasting) in communication networks. However, we are not aware of analytical results concerning flooding in time-dependent stochastic models of dynamic networks. Some works deal with dynamic random graphs where edge probability at a given time is fully independent from the previous topology [6, 12, 11, 18]. In particular, in [11, 18] it is proved that flooding is w.h.p. completed in $O(\log n)$ steps even when every node is connected to a constant number of random neighbors. Furthermore, radio broadcasting in (time-independent) dynamic random graphs has been recently studied in [6]: at every time step, an independent graph is selected according to the standard random graph model $\mathcal{G}_{n,p}$ [3]. They show that, despite the interference phenomenon, *fully-independence* yields w.h.p. logarithmic broadcasting time for any $p \geq 1/n$. We also mention the analytical study of highly-dynamic networks in [6, 16] where edges are managed by a *worst-case* adversary.

Road map. In Section 2, we present (upper) Bound (1); Section 3 is devoted to (lower) Bounds (2) and (3). In Section 4, we study densifying networks. Finally, open problems are discussed in Section 5.

⁵Tightness is intended in the same “computational-complexity” sense of our previous tight bounds.

2 The upper bound

In this section we prove the upper bound (1) stated in the Introduction. The proof has two main ingredients: (i) expansion properties of edge-Markovian dynamic graphs (proved in Subsection 2.1) that are then used (ii) to evaluate the average-rate of new informed nodes as time goes on (Subsection 2.2).

2.1 Expansion properties

Let us consider an edge-Markovian dynamic graph $\mathcal{G} \sim \mathcal{G}(n, p, q, g)$. Let $I \subseteq [n]$. For $t_0, t \in \mathbb{N}$, define the set of nodes (not in I) that have been connected to I in at least one time step $i \in \{t_0 + 1, \dots, t_0 + t\}$

$$H_{t_0, t}(I) = \left\{ v \in [n] \setminus I : \begin{array}{l} \{u, v\} \in E_i \\ \text{for some } u \in I \text{ and for} \\ \text{some } i = t_0 + 1, \dots, t_0 + t \end{array} \right\} \quad (6)$$

For brevity's sake, we will use $H_{t_0, t}$ instead of $H_{t_0, t}(I)$ when set I is clear from the context. The following Lemma is a key-ingredient in proving our upper bound (1): roughly speaking, it states that the probability that a non-informed node is not informed within time step t decreases *exponentially* in t , even though there is a Markovian dependence between the graph topologies of consecutive time-steps.

Lemma 2.1 *Let $v \in [n] \setminus I$, then, for any $t_0 \geq 0$ and $t \geq 1$, it holds*

$$\mathbf{P}(v \notin H_{t_0, t}) \leq [(1 - \xi)(1 - p)^{t-1}]^m$$

where $\xi = \min\{p, 1 - q\}$ and $m = |I|$.

Proof. Let $u \in I$ and $e_u = \{u, v\}$. Then

$$\begin{aligned} \mathbf{P}\left(\bigcap_{i=t_0+1}^{t_0+t} \{e_u \notin E_i\}\right) &= \mathbf{P}(e_u \notin E_{t_0+1}) \prod_{i=t_0+2}^{t_0+t} \mathbf{P}\left(e_u \notin E_i \mid \bigcap_{j=t_0+1}^{i-1} \{e_u \notin E_j\}\right) \\ &= \mathbf{P}(e_u \notin E_{t_0+1}) \prod_{i=t_0+2}^{t_0+t} \mathbf{P}(e_u \notin E_i \mid e_u \notin E_{i-1}) \\ &= \mathbf{P}(e_u \notin E_{t_0+1}) (1 - p)^{t-1} \end{aligned}$$

Furthermore,

$$\begin{aligned} \mathbf{P}(e_u \notin E_{t_0+1}) &= \mathbf{P}(e_u \notin E_{t_0+1} \mid e_u \notin E_{t_0}) \mathbf{P}(e_u \notin E_{t_0}) + \mathbf{P}(e_u \notin E_{t_0+1} \mid e_u \in E_{t_0}) \mathbf{P}(e_u \in E_{t_0}) \\ &= (1 - p) \mathbf{P}(e_u \notin E_{t_0}) + q (1 - \mathbf{P}(e_u \notin E_{t_0})) \\ &= q + (1 - p - q) \mathbf{P}(e_u \notin E_{t_0}) \\ &\leq \max\{q, 1 - p\} \leq 1 - \xi \end{aligned}$$

We thus get

$$\mathbf{P}\left(\bigcap_{i=t_0+1}^{t_0+t} \{e_u \notin E_i\}\right) \leq (1 - \xi)(1 - p)^{t-1}$$

Since Markov chains of different edges are independent, it holds that

$$\begin{aligned}\mathbf{P}(v \notin H_{t_0,t}) &= \mathbf{P}\left(\bigcap_{u \in I} \left(\bigcap_{i=t_0+1}^{t_0+t} \{e_u \notin E_i\}\right)\right) \\ &= \prod_{u \in I} \mathbf{P}\left(\bigcap_{i=t_0+1}^{t_0+t} \{e_u \notin E_i\}\right) \\ &\leq [(1-\xi)(1-p)^{t-1}]^m\end{aligned}$$

□

In what follows we will always assume that $p+q \leq 1$. Case $p+q > 1$ can be managed similarly by using a simple fact shown in Appendix A. When $p+q \leq 1$ then $\xi = p$ and the above lemma implies that

$$\mathbf{P}(v \notin H_{t_0,t}) \leq (1-p)^{mt} \quad (7)$$

In the next three lemmas, we define $m = |I|$.

Lemma 2.2 *Let $1 \leq m \leq \frac{n}{2}$, $t_0 \geq 0$, and $t \geq c \frac{\log n}{nmp}$ for some constant $c > 0$. Then, it holds that*

$$\mathbf{P}(|H_{t_0,t}| \leq \alpha \log n) \leq \frac{1}{n^\gamma}$$

where $\alpha = c/8$ and $\gamma = c/32$.

Proof. Let $v \in [n] \setminus I$, then (7) implies that

$$\mathbf{P}(v \notin H_{t_0,t}) \leq (1-p)^{mt} \leq e^{-pmt}$$

Hence

$$\begin{aligned}\mathbf{P}(v \in H_{t_0,t}) &\geq 1 - e^{-pmt} \geq 1 - e^{-c \frac{\log n}{n}} \\ &\geq \frac{c \frac{\log n}{n}}{1 + c \frac{\log n}{n}} \geq \frac{c \log n}{2n}\end{aligned}$$

where we used the lemma's hypothesis on t and Bound (22) in the Appendix.

Fix $t \geq 1$, for any $v \in [n] \setminus I$ define the random variable Y_v so that $Y_v = 1$ if $v \in H_{t_0,t}$ and $Y_v = 0$ otherwise. We thus have that $|H_{t_0,t}| = \sum_{v \in [n] \setminus I} Y_v$, and we obtain the following inequality for its expected value

$$\mathbf{E}[|H_{t_0,t}|] = \sum_{v \in [n] \setminus I} \mathbf{E}[Y_v] \geq (n-m) \frac{c \log n}{2n} \geq \frac{c}{4} \log n$$

Since the Markov chains of different edges are independent, random variables $\{Y_v : v \in [n] \setminus I\}$ are independent as well. By applying Chernoff's bound (Lemma B.2) with $\mu = \frac{c}{4} \log n$ and $\varepsilon = 1/2$, we get

$$\mathbf{P}\left(|H_{t_0,t}| \leq \frac{c}{8} \log n\right) \leq e^{-\frac{c}{32} \log n}$$

□

Lemma 2.3 *Assume that $t_0 \geq 0$, $1 \leq m \leq \frac{n}{2}$, and $mp \geq 1$. Then it holds that*

$$\mathbf{P}(|H_{t_0,1}| \leq \alpha n) \leq e^{-\gamma n}$$

with $\alpha = 1/8$ and $\gamma = 1/32$.

Proof. Let $v \in [n] \setminus I$, Inequality (7) implies

$$\mathbf{P}(v \notin H_{t_0,1}) \leq (1-p)^m \leq e^{-mp} \leq \frac{1}{e} \leq \frac{1}{2}$$

where we exploit Lemma's hypothesis $mp \geq 1$. Then

$$\mathbf{P}(v \in H_{t_0,1}) \geq \frac{1}{2}$$

As in the proof of Lemma 2.2, we can give a lower bound on the expected size of H_1

$$\mathbf{E}[|H_{t_0,1}|] \geq \frac{1}{2}(n-m) \geq \frac{1}{4}n$$

where we used lemma's hypothesis $m \leq n/2$. By applying Chernoff's bound (Lemma B.2) with $\mu = \frac{1}{4}n$ and $\varepsilon = 1/2$ we finally obtain

$$\mathbf{P}\left(|H_{t_0,1}| \leq \frac{1}{8}n\right) \leq e^{-\frac{1}{32}n}$$

□

Lemma 2.4 *Assume that $t_0 \geq 0$, $c\frac{\log n}{np} \leq m \leq \frac{n}{2}$ for some constant $c > 0$, and $mp \leq 1$. Then,*

$$\mathbf{P}(|H_{t_0,1}| \leq \alpha nmp) \leq \frac{1}{n^\gamma}$$

where $\alpha = 1/8$ and $\gamma = c/32$.

Proof. Let $v \in [n] \setminus I$. As in the proof of the previous Lemma, we have that

$$\mathbf{P}(v \in H_{t_0,1}) \geq 1 - e^{-mp} \geq \frac{mp}{1+mp} \geq \frac{1}{2}mp$$

where we used $mp \leq 1$. So, it holds that

$$\mathbf{E}[|H_{t_0,1}|] \geq \frac{1}{2}mp(n-m) \geq \frac{1}{4}nmp$$

where we used $m \leq n/2$. Again, by applying Chernoff's bound (Lemma B.2) with $\varepsilon = 1/2$ and $\mu = 1/4nmp$, we get

$$\mathbf{P}\left(|H_{t_0,1}| \leq \frac{1}{8}nmp\right) \leq e^{-\frac{1}{32}nmp} \leq e^{-\frac{c}{32}\log n}$$

where we used hypothesis $nmp \geq c \log n$. □

2.2 Flooding time: the upper bound

Thanks to the expansion properties shown in the previous subsection, we are now able to investigate the flooding process in an edge-Markovian dynamic graph $\mathcal{G} \sim \mathcal{G}(n, p, q, g)$.

The *flooding* process can be represented by the sequence $\{I_t : t \in \mathbb{N}\}$ where $I_t \subseteq [n]$ is the subset of nodes defined recursively as follows

$$\begin{cases} I_0 &= \{s\} \\ I_{t+1} &= I_t \cup H_{t,1}(I_t) \end{cases}$$

Nodes in I_t are thus the *informed* nodes at time step t and we define $m_t = |I_t|$.

Notice that from the above definition and from the definition of set $H_{t_0,t}$ in (6), it holds that

1. For any $t \in \mathbb{N}$, $I_{t-1} \subseteq I_t$;
2. For any $t_0, t \in \mathbb{N}$, $I_{t_0} \cup H_{t_0,t}(I_{t_0}) \subseteq I_{t_0+t}$.

In order to evaluate the size of I_{t_0+t} , we will always use Property 2 above. Flooding is completed within time t if $I_t = [n]$, and *flooding time* is defined as the *smallest* t for which flooding is completed.

Proposition 2.5 *After $t = O\left(\frac{\log n}{\log(1+np)}\right)$ time steps, the number m_t of informed node is at least βn w.h.p., where β is a positive constant.*

Proof. We consider two cases. Let c be a positive constant.

Case 1. $p \leq 2c\frac{\log n}{n^2}$. We “organize” the flooding process in *phases*. Every phase is a finite sequence of consecutive time steps. Let M_k be the number of informed node at the beginning of Phase k (we set $M_0 = 1$). From Lemma 2.2, if $M_k \leq n/2$ then, after t_k time steps with

$$t_k = \left\lceil c \frac{\log n}{npM_k} \right\rceil \quad (8)$$

there will be at least $M_k + \alpha \log n$ informed nodes, w.h.p. The value t_k will denote the length of Phase k .

Notice that if $p \leq 2c\frac{\log n}{n^2}$ then $c\frac{\log n}{npM_k} \geq 1$, and so

$$c \frac{\log n}{npM_k} \leq t_k = \left\lceil c \frac{\log n}{npM_k} \right\rceil \leq 2c \frac{\log n}{npM_k}$$

Now, consider the recurrence

$$M_{k+1} \geq \alpha \log n + M_k \quad (9)$$

Its explicit form, since $M_0 = 1$, is

$$M_k \geq k\alpha \log n \quad (10)$$

The recurrence (9) holds for every $M_k < n/2$, w.h.p. Then, it is possible to prove that Inequality (10) holds w.h.p. at least for those k we are interested in, i.e., until $k = \hat{k} = O\left(\frac{n}{\log n}\right)$ so that $M_k \geq n/2$. This is formally stated in next claim.

Claim 1 *For every $k \leq \frac{1}{2\alpha} \frac{n}{\log n}$, it holds that $M_k \geq k\alpha \log n$, w.h.p.*

Proof. Define the events

$$\begin{aligned} \mathcal{E}_k &= \text{“ } M_k \geq \alpha \log n + M_{k-1} \text{ ”} \\ \mathcal{Z}_k &= \text{“ } M_{k-1} < n/2 \text{ ”} \\ \mathcal{F}_k &= \text{“ } M_k < k\alpha \log n \text{ ”} \end{aligned}$$

Let us first observe how the above events are mutually related. It holds that

$$\mathcal{Z}_k \subseteq \mathcal{Z}_{k-1} \quad (11)$$

Indeed, if the informed nodes are less than $n/2$ in phase k , then they were less than $n/2$ in phase $k-1$ too. Moreover, it holds that

$$\bigcap_{i=1}^k \mathcal{E}_i \subseteq \overline{\mathcal{F}_k} \quad (12)$$

Indeed, if Recurrence (9) holds for every phase from 1 to k , then its explicit form (10) must hold. As for complementary sets, (12) implies that

$$\mathcal{F}_k \subseteq \overline{\left(\bigcap_{i=1}^k \mathcal{E}_i\right)} = \bigcup_{i=1}^k \overline{\mathcal{E}_i} \quad (13)$$

Observe also that, for every $k \leq \frac{n}{2\alpha \log n}$, $\mathcal{F}_k \subseteq \mathcal{Z}_k$ and hence $\mathcal{F}_k = \mathcal{F}_k \cap \mathcal{Z}_k$.

Finally, Lemma 2.2 implies that

$$\mathbf{P}(\overline{\mathcal{E}_k} \mid \mathcal{Z}_k) \leq \frac{1}{n^\gamma}$$

with $\gamma = c/32$ and c positive constant that we can choose arbitrarily. Thanks to the above facts, we can conclude that, for every $k \leq \frac{n}{2\alpha \log n}$,

$$\begin{aligned} \mathbf{P}(\mathcal{F}_k) &= \mathbf{P}(\mathcal{F}_k \cap \mathcal{Z}_k) \\ \text{From (13)} &\leq \mathbf{P}\left(\left(\bigcup_{i=1}^k \overline{\mathcal{E}_i}\right) \cap \mathcal{Z}_k\right) = \mathbf{P}\left(\bigcup_{i=1}^k (\overline{\mathcal{E}_i} \cap \mathcal{Z}_k)\right) \leq \sum_{i=1}^k \mathbf{P}(\overline{\mathcal{E}_i} \cap \mathcal{Z}_k) \\ \text{From (11)} &\leq \sum_{i=1}^k \mathbf{P}(\overline{\mathcal{E}_i} \cap \mathcal{Z}_i) = \sum_{i=1}^k \mathbf{P}(\overline{\mathcal{E}_i} \mid \mathcal{Z}_i) \mathbf{P}(\mathcal{Z}_i) \leq \sum_{i=1}^k \mathbf{P}(\overline{\mathcal{E}_i} \mid \mathcal{Z}_i) \\ \text{From Lemma 2.2} &\leq k \frac{1}{n^\gamma} \leq \frac{n}{2\alpha \log n} \frac{1}{n^\gamma} \leq \frac{1}{n^{\gamma-1}} \end{aligned}$$

□

After $\hat{k} = \frac{1}{2\alpha} \frac{n}{\log n}$ phases, there are thus at least $n/2$ informed nodes w.h.p. Since the length of Phase k is t_k , the overall time to get at least $n/2$ informed nodes is

$$\begin{aligned} \sum_{k=1}^{\hat{k}} t_k &= \sum_{k=1}^{\hat{k}} \left\lceil c \frac{\log n}{npM_k} \right\rceil \leq \sum_{k=1}^{\hat{k}} 2c \frac{\log n}{npM_k} \\ &\leq \sum_{k=1}^{\hat{k}} 2c \frac{\log n}{np\alpha k \log n} = \frac{2c}{\alpha np} \sum_{k=1}^{\hat{k}} \frac{1}{k} \\ &\leq \frac{4c}{\alpha} \frac{1}{np} \log \hat{k} \leq \hat{c} \frac{\log n}{np} \end{aligned}$$

for a suitable constant $\hat{c} > 0$.

Since $np \geq \log(1 + np)$ for any choice of $p > 0$, after $O\left(\frac{\log n}{\log(1+np)}\right)$ time steps, there are at least $n/2$ informed nodes w.h.p.

Case 2. $p \geq 2c \frac{\log n}{n^2}$. Observe that we cannot follow the same line of reasoning of the first case. Indeed, we cannot achieve $n/2$ informed nodes from (8): when M_k grows, the term $c \frac{\log n}{npM_k}$ becomes smaller than 1 and we cannot apply the inequality

$$t_k = \left\lceil c \frac{\log n}{npM_k} \right\rceil \leq 2c \frac{\log n}{npM_k}$$

We thus need to analyze the flooding process in three consecutive periods.

- Period 1: If $M_k \leq c \frac{\log n}{np}$ then Recurrence (9) holds w.h.p.⁶. We now make use of the following claim whose proof is similar to Claim 1.

⁶The actual condition would be $M_k \leq \min\{c \frac{\log n}{np}, \frac{n}{2}\}$ but we have $c \frac{\log n}{np} \leq \frac{n}{2}$ since $p \geq 2c \frac{\log n}{n^2}$.

Claim 2 For every $k \leq \frac{c}{\alpha np}$, it holds that $M_k \geq k\alpha \log n$, w.h.p.

Proof. Define the events \mathcal{E}_k and \mathcal{F}_k as in Claim 1, i.e.,

$$\begin{aligned}\mathcal{E}_k &= \text{“ } M_k \geq \alpha \log n + M_{k-1} \text{”} \\ \mathcal{F}_k &= \text{“ } M_k < k\alpha \log n \text{”}\end{aligned}$$

and events \mathcal{Z}_k as follows

$$\mathcal{Z}_k = \text{“ } M_k < c \frac{\log n}{np} \text{”}$$

Observe that Inclusions (11), (12), and (13) hold as well. By using Lemma 2.2 and reasoning like in the proof of Claim 1, it is easy to see that, for each $k \leq \frac{c}{\alpha np}$, it holds

$$\mathbf{P}(\mathcal{F}_k) \leq \frac{1}{n^{\gamma-1}}$$

□

So, after $\hat{k} = \frac{c}{\alpha np}$ phases, there are at least $c \frac{\log n}{np}$ informed nodes w.h.p. Then, the overall number of time steps will be

$$\sum_{k=1}^{\hat{k}} t_k = \hat{c} \frac{\log\left(1 + \frac{1}{np}\right)}{np}$$

- Period 2: If $c \frac{\log n}{np} \leq M_k \leq \min\{1/p, n/2\}$, Lemma 2.4 implies that the recurrence

$$M_{k+1} \geq (1 + \alpha np)M_k \tag{14}$$

holds w.h.p. where the phase length $t_k = 1$ for any $k > 0$. Its explicit form is

$$M_k \geq (1 + \alpha np)^k M_0 \geq (1 + \alpha np)^k \tag{15}$$

Next claim can be proved similarly to Claims 1 and 2.

Claim 3 For every $k \leq \frac{\log(\min\{1/p, n/2\})}{\log(1 + \alpha np)}$, it holds that $M_k \geq (1 + \alpha np)^k$, w.h.p.

Proof. Define the events

$$\begin{aligned}\mathcal{E}_k &= \text{“ } M_k \geq (1 + \alpha np)M_{k-1} \text{”} \\ \mathcal{F}_k &= \text{“ } M_k < (1 + \alpha np)^k \text{”} \\ \mathcal{Z}_k &= \text{“ } M_{k-1} < \min\{1/p, n/2\} \text{”}\end{aligned}$$

Observe that Inclusions (11), (12), and (13) hold as well, and Lemma 2.4 implies

$$\mathbf{P}(\bar{\mathcal{E}}_k | \mathcal{Z}_k) \leq \frac{1}{n^\gamma}$$

with $\gamma = c/32$ and c being a positive constant we can choose arbitrarily.

Reasoning like in previous claim proofs, it is easy to see that, for every $k \leq \frac{\log(\min\{1/p, n/2\})}{\log(1 + \alpha np)}$, it holds

$$\mathbf{P}(\mathcal{F}_k) \leq \frac{1}{n^{\gamma-1}}$$

□

It follows that, after $\frac{\log(\min\{1/p, n/2\})}{\log(1+\alpha np)} = O\left(\frac{\log n}{\log(1+np)}\right)$ time steps, at least $\min\{1/p, n/2\}$ nodes are w.h.p. informed.

- **Period 3:** Observe that if $\min\{1/p, n/2\} = n/2$ (i.e. $p \leq 2/n$), then after Period 2, at least $n/2$ nodes are informed and the proposition is proved. Otherwise the number of informed nodes w.h.p. satisfies the hypothesis of Lemma 2.3: Indeed $1 \leq M_{\hat{k}} \leq n/2$ and $M_{\hat{k}}p \geq 1$. So, thanks to Lemma 2.3, after next time step, at least αn nodes will be informed w.h.p., where α is a positive constant.

Finally, even in this second case ($p \geq 2c\frac{\log n}{n^2}$), $O\left(\frac{\log n}{\log(1+np)}\right)$ time steps suffice to get βn informed nodes, for some constant $\beta > 0$. \square

Proposition 2.6 *Let $0 < \beta < 1$ be any constant and $m \geq \beta n$ be the number of informed nodes at time step $t_0 \geq 0$. Then, after $O\left(\frac{\log n}{np}\right)$ time steps, all nodes will be informed w.h.p.*

Proof. Let I be the set of informed nodes and $|I| = m \geq \beta n$. Let $t \geq c\frac{\log n}{np}$ with c be a constant to be fixed later. Consider set $H_{t_0, t}$ defined as in (6). For any $v \in [n] \setminus I$, it holds that

$$\mathbf{P}(v \notin H_{t_0, t}) = (1-p)^{mt} \leq e^{-pmt} \leq e^{-\beta c \log n} = \frac{1}{n^{\beta c}}$$

where the last inequality follows from the hypothesis on m and t . Then the probability there exists a node $v \in [n] \setminus I$ such that $v \notin H_{t_0, t}$ can be bounded as follows

$$\begin{aligned} \mathbf{P}(\exists v \in [n] \setminus I : v \notin H_{t_0, t}) &\leq \mathbf{P}\left(\bigcup_{v \in [n] \setminus I} \{v \notin H_{t_0, t}\}\right) \\ &\leq \sum_{v \in [n] \setminus I} \mathbf{P}(v \notin H_{t_0, t}) \\ &\leq (n-m) \frac{1}{n^{\beta c}} \leq (1-\beta)n \frac{1}{n^{\beta c}} \leq \frac{1}{n^{\beta c - 1}} \end{aligned}$$

Finally, the thesis follows by fixing $c > 1/\beta$. \square

Theorem 2.7 *Let $\mathcal{G} \sim \mathcal{G}(n, p, q, g)$ be any edge-Markovian dynamic graph. Then, flooding time is w.h.p. $O\left(\frac{\log n}{\log(1+np)}\right)$.*

Proof. Proposition 2.5 states that, after the first $O\left(\frac{\log n}{\log(1+np)}\right)$ time steps, at least βn nodes will be informed w.h.p. Then, thanks to Proposition 2.6, after further $O\left(\frac{\log n}{np}\right)$ time steps, all nodes will be informed w.h.p. Since $np \geq \log(1+np)$, we get

$$O\left(\frac{\log n}{np}\right) \subseteq O\left(\frac{\log n}{\log(1+np)}\right)$$

that completes the proof. \square

3 The lower bounds

In this section, we will prove lower Bounds (2) and (3) stated in the Introduction. We will make use of the following ‘‘monotonicity’’ property: flooding time decreases as birth-rate p increases while it increases as death-rate q increases.

Proposition 3.1 Consider two edge-Markovian dynamic graphs $\mathcal{G} \sim \mathcal{G}(n, p, q, \emptyset)$ and $\mathcal{G}' \sim \mathcal{G}(n, p', q', \emptyset)$ where

$$p \leq p'; \quad q' \leq q; \quad p + q' \leq 1$$

Then, it holds that

$$\mathbf{P}(T(\mathcal{G}) \leq t) \leq \mathbf{P}(T(\mathcal{G}') \leq t) \quad \text{for any } t \in \mathbb{N}$$

Observe that the above result requires the somewhat “artificial” condition $p + q' \leq 1$. Indeed, its proof relies on a suitable application (stated in next lemma) of the *coupling* technique of Markov chains [15] that works only under the above condition. Proving the monotonicity property without this condition by using any other technique is an open problem which seems far from easy.

Lemma 3.2 (Markov-coupling Lemma.) Let $\{Z_t = (X_t, Y_t) : t \in \mathbb{N}\}$ be a Markov chain with state space $\{(0, 0), (0, 1), (1, 1)\}$ and transition matrix

$$M^Z = \left(\begin{array}{c|ccc} & (0, 0) & (0, 1) & (1, 1) \\ \hline (0, 0) & 1 - p' & p' - p & p \\ (0, 1) & q' & 1 - p - q' & p \\ (1, 1) & q' & q - q' & 1 - q \end{array} \right) \quad (16)$$

where $p \leq p'$, $q \geq q'$ and $p + q' \leq 1$. Then, $\{X_t : t \in \mathbb{N}\}$ and $\{Y_t : t \in \mathbb{N}\}$ are Markov chains with state space $\{0, 1\}$ and transition matrices, respectively,

$$M^X = \left(\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1 - p & p \\ 1 & q & 1 - q \end{array} \right) \quad M^Y = \left(\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1 - p' & p' \\ 1 & q' & 1 - q' \end{array} \right)$$

Proof. By definition of Markov chain it holds

$$\mathbf{P}((X_{t+1}, Y_{t+1}) = j \mid (X_t, Y_t) = i_t, \dots, (X_0, Y_0) = i_0) = \mathbf{P}((X_{t+1}, Y_{t+1}) = j \mid (X_t, Y_t) = i_t)$$

for every $j, i_0, \dots, i_t \in \{(0, 0), (0, 1), (1, 1)\}$. So in particular we have

$$\mathbf{P}(X_{t+1} = k \mid (X_t, Y_t) = i_t, \dots, (X_0, Y_0) = i_0) = \mathbf{P}(X_{t+1} = k \mid (X_t, Y_t) = i_t) \quad (17)$$

$$\mathbf{P}(Y_{t+1} = k \mid (X_t, Y_t) = i_t, \dots, (X_0, Y_0) = i_0) = \mathbf{P}(Y_{t+1} = k \mid (X_t, Y_t) = i_t) \quad (18)$$

for every $k \in \{0, 1\}$ and for every $i_0, \dots, i_t \in \{(0, 0), (0, 1), (1, 1)\}$. Now observe that for every $k_0, \dots, k_t, k_{t+1} \in \{0, 1\}$

$$\begin{aligned} & \mathbf{P}(X_{t+1} = k_{t+1} \mid X_t = k_t, \dots, X_0 = k_0) = \\ &= \sum_{h_0, \dots, h_t \in \{0, 1\}} \mathbf{P} \left(X_{t+1} = k_{t+1} \mid \begin{array}{l} (X_t, Y_t) = (k_t, h_t), \dots, \\ (X_0, Y_0) = (k_0, h_0) \end{array} \right) \cdot \\ & \quad \cdot \mathbf{P}(Y_t = h_t, \dots, Y_0 = h_0 \mid X_t = k_t, \dots, X_0 = k_0) \\ &= \sum_{h_0, \dots, h_t \in \{0, 1\}} \mathbf{P}(X_{t+1} = k_{t+1} \mid (X_t, Y_t) = (k_t, h_t)) \cdot \\ & \quad \cdot \mathbf{P}(Y_t = h_t, \dots, Y_0 = h_0 \mid X_t = k_t, \dots, X_0 = k_0) \end{aligned}$$

where in the last equality we used (17). If $\mathbf{P}(X_{t+1} = k_{t+1} \mid (X_t, Y_t) = (k_t, h_t))$ does not depend on h_t , we have the Markov property for $\{X_t : t \geq 0\}$.

By proceeding in the same way for $\{Y_t : t \geq 0\}$ (and using (18) instead of (17)), if $\mathbf{P}(Y_{t+1} = h_{t+1} \mid (X_t, Y_t) = (k_t, h_t))$ does not depend on k_t , then Markov property holds for this random process as well. In conclusion, we have to carry out the following two checks, starting from the given transition matrix M^Z :

1. $\sum_{h'} M_{(k,h),(k',h')}^Z$ does not depend on h and equals $M_{k,k'}^X$;
2. $\sum_{k'} M_{(k,h),(k',h')}^Z$ does not depend on k and equals $M_{h,h'}^Y$.

Check 1:

- $(k, k') = (0, 0)$:
 - $\langle h = 0 \rangle$: $M_{(0,0),(0,0)}^Z + M_{(0,0),(0,1)}^Z = 1 - p' + p' - p = 1 - p$;
 - $\langle h = 1 \rangle$: $M_{(0,1),(0,0)}^Z + M_{(0,1),(0,1)}^Z = q' + 1 - p - q' = 1 - p$;
- $(k, k') = (0, 1)$:
 - $\langle h = 0 \rangle$: $M_{(0,0),(1,1)}^Z = p$;
 - $\langle h = 1 \rangle$: $M_{(0,1),(1,1)}^Z = p$;
- $(k, k') = (1, 0)$:
 - $\langle h = 1 \rangle$: $M_{(1,1),(0,0)}^Z + M_{(1,1),(0,1)}^Z = q' + q - q' = q$;
- $(k, k') = (1, 1)$:
 - $\langle h = 1 \rangle$: $M_{(1,1),(1,1)}^Z = 1 - q$.

Check 2:

- $(h, h') = (0, 0)$:
 - $\langle k = 0 \rangle$: $M_{(0,0),(0,0)}^Z = 1 - p'$;
- $(h, h') = (0, 1)$:
 - $\langle k = 0 \rangle$: $M_{(0,0),(0,1)}^Z + M_{(0,0),(1,1)}^Z = p' - p + p = p'$;
- $(h, h') = (1, 0)$:
 - $\langle k = 0 \rangle$: $M_{(0,1),(0,0)}^Z = q'$;
 - $\langle k = 1 \rangle$: $M_{(1,1),(0,0)}^Z = q'$;
- $(h, h') = (1, 1)$:
 - $\langle k = 0 \rangle$: $M_{(0,1),(0,1)}^Z + M_{(0,1),(1,1)}^Z = 1 - p - q' + p = 1 - q'$;
 - $\langle k = 1 \rangle$: $M_{(1,1),(0,1)}^Z + M_{(1,1),(1,1)}^Z = q - q' + 1 - q = 1 - q'$.

□

Proof of Proposition 3.1. Consider the family

$$\left\{ Z_t(e) : e \in \binom{[n]}{2} \right\}$$

of independent Markov chains where, for every $e \in \binom{[n]}{2}$, $Z_t(e) = (X_t(e), Y_t(e))$ has transition matrix (16) and initial probability distribution $\mathbf{P}(Z_0(e) = (0, 0)) = 1$. Define

$$\begin{aligned} E_t^X &= \left\{ e \in \binom{[n]}{2} : X_t(e) = 1 \right\} \\ E_t^Y &= \left\{ e \in \binom{[n]}{2} : Y_t(e) = 1 \right\} \end{aligned}$$

and $\mathcal{G}^X = \{([n], E_t^X) : t \in \mathbb{N}\}$ and $\mathcal{G}^Y = \{([n], E_t^Y) : t \in \mathbb{N}\}$. Thanks to Lemma 3.2, \mathcal{G}^X and \mathcal{G}^Y are edge-Markovian dynamic graphs with the same probability distributions of \mathcal{G} and \mathcal{G}' , respectively. By definition of Markov Chain $\{Z_t : t \in \mathbb{N}\}$, we have $E_t^X \subseteq E_t^Y$ for any $t \geq 0$. It thus holds that $T(\mathcal{G}^X) \geq T(\mathcal{G}^Y)$. Hence, for any $t \in \mathbb{N}$, we obtain

$$\mathbf{P}(T(\mathcal{G}) \leq t) = \mathbf{P}(T(\mathcal{G}^X) \leq t) \leq \mathbf{P}(T(\mathcal{G}^Y) \leq t) = \mathbf{P}(T(\mathcal{G}') \leq t)$$

□

Observation 3.3 *In order to prove a general lower bound on flooding time, we can prove it in the special case $q = 0$ and then apply Proposition 3.1 to extend it to any $q \in [0, 1]$.*

Notice that, in $\mathcal{G} \sim \mathcal{G}(n, p, 0, \emptyset)$, for each $e \in \binom{[n]}{2}$ and $t \in \mathbb{N}$, it holds that

$$\mathbf{P}(e \in E_t) = 1 - (1 - p)^t$$

Hence, we get the following equivalence result.

Observation 3.4 *Consider any edge-Markovian dynamic graph $\mathcal{G} \sim \mathcal{G}(n, p, 0, \emptyset)$, then random variable G_t and random variable \mathcal{G}_{n, p_t} (i.e. random graph [3]) where $p_t = 1 - (1 - p)^t$ are identically distributed.*

We can thus use the following properties of random graphs

Property 3.5 [[8],[5]] *For random graph $\mathcal{G}_{n, p}$:*

1. *If $p \leq \frac{1}{2} \frac{\log n}{n}$, then w.h.p. $\mathcal{G}_{n, p}$ is not connected;*
2. *There is a constant $c > 1$ such that for $p \geq c \frac{\log n}{n}$, $\mathcal{G}_{n, p}$ is connected and has diameter $\Theta\left(\frac{\log n}{\log(np)}\right)$ w.h.p.*

Lemma 3.6 *If $\mathcal{G} \sim \mathcal{G}(n, p, 0, \emptyset)$, then $T(\mathcal{G}) = \Omega\left(\frac{\log n}{np}\right)$, w.h.p.*

Proof. As for $p > \frac{\log n}{n}$, it is sufficient to note that $T(\mathcal{G}) \geq 1 > \frac{\log n}{np}$. Assume now $p \leq \frac{\log n}{n}$. We will prove that in this case there is w.h.p. a node v such that v and the source node are in different components in each graph G_t where $t \leq \tilde{t} = \left\lfloor \frac{1}{4} \frac{\log n}{np} \right\rfloor$. This implies that flooding w.h.p. requires more than $\tilde{t} = \Omega\left(\frac{\log n}{np}\right)$ time steps.

Notice that $E_{t-1} \subseteq E_t$ for $q = 0$. So, in order to prove the existence of node v , it is sufficient to prove that $G_{\tilde{t}}$ is w.h.p. not connected. Moreover, from Observation 3.4, this is equivalent to prove that $\mathcal{G}_{n,p_{\tilde{t}}}$ is w.h.p. not connected. This is a consequence of Property 3.5-1 since, for $p \leq \frac{\log n}{n}$, it holds

$$\begin{aligned} p_{\tilde{t}} &= 1 - (1-p)^{\tilde{t}} \leq 1 - e^{-\frac{p\tilde{t}}{1-p}} \leq \frac{p\tilde{t}}{1-p} \\ &\leq \frac{1}{4} \frac{\log n}{n} \frac{1}{1-p} \leq \frac{1}{2} \frac{\log n}{n} \end{aligned}$$

where in the last inequality we use $\frac{1}{1-p} \leq 2$ that holds because $p \leq \frac{\log n}{n}$. \square

Lemma 3.7 *If $\mathcal{G} \sim \mathcal{G}(n, p, 0, \emptyset)$ and $p \geq \frac{\log n}{n}$, then $T(\mathcal{G}) = \Omega\left(\frac{\log n}{\log(np)}\right)$, w.h.p.*

Proof. Let c be the constant occurring in Property 3.5-2. For $p > \frac{1}{2c}$ it is sufficient to note that for n sufficiently large we have

$$T(\mathcal{G}) \geq 1 > \frac{\log n}{2c \log\left(\frac{n}{2c}\right)} \geq \frac{\log n}{2c \log(np)} = \Omega\left(\frac{\log n}{\log(np)}\right)$$

Assume now $p \leq \frac{1}{2c}$. We use the following simple

Claim 4 *If $q = 0$ and flooding terminates in t steps, then G_t has diameter at most $2t$.*

Proof. We will prove that any informed node in G_t has distance at most t from the source. We argue by induction on t . For $t = 0$ the claim is obvious so assume it holds for $t - 1$ and prove it for t . Let v be an informed node in G_t . If v is informed also in G_{t-1} , then the thesis follows by inductive hypothesis noticing that $E_{t-1} \subseteq E_t$ for $q = 0$. If on the contrary v is informed for the first time in G_t , then v has an informed neighbour in G_{t-1} . By inductive hypothesis this neighbor has distance at most $t - 1$ from the source in G_{t-1} . Thus v has distance from the source at most t in G_{t-1} and since $E_{t-1} \subseteq E_t$ for $q = 0$ the same is true in G_t . \square

Due this claim, in order to prove the lemma we show that w.h.p. graphs G_t , with $t \leq \tilde{t} = \left\lfloor 4c \frac{\log n}{\log(np)} \right\rfloor$, are not connected or have diameter $\Omega\left(\frac{\log n}{\log(np)}\right)$. Since $E_t \subseteq E_{t+1}$ for $q = 0$, it is sufficient to show that $G_{\tilde{t}}$ has w.h.p. diameter $\Omega\left(\frac{\log n}{\log(np)}\right)$. Moreover, from Observation 3.4, this is equivalent to prove that $\mathcal{G}_{n,p_{\tilde{t}}}$ has this property. This follows from Property 3.5-2: indeed, for $\frac{\log n}{n} \leq p \leq \frac{1}{2c}$, we get

$$\begin{aligned} p_{\tilde{t}} &= 1 - (1-p)^{\tilde{t}} \geq 1 - e^{-p\tilde{t}} \geq \frac{p\tilde{t}}{1+p\tilde{t}} \\ &\geq c \frac{\log n}{n} \frac{4 \log n}{2 \log(np) + 2 \log n} \geq c \frac{\log n}{n} \end{aligned} \tag{19}$$

Thus $\mathcal{G}_{n,p_{\tilde{t}}}$ has w.h.p. diameter

$$d \geq c' \frac{\log n}{\log(np_{\tilde{t}})} \tag{20}$$

for a given constant c' .

Moreover, from (19), $\frac{\log n}{n} \leq p \leq \frac{1}{2c}$ and n sufficiently large, we have

$$\begin{aligned} p_{\tilde{t}} &= 1 - (1-p)^{\tilde{t}} \leq 1 - e^{-\frac{p\tilde{t}}{1-p}} \leq \frac{p\tilde{t}}{1-p} \\ &\leq \frac{3}{2}p \left(\frac{4c \log n}{\log(np)} + 1 \right) \leq \frac{3}{2}p \left(\frac{4c \log n}{\log \log n} + 1 \right) \\ &\leq p \log n \leq np^2 \end{aligned}$$

Hence, from (20) and the above inequality, we conclude

$$d \geq c' \frac{\log n}{\log(np)^2} = \Omega\left(\frac{\log n}{\log(np)}\right)$$

□

We can now consider arbitrary death-rate q . Indeed, from Observation 3.3, Proposition 3.1, Lemma 3.6, and Lemma 3.7, we obtain the lower bounds stated in the Introduction.

Theorem 3.8 *If $\mathcal{G} \sim \mathcal{G}(n, p, q, \emptyset)$, then $T(\mathcal{G}) = \Omega\left(\frac{\log n}{np}\right)$, w.h.p.*

Theorem 3.9 *If $\mathcal{G} \sim \mathcal{G}(n, p, q, \emptyset)$ and $p \geq \frac{\log n}{n}$, then $T(\mathcal{G}) = \Omega\left(\frac{\log n}{\log(1+np)}\right)$, w.h.p.*

4 Densifying networks

We now consider the case $q = 0$, i.e. densifying networks.

Lemma 4.1 *If $\mathcal{G} \sim \mathcal{G}(n, p, 0, \emptyset)$ and $p \leq \frac{\log \log n}{n}$, then $T(\mathcal{G}) = \Theta\left(\frac{\log n}{np}\right)$ w.h.p.*

Proof. The lower bound follows from Lemma 3.6. As for the upper bound, notice that $E_{t-1} \subseteq E_t$ for $q = 0$. Hence, if at a given step t the graph G_t is connected and has diameter d , then at step $t+d$ flooding is surely complete. Thus, we can prove the Lemma by showing that graph $G_{\tilde{t}}$, $\tilde{t} = \left\lceil 4c \frac{\log n}{np} \right\rceil$ (where c is the constant occurring in Property 3.5-2) is connected and has diameter $O\left(\frac{\log n}{np}\right)$, w.h.p. From Observation 3.4, this is equivalent to prove that $\mathcal{G}_{n, p_{\tilde{t}}}$ is connected and has diameter $O\left(\frac{\log n}{np}\right)$, w.h.p. For n sufficiently large, we get

$$\begin{aligned} p_{\tilde{t}} &= 1 - (1-p)^{\tilde{t}} \geq 1 - e^{-p\tilde{t}} \geq \frac{p \cdot \tilde{t}}{1 + p\tilde{t}} \\ &\geq c \frac{\log n}{n} \cdot \frac{4}{2 + \frac{4c \log n}{n}} \geq c \frac{\log n}{n} \end{aligned}$$

So, Property 3.5-2 implies that $\mathcal{G}_{n, p_{\tilde{t}}}$ is connected and has diameter $d \leq c' \frac{\log n}{\log(np_{\tilde{t}})}$ for some constant c' , w.h.p. Finally, from (21) and $p \leq \frac{\log \log n}{n}$, we have

$$d \leq c' \frac{\log n}{\log \log n} \leq c' \frac{\log n}{np} = O\left(\frac{\log n}{np}\right)$$

□

Lemma 4.2 *If $\mathcal{G} \sim \mathcal{G}(n, p, 0, \emptyset)$ and $\frac{\log \log n}{n} \leq p \leq \frac{\log n}{n}$, then $T(\mathcal{G}) = \Theta\left(\frac{\log n}{\log \log n}\right)$ w.h.p.*

Proof. Let $\mathcal{G} \sim \mathcal{G}(n, p, 0, \emptyset)$ and $\mathcal{G}' \sim \mathcal{G}(n, p', 0, \emptyset)$ be two edge-Markovian dynamic graphs where $p \leq p'$. Proposition 3.1 implies that, for each $t \geq 1$ and $p \leq p'$,

$$\mathbf{P}(T(\mathcal{G}) \leq t) \leq \mathbf{P}(T(\mathcal{G}') \leq t)$$

Thus, to prove the Lemma, it is sufficient to show that

$$T(\mathcal{G}) = \begin{cases} \Omega\left(\frac{\log n}{\log \log n}\right) & \text{if } p = \frac{\log n}{n} \\ O\left(\frac{\log n}{\log \log n}\right) & \text{if } p = \frac{\log \log n}{n} \end{cases}$$

The lower bound follows from Lemma 3.7 since for $p = \frac{\log n}{n}$ we get $\frac{\log n}{\log(np)} = \frac{\log n}{\log \log n}$. The upper bound follows from Lemma 4.1 since for $p = \frac{\log \log n}{n}$ we get $\frac{\log n}{np} = \frac{\log n}{\log \log n}$. \square

5 Open problems

An important consequence of our results is that flooding time does not asymptotically depend on the “death-rate” q whenever p does not fall into the range $[1/n, \log n/n]$. An interesting open question is thus that of determining the dependence of flooding time w.r.t. q inside that range.

We also proved that whenever the edge-Markovian dynamic graph is almost dense, i.e. when $p \geq (1/n)^\delta$, flooding time does not depend on the initial probability distribution. Another challenging research issue is to study flooding time, as function of the initial probability distribution, in edge-Markovian dynamic graphs that are not almost dense. A relevant scenario is that yielded by choosing the initial graph according to the *stationary* distribution of the Markov chain.

As for more general research issues, we mention that of investigating constrained communication mechanisms in our dynamic network model: *multi-access* channel (e.g the radio network model [2]), and the telephone-call model [12].

References

- [1] C. Avin, M. Koucký, and Z. Lotker. How to Explore a Fast-Changing World. In *Proc. of 35th International Colloquium on Automata, Languages and Programming (ICALP)*, to appear, 2008.
- [2] R. Bar-Yehuda, O. Goldreich, and A. Itai. On the time-complexity of broadcast in multi-hop radio networks: an exponential gap between determinism and randomization. *Journal of Computer and System Sciences*, 45(1):104-126, 1992
- [3] B. Bollobás. *Random Graphs*. Cambridge University Press, 2001.
- [4] N.B. Chang and M. Liu. Optimal controlled flooding search in a large wireless network. In *Proc. of 3rd International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WIOPT)*, pages 229-237, 2005.
- [5] F. Chung and L. Lu. The Diameter of Sparse Random Graphs. *Advances in Applied Mathematics*, 26:257-279, 2001
- [6] A.E.F. Clementi, A. Monti, F. Pasquale, and R. Silvestri. Communication in Dynamic Radio Networks. In *Proc. of 26th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC)*, pages 205-214. ACM Press, 2007.
- [7] S.N. Dorogovtsev and J.F.F. Mendes. *Evolution of Networks*. Oxford University Press, 2003.
- [8] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5:17-61, 1960.
- [9] M.M.C. Gkantsidis and A. Saberi. Hybrid search schemes for unstructured peer-to-peer networks. In *Proc. of 24th INFOCOM*, pages 1526-1537. IEEE Computer Society, 2005.
- [10] Gnutella. Gnutella rfc. <http://rfc-gnutella.sourceforge.net>, 2002
- [11] S.M. Hedetniemi, S.T. Hedetniemi, and A.L. Liestman. A survey of gossiping and broadcasting in communication networks. *Networks*, 18(4):319-349, 1988.
- [12] D. Kempe and J. Kleinberg. Protocols and Impossibility Results for Gossip-Based Communication Mechanisms. In *Proc. of 43rd Symposium on Foundations of Computer Science (FOCS)*, pages 471-480. IEEE Computer Society, 2002.
- [13] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graphs over time: Densification laws, Shrinking diameters, and possible explanations. In *Proc. of 11th ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 177-187. ACM Press, 2005.
- [14] Q. Lv, P. Cao, E. Cohen, K. Li, and S. Shenker. Search and replication in unstructured peer-to-peer networks. In *Proc. of 16th International Conference on Supercomputing (ICS)*, pages 84-95, 2002.
- [15] M. Mitzenmacher and E. Upfal. *Probability and Computing*. Cambridge University Press, 2005.
- [16] R. O'Dell and R. Wattenhofer. Information Dissemination in Highly Dynamic Graphs. In *Proc. of ACM DIALM-POMC*, pages 104-110, 2005.

- [17] K. Oikonomou and I. Stavrakakis. Performace analysis of probabilistic flooding using random graphs. In *Proc. of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 1-6. IEEE Computer Society, 2007.
- [18] B. Pittel. On spreading a rumor. *SIAM Journal on Applied Mathematics*, 47(1):213-223, 1987.
- [19] C. Scheideler. Models and Techniques for Communication in Dynamic Networks. In *Proc. of 19th Symposium on Theoretical Aspects of Computer Science (STACS)*, volume 2285 of *LNCS*, pages 27-49. Springer, 2002.

Appendix

A Case $p + q > 1$

Going back to Lemma 2.1, we can see that if $\xi = \min\{p, 1 - q\} = 1 - q$ then the Lemma states that

$$\mathbf{P}(v \notin H_{t_0, t}) \leq [q(1 - p)^{t-1}]^m$$

for each $v \in [n] \setminus I$. Since $q \leq 1$, we get $[q(1 - p)^{t-1}]^m \leq (1 - p)^{m(t-1)}$; so

$$\mathbf{P}(v \notin H_{t_0, t+1}) \leq (1 - p)^{mt}$$

In other words, we can obtain the same inequalities of Section 2 just *waiting* one more time step, and of course this fact does not change the asymptotic behaviour of flooding time.

B Known inequalities

Lemma B.1 *If $|x| < 1$ then the following inequalities hold*

$$e^{-\frac{x}{1-x}} \leq 1 - x \leq e^{-x} \tag{21}$$

$$\frac{x}{1+x} \leq 1 - e^{-x} \leq x \tag{22}$$

Lemma B.2 (Chernoff's bound) *Let be $X = \sum_{i=1}^n X_i$ where X_1, \dots, X_n are independent Bernoulli random variables and let be $0 < \varepsilon < 1$. If $0 < \mu \leq \mathbf{E}[X]$, then it holds*

$$\mathbf{P}\{X \leq (1 - \varepsilon)\mu\} \leq e^{-\frac{\varepsilon^2}{2}\mu}$$