Information spreading in dynamic graphs

Andrea Clementi · Riccardo Silvestri · Luca Trevisan

Received: 13 October 2012 / Accepted: 15 May 2014 / Published online: 6 June 2014 © Springer-Verlag Berlin Heidelberg 2014

Abstract We present a general approach to study the flooding time (a measure of how fast information spreads) in dynamic graphs (graphs whose topology changes with time according to a random process). We consider arbitrary ergodic Markovian dynamic graph processes, that is, processes in which the topology of the graph at time $t$ depends only on its topology at time $t-1$ and which have a unique stationary distribution. The most well studied models of dynamic graphs are all Markovian and ergodic. Under general conditions, we bound the flooding time in terms of the mixing time of the dynamic graph process. We recover, as special cases of our result, bounds on the flooding time for the random trip model and the random path one; previous analysis techniques provided bounds only in restricted settings for such models. Our result also provides the first bound for the random waypoint model whose analysis had been an important open question. The bound is tight for the most realistic ranges of the network parameters.

Keywords Dynamic graphs · Flooding protocols · Markov chains · Mobility models

1 Introduction

A dynamic graph is a probabilistic process that describes a graph whose topology changes with time. Dynamic graphs are appropriate models of wireless networks, peer-to-peer networks, social networks, and so on [9,32,33]. There are several interesting problems on dynamic graph processes, for example load balancing, studied in [18,30]. Here, we are interested in the speed of information spreading, a question that can model the spread of disease, the broadcast of files on peer-to-peer networks, of memes in social networks, etc. [17,19,22,24,25].

The simplest model of information spreading is the process of flooding [4,11,14,16,19,20,24,28,29], which begins with one node in the network being given a certain piece of information, and then the application of a protocol in which, at each time step, every node that has the information spreads it to its neighbors. Recall that the neighborhood of a node changes with time, and so even though the flooding algorithm is deterministic, the process of information spreading is probabilistic.

Data propagation in mobile ad-hoc networks [7,26] is one of the main applications that has motivated the study of the flooding process in geometric models of dynamic graphs in which, at every time step, every node is mapped to a point in a metric space, and two nodes are connected if their distance is
smaller than a given communication radius. The underlying metric space is usually a bounded portion of the plane, for example a square or the unit disk, and the dynamics come from independent random walks performed by the individual nodes via local moves [12,13,16,22,25,28,29]. Thus, the position of each node is a random variable whose distribution at every time step is determined by one independent copy of a fixed Markov chain which has a unique stationary distribution (typically called the stationary positional distribution) [1]. For example, a representative model of this type is the random walk model: $n$ nodes are placed on a $m \times m$ grid; at each time step, every node $v$ independently moves to a point in the grid randomly chosen among the points adjacent to the one that $v$ occupied at the previous time step; at each time step, the edge $(u,v)$ is present in the dynamic graph if $u$ and $v$ are located within distance $r$ in the grid. In this model, the state set of the node’s Markov chain is the set of all points of the grid and the stationary positional distribution turns out to be essentially uniform [1]. The flooding time will depend on the initial locations, on $r$, $m$ and $n$. Usually, one is interested in a worst-case analysis with respect to the initial locations, and in a bound dependent only on $r$, $m$ and $n$. Known analyses of (several versions) of the random walk model rely rather strongly on the uniformity property of its stationary positional distribution.

The random waypoint model [6–8,26] is another classic model of networks of mobile agents, and it is probably the most well studied one. In this model, every node chooses a random destination point in the mobility space, then it travels over the shortest path till it reaches the destination, and so on. Some analytical properties of this model have been derived such as the mixing time, the stationary positional distribution, etc. [6,8,26]. However, bounding its flooding time (or any basic communication task such as routing, data collection, etc.) is still a fundamental open problem. The techniques adopted for the random walk model do not work in the random waypoint model, mainly because, in the latter model, the nodes spend long periods of time in deterministic trajectories and because the two models have very different stationary positional distributions. For instance, the stationary positional distribution of the random waypoint model over a square is far from uniform and it is highly biased towards the center of the square.

The two above families of models define a probabilistic process over the nodes, which then implies which pairs of nodes are connected by an edge. There are also models of probabilistic processes that are directly defined over the edges. A very general model is provided by Markovian Evolving Graphs (MEGs) introduced in [2] (see later for a formal definition). However, available techniques to analyze information spreading only concern very restricted subclasses of MEGs, such as that studied in [11], where the state (i.e. on/off) of the edges is ruled by independent copies of a simple two-state Markov chain.

1.1 Our work

1.1.1 General dynamic networks

We provide an upper bound to the flooding time of any Markovian Evolving Graph (MEG) [2] that admits a unique stationary distribution. Formally, let $V$ be a set of $n$ nodes. If we call $G_t = (V,E_t)$ the random variable describing the dynamic graph at time $t$, the process is a MEG if the sequence of random variables $G([n],\{E_t\}_{t\geq0}) = G_0,\ldots,G_t,\ldots$ is Markovian, that is, if the distribution of $G_t$ is completely determined by the distribution of $G_{t-1}$, via a transformation independent of $t$. The class MEG is very general and it subsumes all above-mentioned network models such as random walk and random waypoint: indeed, it is easy to verify that all such models yields a sequence of random graphs $G([n],\{E_t\}_{t\geq0})$ which is Markovian. We require the Markovian sequence to be ergodic, i.e., it converges to a unique stationary distribution, for every initial choice $G_0$. This stationary distribution is a probability distribution over $n$-nodes graphs and it is called the stationary (random) graph $G_n$. Observe that all relevant models of dynamic random graphs proposed in the literature are ergodic.

Our upper bound to the flooding time is a function of: i) the (worst-case) edge-probability in $G_n$ (that determines the expected density of the stationary graph) ii) the degree of independence among edges in $G_n$ and, iii) the mixing time of the Markov chain $G([n],\{E_t\}_{t\geq0})$. More precisely, given an edge $e$, a node $i$, and a node subset $A$, let $p(e)$ be the probability that $e$ exists in $G_n$ and let $e(i, A)$ be the binary random variable returning 1 if an edge exists connecting $i$ to some node in $A$ in $G_n$. Let $M$ be an upper bound on the mixing time of $G([n],\{E_t\}_{t\geq0})$. Let us assume that, for some arbitrary positive reals $\alpha$ and $\beta$, (i) $p(i, j) \geq \alpha$ for every edge $(i, j)$ and\footnote{With an abuse of notation, event probabilities such as $P(e(i, A) = 1)$ will be shortly denoted as $P(e(i, A))$.} (ii) $P(e(i, A) \cdot e(j, A)) \leq \beta P(e(i, A)) \cdot P(e(j, A))$ for arbitrary nodes $i$, $j$ and an arbitrary node subset $A$ (not containing $i$ or $j$) (informally speaking, during the dynamic-graph process, Conditions (i) and (ii) are “only” required to hold every $M$ time steps). Then, we prove that the flooding time is w.h.p.\footnote{We say that an event holds with high probability if it holds with probability at least $1 - 1/n$.}

\[ O\left(M \left(\frac{1}{n\alpha} + \frac{1}{\beta}\right)^2 \log n\right) \] (1)
Our methods can be applied to non-Markovian processes as well, although the results are more complex to state (see Sects. 2, 3). To apply our result to a specific model of dynamic networks, we need an upper bound \( M \) to the mixing time: several recent studies give bounds on mixing times \([1, 11, 26]\), and our result allows to use any such bound, as well as future such results, in order to bound the flooding time.

Conditions (i) and (ii) are rather mild. We can obtain strong bounds on the flooding time even given rather weak bounds on the density and independence parameters, say \( \alpha = \Theta(1/n) \) and \( \beta = O(\text{polylog}(n)) \); graph dynamics with such parameters may generate, at each time \( t \), a graph \( G_t \) that it is not even connected, and in which it may even be possible that a constant fraction of the nodes have degree zero. Our bound implies that even in such sparse and disconnected topologies, the flooding time can be just a poly-logarithmic factor away from the mixing time of the MEG.

### 1.1.2 Node-MEGs

Our general approach finds a natural application in a subclass of MEGs that we call node-Markovian Evolving Graphs, denoted as node-MEGs. In this class of dynamics, at every time step, every node has an associated state. The state of a node is an arbitrary piece of information that, for example, may represent geometric position, destination, trajectory phase, social role and so on. The requirements of the model is that, (i) for each node, the state of the node evolves with time as a Markov chain \( \mathcal{M} \); (ii) the states of different nodes evolves independently; and (iii) at each time step \( t \), the presence of an edge \((i, j)\) in the graph \( G_t \) is only a function of the state of node \( i \) and the state of node \( j \) at time \( t \) (and the function is the same for each pair of nodes \( i, j \)).

Node-MEGs are a rather general class of MEGs that includes every mobility model where nodes act independently over a discrete space (such as an arbitrary graph or any geometric region discretized, for instance, by using a grid of suitable resolution\(^3\)). Random walk, random waypoint, and random trip models \([7, 15, 26]\) have a natural formulation as node-MEGs: for instance, a formulation of the random waypoint as a node-MEGs is described in Sect. 4.2. Notice that in node-MEGs, nodes are indistinguishable so Condition (i) is easy-to-check: if it is satisfied in the stationary graph for a specific edge then it is satisfied for all edges. We then prove that if incident edges are almost pairwise independent in the stationary graph yielded by a node-MEG then Condition (ii) is satisfied for some constant \( \beta \): so, checking Condition (ii) can here be reduced to check pairwise independence of the edges incident on a generic node.

---

\(^3\) The level of resolution does not affect the obtained bound on the flooding time, provided the resolution is high enough.

### 1.1.3 Geometric mobility models

In the case of node-MEGs defined over geometric spaces (such as in the random waypoint model), we state an easy-to-check condition implying Condition (ii): we transform the pairwise-independence condition on incident edges into mild uniformity conditions on the single-node positional stationary distribution (see Corollary 1). We then show that such uniformity conditions are satisfied by a wide class of random mobility models, including the random waypoint model over a square: we thus get the first known upper bound for its flooding time. Furthermore, when the stationary graph is sparse, the obtained bound is almost tight, i.e., \( O((L/v)\text{polylog}(n)) \) where \( L \) is the diameter of the square and \( v \) is the node speed. In particular, our bound is almost tight whenever the transmission radius and the node speed are absolute constants: this setting is a good fit for opportunistic delay-tolerant Mobile Ad-hoc Networks \([20, 21, 29]\).

### 1.1.4 Graph mobility models

Our upper bound holds even when the mobility space is an arbitrary graph (the vertices of such graph are called points): in this case nodes choose randomly their trajectories from some fixed families of simple paths of the mobility graph and the mixing time is proportional to its hop-diameter. This mobility model is called random path (on graphs). In this model, the parameter \( \beta \) in our bound (1) is determined by the point congestion yielded by the feasible paths of the mobility graph: informally speaking, \( \beta \) is small whenever the random paths do not yield a high point congestion. This condition reduces to a uniformity condition on the (single-node) positional stationary distribution. The bound on the flooding time that we obtain in this way from (1) is almost tight.

The random path model is a generalization of the random walk model on graphs; if we apply our results for the random path model to the random walk model, we obtain bounds on the flooding time in the latter model that improve over previous results \([16]\) for the class of graphs where the mixing time of a random walk is smaller than the meeting time of two random walks.

### 1.1.5 Link-based dynamic networks

We show that our method can also be applied to a subclass of MEGs in which each edge evolves independently according to an arbitrary (hidden) Markov chain. Previously \([5, 11]\), such link-based dynamic models had been studied only in the case in which the edge Markov chain is very simple.
1.2 Previous works

As mentioned before, information spreading in dynamic networks has been extensively studied in the literature under a variety of scenarios and objectives (for a recent good survey see [24]). For brevity’s sake, we restrict our attention to those results which are directly related to our work. Previous models and results can be roughly classified in two main classes: link-based dynamic graphs and mobility models.

Regarding the first class, radio broadcasting is analyzed in [10] on a dynamic graph managed by a worst-case dynamic adversary and on a sequence of independent Erdős-Rény graphs. In [11], an upper bound on the flooding time for the restricted model edge-MEG has been derived. The flooding time of another simple version of edge-MEGs has been studied in [5]. The general MEG model has been introduced in [2] where some results are obtained for the cover time and the hitting time of random walks. Flooding time in stationary MEGs is studied in [12]; the method in [12] only holds for stationary graphs which are connected and good expanders while, as discussed above, our method also works when the stationary graph is highly disconnected (and, thus, it is not a good expander).

A worst-case model of dynamic graphs has been introduced in [23]. The analysis of certain communication tasks is presented under a strong stability condition called $T$-interval connectivity (for $T \geq 1$) which stipulates that for every $T$ consecutive steps a stable connected spanning subgraph must exist.

Regarding mobility models, almost tight bounds on the flooding time for the random walk model have been obtained in [12, 13, 22, 25, 28, 29]. As discussed above, their techniques strongly rely on specific properties of the adopted version of the random walk. The case of general mobility graphs has been considered in [16]: the obtained results are discussed and compared to our results in Sect. 4.2. An upper bound on a variant of the random waypoint model has been derived in [14]. In this version, nodes follow Manhattan paths. Similarly to the works on the random walk models, the ad-hoc analysis in [14] strongly relies on the particular node trajectories and on some geometric properties of the specific positional distribution yielded by this variant. For instance, the analysis of the information spreading uses the fact that nodes only follow horizontal and vertical path thus increasing the probability to meet each other: this property does not hold for the classic version of the random way point. On the contrary our approach exploits more general properties and it obtains bounds for any version of the random waypoint model.

1.3 Paper organization

Section 2 formalizes the general model of dynamic graphs and the flooding process. Section 3 provides the main theorem bounding the flooding time in the most general setting. The node-MEG model is described in Sect. 4, where we give an upper bound on the flooding time for this model. The representations of the random trip and the random path models as specific instances of the node-MEG model are given in Sect. 4.2. There, the flooding-time bound for node-MEGs is transformed into two useful bounds on the flooding time: the first one for the random trip and the second one for the random path. The application of the main theorem to general edge-MEGs is described in Sect. 5 while conclusions with open questions are discussed in Sect. 6.

2 Preliminaries

For any positive $n$, the set of nodes $V$ will be represented by $[n] = \{1, 2, \ldots, n\}$. A dynamic graph $G([n], \{E_t\}_{t \geq 0})$, with node set $[n]$, is a stochastic process represented by a sequence of random variables $E_0, E_1, \ldots, E_t, \ldots$ such that, for every $t$, $E_t$ is the set of edges of the dynamic graph at time $t$. The speed of information spreading can be studied in terms of the flooding time. Given a dynamic graph $G([n], \{E_t\}_{t \geq 0})$ and a node $s \in [n]$, the flooding process with source $s$ is defined as follows. At time $t = 0$, $s$ is the only informed node; then a node $v$ gets informed at time $t + 1$ iff an edge $(u, v) \in E_t$ exists connecting a node $u$ that was informed at time $t$ to the node $v$. Flooding over a dynamic graph is represented by the stochastic process $\{I_t\}_{t \geq 0}$ where

$$ I_0 = \{s\} \quad \text{and} \quad \forall t \geq 1 $$

$$ I_t = I_{t-1} \cup \{j \in [n] | \exists i \in I_{t-1} : (i, j) \in E_t\} $$

The random variable $I_t$ is the set of informed nodes at time $t$. Clearly, it holds that

$$ I_0 \subseteq I_1 \subseteq I_2 \subseteq \cdots \subseteq I_t \subseteq \cdots $$

The flooding time with source $s$ is the random variable $F(G, s) = \min\{t \geq 0 | I_t = [n]\}$ and the flooding time is the random variable $F(G) = \max_s F(G, s)$.

Given a dynamic graph $G([n], \{E_t\}_{t \geq 0})$, we define the following random variables. For every time $t$, for every pair of nodes $i, j \in [n]$ and for every subset of nodes $A \subseteq [n]$, let

$$ e_{i, j}^t = \begin{cases} 1 & \text{if } (i, j) \in E_t \\ 0 & \text{otherwise} \end{cases} $$

$$ e_{A, t}^t = \begin{cases} 1 & \text{if } \exists j \in A : (i, j) \in E_t \\ 0 & \text{otherwise} \end{cases} $$

Moreover, for any binary random variable $X$, the notation $P(X | E_t \leq \alpha) \leq (\geq) \alpha$ stands for
\( \forall \) sequence of edge subsets \( L_0, \ldots, L_T \), with

\[
P \left( \bigwedge_{t=0}^T (E_t = L_t) \right) > 0, \quad \text{it holds}
\]

\[
P \left( X = 1 \left| \bigwedge_{t=0}^T (E_t = L_t) \right) \right) \leq (\text{or } >) \alpha
\]

In what follows, we shortly review some basic concepts of finite discrete-time Markov chains that will be used in our paper. A finite (discrete-time) Markov chain (see [1]) is a random process \( X_0, X_1, \ldots, X_T \), defined by a pair \( M = (S, P) \) where \( S \) is the set of states and \( P : S \times S \to \mathbb{R}^+ \) is the matrix of state-transition probabilities. In particular, for any time step \( t \geq 0 \), \( P(s_i, s_j) \) is the probability that the process moves from state \( X_t = s_i \) to state \( X_{t+1} = s_j \), independently of the previous history of the process.

A state \( s_i \) is periodic if an integer \( \Delta > 1 \) exists such that \( P(X_{t+\ell} = s_i | X_t = s_i) = 0 \) unless \( \ell \) is divisible by \( \Delta \). A state is aperiodic if it is not periodic. A finite Markov chain is aperiodic if all its states are aperiodic. A finite Markov chain is irreducible if, for every possible pair of states \( (s_i, s_j) \), there is positive probability that the process starting from \( s_i \) reaches \( s_j \). A stationary distribution for \( M \) is a probability distribution \( \pi : S \to \mathbb{R}^+ \) such that \( \pi = \pi \cdot P \).

The fundamental theorem (see [1]) of finite Markov chains states that any finite aperiodic irreducible Markov chains always converges to a unique stationary distribution. Finite converging Markov chains are also called ergodic. The worst-case time (with respect to all possible starting distributions) required to reach a distribution which is “close” to the stationary distribution is known as mixing time. More formally, consider two arbitrary probability distributions \( \psi \) and \( \zeta \) over the same discrete domain \( \Omega \). The Total Variation Distance [1] is defined as

\[
||\psi - \zeta||_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\psi(x) - \zeta(x)|
\]

Let \( \pi_t \) be the distribution of \( M \) at time \( t \) starting from state \( q \) and define:

\[
\Delta_q(t) = ||\pi_t - \pi||_{TV}, \quad \Delta(t) = \max_{q \in S} \Delta_q(t),
\]

\[
\tau_q(\epsilon) = \min\{t : \Delta_q(t) \leq \epsilon\}, \quad \tau(\epsilon) = \max_{q \in S} \tau_q(\epsilon)
\]

Then, in our setting, the (asymptotic) mixing time of \( M \) is given by \( \tau(\epsilon) \) for \( \epsilon = O(\frac{1}{n}) \), where \( n \) is the number of nodes of the dynamic graph.

The following lemma will be used in Sect. 4 and can be obtained by a simple application of the definition of Total Variation Distance.

**Lemma 1** For every \( i = 1, \ldots, k \), let \( \psi_i \) and \( \zeta_i \) be any two probability distributions over any domain \( \Omega_i \). Let \( \psi \) and \( \zeta \) denote the product probability distributions over \( \prod_{i=1}^k \Omega_i \) of \( \psi_i, \psi \) and \( \zeta_i, \zeta \), respectively. Then, it holds

\[
||\psi - \zeta||_{TV} \leq \sum_{i=1}^k ||\psi_i - \zeta_i||_{TV}
\]

**Proof** Let \( \Omega = \prod_{i=1}^k \Omega_i \). We denote \( (x_1, \ldots, x_k) \) by \( x \).

For every \( i \), let \( \Omega_i^{-1} = \prod_{j \neq i} \Omega_j \) and let \( \tau_i \) denote \( (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k) \). It holds

\[
||\psi - \zeta||_{TV}
\]

\[
= \frac{1}{2} \sum_{i=1}^k |\psi(x_i) - \zeta(x_i)|
\]

\[
= \frac{1}{2} \sum_{i=1}^k \prod_{j=1}^k |\psi_j(x_j) - \zeta_j(x_j)|
\]

\[
= \frac{1}{2} \sum_{i=1}^k \prod_{j=1}^k |\psi_j(x_j) + \sum_{i=2}^k \zeta_j(x_j) \prod_{j=i+1}^k \psi_j(x_j) - \sum_{i=2}^k \zeta_j(x_j) \prod_{j=i}^k \psi_j(x_j) - \sum_{i=2}^k \zeta_j(x_j) \prod_{j=i}^k \psi_j(x_j)|
\]

\[
= \frac{1}{2} \sum_{i=1}^k \prod_{j=1}^k \zeta_j(x_j) \prod_{j=i+1}^k |\psi_j(x_j) - \zeta_j(x_j)|
\]

\[
\times \prod_{j=1}^k \zeta_j(x_j) \prod_{j=i+1}^k |\psi_j(x_j)|
\]

Observe that

\[
\sum_{i=1}^k \prod_{j=1}^k \zeta_j(x_j) \prod_{j=i+1}^k |\psi_j(x_j)|
\]
\[ \sum_{(x_1, \ldots, x_k) \in \prod_{j=1}^{\frac{n}{2}} \Omega_j} \prod_{j=1}^{k} \psi_j(x_j) \]

\[ \sum_{(x_1, \ldots, x_k) \in \prod_{j=1}^{\frac{n}{2}} \Omega_j} \prod_{j=1}^{k} \xi_j(x_j) = 1 \]

Finally, we provide some well-known inequalities that will be used in the rest of the paper.

**Lemma 2** (The Paley–Zigmund Inequality) If \( X \geq 0 \) is a random variable with finite variance, and if \( 0 < \theta < 1 \), then

\[ P (X \geq \theta E[X]) \geq (1 - \theta^2) \frac{E[X]^2}{E[X^2]} \]  

**Lemma 3** (Lemma 3.1 in [3]) Let \( X_1, \ldots, X_n \) be a sequence of random variables with values in an arbitrary domain, and let \( Y_1, \ldots, Y_n \) be a sequence of binary random variables, with the property that \( Y_1 = Y_1(X_1, \ldots, X_n) \). If

\[ P (Y_i = 1 \mid X_1, \ldots, X_{i-1}) \geq p \]

then

\[ P \left( \sum_{i=1}^{n} Y_i \leq k \right) \leq P \left( B(n, p) \leq k \right) \]

where \( B(n, p) \) is a binomially distributed random variable with parameters \( n \) and \( p \).

**Lemma 4** (The Chernoff Bound) Let \( X_1, \ldots, X_n \) be independent binary random variables and let \( X = \sum_{i=1}^{n} X_i \) and \( \mu = E[X] \). Then, for any \( \delta > 0 \),

\[ P (X < (1 - \delta)\mu) < \exp \left( -\frac{\delta^2 \mu}{2} \right) \]

### 3 Flooding in dynamic graphs

Our goal is to evaluate the flooding time of a dynamic graph as a function of certain properties of its edges. These properties are not required to be satisfied at every point in time; instead, it will suffice that they hold at every *epoch* step, i.e. at the beginning of every time interval of length \( M \), independently of what happened during the interval (see the next definition). When the dynamic graph is a Markov chain admitting a stationary graph, we consider the expansion properties of the stationary graph and the parameter \( M \) equals the mixing time of the dynamic graph. However, aiming for maximal generality, we first introduce such concepts for general (non-Markovian) processes.

**Definition 1** (Density and independence conditions) Let \( M \) be a positive integer and let \( \alpha \) and \( \beta \) be two positive reals such that \( \alpha \leq 1 \). A dynamic graph \( \mathcal{G}([n], \{E_t\}_{t \geq 0}) \) is \((M, \alpha, \beta)\)-stationary if \( \forall t \geq 1, \forall i, j \in [n] \) with \( i \neq j \), \( \forall A \subseteq [n] - \{i, j\} \), the following two conditions hold:

1. \( P \left( E_{i,j}^M \mid E_i \leq (t-1)M \right) \geq \alpha \) (Density Condition);
2. \( P \left( E_{i,A}^M \cdot E_{j,A}^M \mid E_i \leq (t-1)M \right) \leq \beta P \left( E_{i,A}^M E_i \leq (t-1)M \right) \times P \left( E_{j,A}^M E_i \leq (t-1)M \right) \) (\( \beta \)-Independence Condition).

The time steps \( tM \) (\( t = 1, 2, \ldots \)) will be called *epoch* steps.

Let us observe that the density condition determines a lower bound on the (worst-case) edge-probability in \( \mathcal{G}_n \) and so on the expected edge-density of the stationary graph, while the \( \beta \)-independence property refers to the degree of independence among edges in \( \mathcal{G}_n \).

The motivation for such definitions is as follows: we would like to show that, if at time \( t \) we have \( k \) informed nodes, then between time \( t + 1 \) and time \( t + M \) we have, say, a constant probability that \( \Omega(k) \) additional nodes become informed (then we would be able to say that the flooding time is \( O(M \log n) \)). There are \( k \cdot (k - n) \) possible edges connecting nodes that are informed at time \( t \) to not-yet-informed nodes: the density condition tells us that, provided that \( t \) is larger than the mixing time, many of these edges will appear in the dynamic graph, and the independence condition allows us to say that they will be somewhat evenly spread among the non-informed nodes.

The following theorem, which formalizes the above intuition, is our main result.

**Theorem 1** (flooding time) If \( \mathcal{G}([n], \{E_t\}_{t \geq 0}) \) is \((M, \alpha, \beta)\)-stationary for some positive reals \( M, \alpha \leq 1 \) and \( \beta \), then with high probability the flooding time in \( \mathcal{G} \) is

\[ O \left( M \left( \frac{1}{\alpha M} + \beta \right)^2 \log n \right) \].

#### 3.1 Proof of Theorem 1

The epoch steps \( tM \) (\( t = 1, 2, \ldots \)) will be abbreviated by \( t \) (i.e. \( t \) will stand for \( tM \), with respect to a fixed \((M, \alpha, \beta)\)-stationary dynamic graph \( \mathcal{G}([n], \{E_t\}_{t \geq 0}) \)). Thus, we write \( E_t, e_t^{i,j} \) and \( e_t^{i,A} \) to denote \( E_t M, e_t^{i,j} \) and \( e_t^{i,A} \), respectively. These abbreviations will be also used later in Sect. 3.1.2 for other random variables such as \( I_{t,M} \), the set of informed nodes at epoch step \( t \).
3.1.1 Expansion properties

In what follows, we prove certain expansion-like properties of \((M, \alpha, \beta)\)-stationary dynamic graphs. By this, we do not mean that a sample \(G \sim (M, \alpha, \beta)\)-stationary is likely to be an expander graph, in fact, under our conditions, it is possible for \(G\) to be disconnected with high probability. Such properties will then be exploited in the analysis of the flooding process.

Let \(\deg_{i,A}^r\) be the random variable counting the number of nodes in \(A\) connected to \(i\) at epoch step \(r\), i.e., \(\deg_{i,A}^r = |\{j \in A \mid \{i, j\} \in E_r\}|\). Observe that, thanks to Condition (1), the expected value of \(\deg_{i,A}^r\) is at least \(|A|\alpha\); the following lemma provides a concentration result as function of the “independence” parameter \(\beta\).

**Lemma 5** If \(\mathcal{G}([n], \{E_t\}_{t \geq 0})\) is \((M, \alpha, \beta)\)-stationary then, \(\forall r \geq 1, \forall i \in \{n\}, \) and \(\forall A \in \{n\} - \{i\}\),

\[
P \left( \deg_{i,A}^r \geq \frac{|A|\alpha}{2} \mid E_{t-1} \right) \geq \frac{|A|\alpha}{2 + 2|A|\alpha\beta}
\]

**Proof** Firstly we bound the expected square of the degree. For the sake of brevity, we omit the conditioning on \(E_{t-1}\).

It holds that

\[
\mathbb{E} \left[ (\deg_{i,A}^r)^2 \right] = \mathbb{E} \left[ \left( \sum_{j \in A} e_{i,j}^r \right)^2 \right] = \sum_{j,k \in A} \mathbb{E} \left[ e_{i,j}^r \cdot e_{i,k}^r \right] \quad \text{(by linearity of expectation)}
\]

\[
\leq \sum_{j,k \in A, j \neq k} \mathbb{E} \left[ \left( e_{i,j}^r \right)^2 \right] + \mathbb{E} \left[ \deg_{i,A}^r \right] \quad \text{(since } (e_{i,j}^r)^2 = e_{i,j}^r)\]

\[
\leq \beta \mathbb{E} \left[ e_{i,j}^r \cdot e_{i,k}^r \right] \quad \text{(by Def. 1.2)}
\]

\[
\leq \beta \mathbb{E} \left[ \deg_{i,A}^r \right] + \mathbb{E} \left[ \deg_{i,A}^r \right]^2 \quad \text{(by Def. 1.2)}
\]

Moreover, from Condition (1) of Definition 1, it follows that

\[
\mathbb{E} \left[ \deg_{i,A}^r \right] \geq \frac{|A|\alpha}{2}
\]

Thus, from the Paley–Zigmund inequality (i.e. Eq. (3) with \(\theta = 1/2\)) and the above bounds, we obtain

\[
P \left( \deg_{i,A}^r \geq \frac{|A|\alpha}{2} \right) \geq \mathbb{E} \left[ \left( \frac{\deg_{i,A}^r}{|A|\alpha} \right)^2 \right] \geq \frac{1}{2} \mathbb{E} \left[ \deg_{i,A}^r \right]^2
\]

\[
\geq 1 + \frac{2}{\beta} \mathbb{E} \left[ \deg_{i,A}^r \right] \quad \text{and } \mathbb{E} \left[ \deg_{i,A}^r \right] \geq \frac{|A|\alpha}{2 + 2|A|\alpha\beta}
\]

Since the function \(\frac{1}{2 + 2|A|\alpha\beta}\) is increasing, we have that

\[
P \left( \deg_{i,A}^r \geq \frac{|A|\alpha}{2} \right) \geq \mathbb{E} \left[ \deg_{i,A}^r \right] \geq \frac{2 + 2|A|\alpha}{2 + 2|A|\alpha\beta}
\]

\(\square\)

The next lemma extends Lemma 5 to the expansion of an arbitrary node subset. For every epoch step \(r\) and for every \(A, B \subseteq \{n\}\), we consider the random variable

\[
\deg_{A,B}^r = |\{j \in B \mid \exists i \in A : \{i, j\} \in E_r\}|
\]

**Lemma 6** If \(\mathcal{G}([n], \{E_t\}_{t \geq 0})\) is \((M, \alpha, \beta)\)-stationary then, \(\forall r \geq 1, \forall A \subseteq \{n\}, \) and \(\forall B \subseteq \{n\} - A\), it holds that

\[
P \left( \deg_{A,B}^r \geq \frac{|A||B|\alpha}{4 + 4|A|\alpha\beta} \mid E_{t-1} \right) \geq \frac{|A||B|\alpha}{4 + 6|A||B|\alpha\beta}
\]

**Proof** The proof argument is very similar to that of the proof of Lemma 5. We first bound the expected square of the expansion. For the sake of brevity, we omit the conditioning by \(E_{t-1}\).

\[
\mathbb{E} \left[ (\deg_{A,B}^r)^2 \right] = \mathbb{E} \left[ \left( \sum_{j \in B} e_{j,A}^r \right)^2 \right] = \sum_{j,k \in B} \mathbb{E} \left[ e_{j,A}^r \cdot e_{k,A}^r \right] \quad \text{(by Def. 1.2)}
\]

\[
\leq \beta \sum_{j,k \in B} \mathbb{E} \left[ e_{j,A}^r \cdot e_{k,A}^r \right] \quad \text{(by Condition (2))}
\]

Moreover, from Lemma 5 we get

\[
\mathbb{E} \left[ \deg_{A,B}^r \right] = \sum_{j \in B} \mathbb{E} \left[ e_{j,A}^r \right] = \sum_{j \in B} \mathbb{P} \left( e_{j,A}^r > 0 \right) \geq \sum_{j \in B} \frac{|A||B|\alpha}{2 + 2|A||B|\alpha\beta}
\]
From the Paley–Zigmund inequality (i.e. Eq. (3) with \( \theta = 1/2 \)) and the above bounds, we have that
\[
P \left( \deg_{A,B}^r \geq \frac{\left| A \right| \left| B \right| |R|}{4 + 4 \left| A \right| \alpha \beta} \right) \\
\geq 2 \mathbb{E} \left( \deg_{A,B}^r \right)^2 \\
\geq 2 \mathbb{E} \left( \deg_{A,B}^r \right)^2 \\
= 4 + 6 \left| A \right| \left| B \right| |R| \\
\]
\( \square \)

When the dynamic graph is sparse, the expansion rate obtained by considering a single snapshot of the process (i.e. the expansion of a node subset at time \( \tau \)) does not suffice to get a good number of new informed nodes. In this case, a “dynamic” version of the expansion properties is required. For every epoch step \( \tau \), for every \( T \geq 1 \), and for every \( A \subseteq \{i\} \), define
\[
\text{spread}_{A}^{r,T} = \left\{ j \in \{n\} - A \mid \exists r' \exists i \in A : \tau < r' \leq \tau + T \land \{i, j\} \in E_{r'} \right\}
\]
That is, \( \text{spread}_{A}^{r,T} \) is the number of nodes outside \( A \) that get connected to nodes in \( A \) during the epoch steps in the interval \( (\tau, \tau + T) \).

**Lemma 7** If \( G(n, \{ E_i \}_{i \geq 0}) \) is \((M, \alpha, \beta)\)-stationary then, \( \forall \tau \geq 1, \forall A \subseteq \{n\} \) with \( |A| \leq n/4 \), and \( \forall \tau \geq 0 \),
\[
P \left( \text{spread}_{A}^{r,T} < |A| \mid E_{\leq \tau} \right) < \exp(-t),
\]
where
\[
T = 256 \left( \frac{1}{|A|^2 n^2 |\alpha|^2} + \frac{\beta}{n \alpha} + \frac{|A| |B|^2}{n} \right) + \left( \frac{4 |A| n \alpha}{|A|^2 n^2 |\alpha|^2} + 3 \beta \right) t
\]

**Proof** For brevity’s sake, we omit the conditioning by \( E_{\leq \tau} \). Let \( S_{\tau} \) be the set of nodes outside \( A \) that get connected to nodes in \( A \) in the epoch step \( \bar{\tau} \in (\tau, \tau + T) \). Formally,
\[
S_0 = \emptyset \quad \text{and} \quad S_{\bar{\tau}} = S_{\bar{\tau}+1} \cup \{ j \in \{n\} - A \mid \exists i \in A : \{i, j\} \in E_{\bar{\tau}+1} \}
\]
Clearly, \( |S_{\bar{\tau}}| = \text{spread}_{A}^{r,T} \). Let
\[
\gamma = \frac{|A| n \alpha}{8 + 8 |A| \alpha \beta}
\]
Define
\[
Y_{\ell} = \begin{cases} 1 & \text{if } |S_{\ell-1}| \geq |A| \text{ or } |S_{\ell}| \geq |S_{\ell-1}| + \gamma \\ 0 & \text{otherwise} \end{cases}
\]
Observe that \( Y_{\ell} = f_{\ell}(E_{\ell+1}, \ldots, E_{\ell+T}) \) for a suitable function \( f_{\ell} \). From the inequality
\[
P \left( A \lor B \mid H \right) \geq P \left( B \mid H \land \bar{A} \right)
\]
we get that
\[
P \left( Y_{\ell} = 1 \mid E_{\ell+1} \right) = P \left( |S_{\ell-1}| \geq |A| \lor |S_{\ell}| \geq |S_{\ell-1}| + \gamma \mid E_{\ell+1} \right) \\
\geq P \left( |S_{\ell}| \geq |S_{\ell-1}| + \gamma \mid E_{\ell+1} \land |S_{\ell-1}| < |A| \right)
\]
where \( E_{\ell+1} \) stands for \( E_{\ell+1}, \ldots, E_{\ell+T} \). Assume that \( |S_{\ell-1}| < |A| \) and let \( W = [n] - A - S_{\ell-1} \). Since \( |A| \leq n/4 \), it holds that \( |W| \geq \frac{n}{2} \) and
\[
|S_{\ell}| \geq |S_{\ell-1}| + \gamma \quad \Leftrightarrow \quad \text{deg}_{A,W}^{r+T} \geq \gamma
\]
\[
\Leftrightarrow \text{deg}_{A,W}^{r+T} \geq \gamma
\]
From Lemma 6 we have that
\[
P \left( \deg_{A,W}^{r+T} \geq \gamma \mid E_{\ell+1} \land |S_{\ell-1}| < |A| \right) \\
\geq P \left( \deg_{A,W}^{r+T} \geq \frac{|A||W| |\alpha|^2}{4 + 4 |A| \alpha \beta} \mid E_{\ell+1} \land |S_{\ell-1}| < |A| \right) \\
\geq \frac{|A||W| |\alpha|^2}{4 + 6 |A||W| |\alpha|^2} \geq \frac{|A| n \alpha}{8 + 6 |A| n \alpha \beta} =: p
\]
(6)

From Eqs. (4), (5), and (6) we get
\[
P \left( Y_{\ell} = 1 \mid E_{\ell+1} \right) \geq p
\]
We can now apply Lemma 3 to the random variables \( E_{\ell+1}, \ldots, E_{\ell+T} \) and to the random variables \( Y_1, \ldots, Y_T \)
\[
P \left( \sum_{\ell=1}^{T} Y_{\ell} < \frac{|A|}{\gamma} \right) \leq P \left( B(T, p) < \frac{|A|}{\gamma} \right)
\]
Since \( \text{spread}_{A}^{r,T} < |A| \) \( \Rightarrow \) \( \sum_{\ell=1}^{T} Y_{\ell} < \frac{|A|}{\gamma} \), from the above inequality we obtain
\[
P \left( \text{spread}_{A}^{r,T} < |A| \right) \leq P \left( B(T, p) < \frac{|A|}{\gamma} \right)
\]
By applying Chernoff’s Bound (Lemma 4), after simple calculations, we get
\[
P \left( B(T, p) < \frac{|A|}{\gamma} \right) \leq \exp(-t)
\]
for
\[
T = 256 \left( \frac{1}{|A|^2 n^2 |\alpha|^2} + \frac{\beta}{n \alpha} + \frac{|A| |B|^2}{n} \right) + \left( \frac{4 |A| n \alpha}{|A|^2 n^2 |\alpha|^2} + 3 \beta \right) t
\]
\( \square \)
The next result still concerns the “dynamic” expansion of any subset of nodes. It will be applied when the subset of informed nodes is large (say at least \(n/2\)) in order to get a good bound on the completion time of the last phase of the flooding process. Let us define the following random variables

\[
e_{i,A}^{\tau,T} = \begin{cases} 1 & \text{if } \exists \tau' \in A : \tau < \tau' \leq \tau + T \text{ and } \{i, j\} \in E_{\tau'} \leq \tau \leq \tau + T \text{ and } \{i, j\} \in E, \\ 0 & \text{otherwise} \end{cases}
\]

\[
e_{i,A}^{\tau,T} = \begin{cases} 1 & \text{if } \exists \tau' \in A : \tau < \tau' \leq \tau + T \text{ and } \{i, j\} \in E_{\tau'} \leq \tau \leq \tau + T \text{ and } \{i, j\} \in E, \\ 0 & \text{otherwise} \end{cases}
\]

\[
(7)
\]

**Lemma 8** If \(G([n], (E_t)_{t \geq 0})\) is \((M, \alpha, \beta)\)-stationary then, for every \(\tau, t \geq 1\), for every \(A \subseteq [n]\), and for every \(i \in [n] \setminus A\), it holds that

\[
P(e_{i,A}^{\tau,T} = 0 \mid E_{\leq \tau}) \leq \exp(-t), \text{ where } T = 2 \left( \frac{1}{\exp(1)\alpha + \beta} \right) \]

Proof For brevity’s sake, we omit the conditioning by \(E_{\leq \tau}\) in the notation \(P(\cdot)\). For \(s = 1, \ldots, T\), define the random variables

\[
Y_s = e_{i,A}^{\tau+s}
\]

Observe that \(Y_s = f_s(E_{\leq \tau+s})\) for a suitable function \(f_s\). Lemma 5 implies

\[
P(Y_s = E_{\leq \tau+s} \mid 1) = P(\deg_{E_{\leq \tau+s}} > 0 \mid E_{\leq \tau+s+1}) \geq p = \frac{1}{2 + 1/\alpha + \beta}
\]

By applying Lemma 3 to \(E_{\leq \tau+1}, \ldots, E_{\leq \tau+7}\) and \(Y_1, \ldots, Y_T\), we get

\[
P\left(\sum_{s=1}^{T} Y_s = 0\right) \leq P(B(T, p) = 0) = (1 - p)^T \leq \exp(pT) = \exp(-t)
\]

\[
3.1.2 \text{ A bound on the flooding time}
\]

We first provide a bound for the number of epoch steps required to obtain at least \(n/2\) informed nodes. This is the spreading phase.

**Lemma 9** (Spreading Phase) If \(G([n], (E_t)_{t \geq 0})\) is \((M, \alpha, \beta)\)-stationary then, \(\forall \tau \geq \hat{T}\) with \(\hat{T} = O\left(\left(\frac{1}{n^2} + \beta\right)^2 \log n\right)\) it holds that\(^4\)

\[
P\left(|I_\tau| < \frac{n}{2}\right) \leq \frac{1}{n^2}
\]

\[^4\] Recall that we are still using the abbreviations like \(I_\tau = I_{\tau,M}\).

\[
\text{Proof} \ \text{Observe that, for any } |A| \leq n/4, \text{ the bound on } T, \text{ for } t = 1 \text{ in Lemma 7, satisfies}
\]

\[
T = \Theta\left(\left(\frac{1}{|A|n^2\alpha^2} + \beta + \frac{|A|\beta^2}{n}\right) + \left(\frac{1}{|A|n\alpha} + \beta\right)\right) = O\left(\left(\frac{1}{n^2\alpha} + \beta\right)^2\right)
\]

From Lemma 7, there is positive constant probability (let’s say greater than \(c > 0\)) that, after every interval of \(T\) epoch steps, the size of the set of informed nodes at least doubles as long as it is smaller than \(n/2\). We say that an interval of \(T\) epoch steps is a success if, at the end of the interval, the number of informed nodes has doubled. So, the probability of a success is greater than \(c\) independently of the previous “history” of the process. By applying Lemma 3 and the Chernoff’s bound, we have that, after a suitable number \(k = \Theta(\log n)\) of time intervals of \(T\) epoch steps, w.h.p. (let’s say with probability larger than \(1 - 1/n^2\)), there will be a logarithmic number of successes: this suffices to get at least \(n/2\) informed nodes.

\[
\text{Lemma 10 \ (saturation phase) Let } G([n], (E_t)_{t \geq 0}) \text{ be } (M, \alpha, \beta)-\text{stationary and assume that the flooding process is in some epoch step } \tau \text{ such that } |I_\tau| \geq n/2. \text{ Then, w.h.p., all nodes get informed within } O\left(\left(\frac{1}{n^2\alpha} + \beta\right)^2 \log n\right) \text{ epoch steps.}
\]

\[
\text{Proof} \ \text{Let us consider any epoch step } \tau \text{ where } |I_\tau| \geq n/2. \text{ Then, by choosing } A = I_\tau \text{ and } t = \Theta(\log n), \text{ Lemma 8 implies that w.h.p. every node gets informed within } O\left(\left(\frac{1}{n^2\alpha} + \beta\right)^2 \log n\right) \text{ epoch steps. Then, by the Union Bound, all nodes get informed after the same number of epochs.}
\]

\[
\text{Proof of Theorem 1} \ \text{Thanks to Lemma 9, w.h.p., after } O\left(M\left(\frac{1}{n^2\alpha} + \beta\right)^2 \log n\right) \text{ steps the set of informed nodes will be at least } n/2. \text{ Then, Lemma 10 implies that, after further } O\left(M\left(\frac{1}{n^2\alpha} + \beta\right) \log n\right) \text{ steps, all nodes will get informed, w.h.p.}
\]

\[
4 \text{ Node Markovian evolving graphs}
\]

We introduce the general class of Node Markovian Evolving Graphs (in short, node-MEGs) where the behavior of the nodes is ruled by independent Markov chains. To every node is associated an independent copy of the same ergodic Markov chain whose states contain enough information to determine whether two nodes are connected or not. A popular instance of node-MEGs is given by the random walk model described in the Introduction: in this simple model, the state of the Markov chain (only) encodes the geometric position of the node since the presence of a link between two nodes only depends on their relative positions. A more complex example of node-MEG is the random waypoint model.
In this case, even though the links are still determined by the node positions, the state of a node must encode the current position and the source-destination path the node is running over. The latter information is required in order to get the next state of the node, i.e., to correctly define the state-transition matrix of the Markov chain.

Formally, in order to define a node-MEG, we need to specify the following items:

- **The Markov chain of the nodes.** Let $\mathcal{M} = (S, P) = (S, P)$ be an ergodic, discrete-time Markov chain where $S$ is the set of states and $P : S \times S \rightarrow \mathbb{R}$ are the transition probabilities. Each node $i$ has an initial state $s_i^0$ at time zero, and the state of node $i$ evolves according to $P$. The Markov chains describing the states of different nodes evolve independently. Notice that the Markov chain $\mathcal{M}$ (both the set $S$ and the transition probabilities $P$) may depend on the number of nodes.

- **Dynamic edges.** The dynamic topology is determined by a fixed symmetric map $C : S \times S \rightarrow \{0, 1\}$: any two nodes $i, j$ are connected at a given time $t$ if $C(s_i^t, s_j^t) = 1$, where $s_i^t$ and $s_j^t$ are the states of nodes $i$ and $j$ at time $t$, respectively. The symmetric map $C(\cdot, \cdot)$ is also called the (static) connection map of $\mathcal{M}$. Observe that this connection map does not refer at all to the directed graph representing the Markov chain $\mathcal{M}$.

- **Initial global-state distributions.** The initial state $s_i^0$ of each node $i$ is chosen at random according to a probability distribution $\iota_i$ over the set of states $S$. We denote by $t_i$ the global initial probability distribution determined by the product of the probability distributions $\iota_i$. The state of a node $i$ at time $t$ is thus a random variable $s_i^t$ fully determined by the initial distribution $t_i$ and the $i$-th copy of $\mathcal{M}$.

- **The dynamic graph.** A node-MEG $\text{NM}(n, \mathcal{M}, C)$ together with an initial global probability distribution $t$ determines a dynamic graph

  $$G([n], \{E_t\}_{t \geq 0}) \text{ where, for any } t \geq 0, E_t = \{(i, j) \mid C(s_i^t, s_j^t) = 1\}$$

  Observe that, while the $n$ random variables $s_i^t$ are mutually independent, the binary random variables $e_{i, j}^t$ (representing the dynamic edges) are in general mutually dependent.

In Sect. 4.2, we provide a geometric example of node-MEG in which the set $S$, the transition probabilities $P : S \times S \rightarrow \mathbb{R}$, and the connection map $C(\cdot, \cdot)$ are specified.

**Definition 2** A node-MEG $\text{NM}(n, \mathcal{M}, C)$ in its stationary phase is the random graph distribution (in short, random graph) determined by first choosing the state $s_i$ of each node $i$ independently at random according to the stationary distribution $\pi$ of $\mathcal{M}$ and, then, by looking at the values $(s_i, s_j)$’s in order to determine the edge distribution of the random graph.

Clearly, as in any model of random graphs, we can consider the probability that a fixed subset of nodes is connected in the stationary phase of any node-MEG.

**Definition 3** Given any node-MEG $NM = \text{NM}(n, \mathcal{M}, C)$ in its stationary phase, we define:

- $P_{\text{NM}}(i, j)$ as the probability that a fixed pair $(i, j)$ of nodes is connected, and
- $P_{\text{NM}^2}(i, j, k)$ as the probability that two fixed nodes $i, j$ are both connected to another fixed node $k$.

It is then easy to verify that node-MEGs enjoy the following useful property.

**Fact 4** Consider any node-MEG $\text{NM} = \text{NM}(n, \mathcal{M}, C)$ in its stationary phase, then both $P_{\text{NM}}(i, j)$ and $P_{\text{NM}^2}(i, j, k)$ do not depend on the choice of the fixed nodes: they are functions of the stationary distribution $\pi$ of $\mathcal{M}$ and of the symmetric map $C(\cdot, \cdot)$, only. More precisely, let $\Gamma(x) = \{y \in S \mid C(x, y) = 1\}$, then it holds

$$P_{\text{NM}} = \sum_{x \in S} \pi(x)\pi(\Gamma(x)) \text{ and } P_{\text{NM}^2} = \sum_{x \in S} \pi(x)\pi(\Gamma(x))^2$$

In Sect. 4.2, we will show that Node-MEGs subsume a wide class of mobility models.

**Flooding in Node-MEGs** We now derive some simple properties ensuring that a node-MEG $\text{NM}(n, \mathcal{M}, C)$ is a $(\mathcal{M}, \alpha, \beta)$-stationary dynamic graph. Since edges are not independent in a node-MEG, it is non-trivial to verify that the $\beta$-independence condition holds. The models at hand are Markovian so the idea is to consider the model during its stationary state, that is, the time $\mathcal{M}$ between two consecutive epoch steps is proportional to the mixing time of the Markov chain $\mathcal{M}$. Another idea is to use the independence among the node evolutions and Fact 4 to reduce the $\beta$-independence condition (which concerns arbitrary sets $A$ of edges) to a condition about the almost pairwise independence of edges incident on the same node.

**Theorem 2** Let $\text{NM} = \text{NM}(n, \mathcal{M}, C)$ be a node-MEG such that $P_{\text{NM}} \geq \frac{1}{n}O(1)$ and $P_{\text{NM}^2} \leq \eta(P_{\text{NM}})^2$, for some $\eta \geq 1$. Then, with high probability, the flooding time is

$$O\left(\frac{T_{\text{mix}}}{{n}P_{\text{NM}} + \eta}^2 \log^2 n\right)$$

where $T_{\text{mix}}$ is the mixing time of the Markov chain $\mathcal{M}$.

The crux of the proof of the above theorem lies in showing that the condition $P_{\text{NM}^2} \leq \eta(P_{\text{NM}})^2$ implies a $\beta$-independence property which is sufficient to apply Theorem 1.
4.1 Proof of Theorem 2

We need some preliminary notations. For every \( x \in S \), recall that
\[
\Gamma(x) = \{ y \in S : C(x, y) = 1 \}
\]

In words, \( \Gamma(x) \) is the set of states that are one hop away from state \( x \). For any node \( i \in [n] \), let \( v_i \) be a probability distribution over the set of states \( S \). The symbol \( v \) (without subscript) will denote the product probability distribution over \( \prod_{i \in [n]} S \). By assuming that the states of the nodes are chosen randomly according to the probability distribution \( v \), for every \( i, j \in [n] \) and for every \( A \in \{ [n] \setminus \{ i \} \} \), define binary random variables \( e_{i,j}^v \) and \( e_{i,A}^v \),
\[
e_{i,j}^v = 1 \quad \text{if nodes } i \text{ and } j \text{ are connected}
\]
\[
e_{i,A}^v = 1 \quad \text{if node } i \text{ is connected to at least one node in } A
\]
It is easy to verify that the following equations hold:
\[
\mathbb{P}(e_{i,j}^v) = \sum_{x \in S} v_i(x) v_j(\Gamma(x)) \quad \text{and} \quad \mathbb{P}(e_{i,A}^v) = \sum_{x \in S} v_i(x) \prod_{j \in A} (1 - v_j(\Gamma(x))) ,
\]
where \( \overline{X} \) is the negation of \( X \) defined as \( \overline{X} = 1 \) iff \( X = 0 \). Notice that \( v_j(\Gamma(x)) \) is the probability that node \( j \) is connected to a fixed node in state \( x \). Let \( \pi \) be the stationary probability distribution of the Markov chain \( M \). With an abuse of notation, we denote by \( \pi \) also the probability distribution over \( \prod_{i \in [n]} S \) given by the product of the \( \pi \)'s. For the sake of brevity, we write \( e_{i,j} \), \( e_{i,A} \) to mean, respectively \( e_{i,j}^\pi \), \( e_{i,A}^\pi \). The next lemma shows that, under the second condition of Theorem 2 (i.e. \( P_{NM2} \leq \eta (P_{NM})^2 \), for some \( \eta \geq 1 \)), the node-MEG satisfies a \( \beta \)-independence property as in the assumptions of Theorem 1.

**Lemma 11** Let \( NM = NM(n, M, C) \) be a node-MEG such that
\[
P_{NM2} \leq \eta (P_{NM})^2 \quad \text{for some } \eta \geq 1
\]
Then, it holds that
\[
\forall i, j \in [n] \ \forall A \subseteq \{ [n] \} \setminus \{ i, j \} , \ \mathbb{P}(e_{i,A} \cdot e_{j,A}) \leq 17 \eta \mathbb{P}(e_{i,A}) \mathbb{P}(e_{j,A})
\]

**Proof** For the sake of brevity, \( q(x) \) and \( q(x, y) \) denote, respectively, \( \pi(\Gamma(x)) \) and \( \pi(\Gamma(x) \cup \Gamma(y)) \). We also define the following subset of states
\[
V = \left\{ x \in S : q(x) > \frac{1}{\sqrt{|A|}} \right\}
\]
We now prove two claims that represent the key ingredients of the proof.

**Claim** For every \( k \in [n] - A \), it holds that
\[
\mathbb{P}(e_{k,A}) \geq \frac{\sqrt{|A|}}{2} \left( P_{NM} - \sum_{x \in V} \pi(x)q(x) \right) + \frac{1}{2} \pi(V)
\]

**Proof** We provide an upper bound \( \mathbb{P}(e_{k,A}) \). Clearly, it holds that
\[
\mathbb{P}(e_{k,A}) = \prod_{x \in V} \pi(x)(1 - q(x))^{|A|}
\]
\[
+ \sum_{x \in S - V} \pi(x)(1 - q(x))^{|A|}
\]
If \( x \in V \), then
\[
(1 - q(x))^{|A|} \leq e^{-q(x)|A|} \leq e^{-\sqrt{|A|}} \leq e^{-1}
\]
If \( x \in S - V \), then
\[
(1 - q(x))^{|A|} \leq e^{-q(x)|A|} \leq 1 - \frac{q(x)|A|}{2\sqrt{|A|}}
\]
\[
= 1 - \frac{\sqrt{|A|}}{2} q(x),
\]
where we used the inequality \( e^{-x} \leq 1 - x/(2a) \) (that holds for \( a \geq 1 \) and any \( 0 \leq x \leq a \)). By combining Eqs. (8), (9), and (10), we obtain
\[
\mathbb{P}(e_{k,A}) \leq 1 - \frac{\sqrt{|A|}}{2} q(x)
\]
and the claim immediately follows.
Claim For every $i, j \in [n] - A$, it holds that

$$P(e_{i,A} \cdot e_{j,A}) \leq P(e_{i,A}) P(e_{j,A}) + \eta |A| (P_{NM})^2$$

Proof For any two binary random variables $X$ and $Y$, it holds that $X \cdot Y = (X \lor Y)$ and thus $P(X \cdot Y) = P(X \lor Y) + 1 - P(X) - P(Y)$. Hence,

$$P(e_{i,A} \cdot e_{j,A}) = P(e_{i,A} \cdot e_{j,A}) + 1 - P(e_{i,A}) - P(e_{j,A}) \tag{11}$$

First, we bound $P(e_{i,A} \cdot e_{j,A})$. It holds that

$$P(e_{i,A} \cdot e_{j,A}) = \sum_{x \in S} \sum_{y \in S} \pi(x) \pi(y)$$

$$\times \prod_{h \in A} \pi(S - (\Gamma(x) \cup \Gamma(y)))$$

$$= \sum_{x, y \in S} \pi(x) \pi(y) (1 - \pi(\Gamma(x) \cup \Gamma(y)))^{A}$$

We consider the set $R$ of all pairs of states in $S$ having hop distance at most 2 (w.r.t. the Markov chain $M$) and its complement set:

$$R = \{(x, y) \in S \times S \mid \Gamma(x) \cap \Gamma(y) \neq \emptyset\} \quad \text{and} \quad \overline{R} = S \times S - R$$

Observe that if $(x, y) \in \overline{R}$ then $q(x, y) = q(x) + q(y)$. Thus, it holds that

$$P(e_{i,A} \cdot e_{j,A}) = \sum_{(x, y) \in \overline{R}} \pi(x) \pi(y) (1 - q(x, y))^{A} +$$

$$\sum_{(x, y) \in R} \pi(x) \pi(y) (1 - q(x, y))^{A}$$

$$= \sum_{(x, y) \in \overline{R}} \pi(x) \pi(y) (1 - q(x) - q(y))^{A} +$$

$$\sum_{(x, y) \in R} \pi(x) \pi(y) (1 - q(x, y))^{A}$$

$$= \Lambda_1 + \Lambda_2 \tag{12}$$

where

$$\Lambda_1 = \sum_{x, y \in S} \pi(x) \pi(y) (\max\{1 - q(x) - q(y), 0\})^{A}$$

$$\Lambda_2 = \sum_{(x, y) \in \overline{R}} \pi(x) \pi(y) \left[ (1 - q(x, y))^{A} - (\max\{1 - q(x) - q(y), 0\})^{A} \right]$$

In order to bound $\Lambda_1$ we use the inequality $\max\{1 - a - b, 0\} \leq (1 - a)(1 - b)$, that holds for any $a$ and $b$ with $0 \leq a, b \leq 1$:

$$\Lambda_1 \leq \sum_{x, y \in S} \pi(x) \pi(y) (1 - q(x))^{A} (1 - q(y))^{A}$$

$$= \sum_{x \in S} \pi(x) (1 - q(x))^{A} \sum_{y \in S} \pi(y) (1 - q(y))^{A}$$

$$= P(e_{i,A}) P(e_{j,A}) \tag{13}$$

By combining Eqs. (11), (12), and (13), we obtain

$$P(e_{i,A} \cdot e_{j,A}) \leq 1 - P(e_{i,A}) - P(e_{j,A})$$

$$+ P(e_{i,A}) P(e_{j,A}) + \Lambda_2$$

$$= (1 - P(e_{i,A})) (1 - P(e_{j,A})) + \Lambda_2$$

$$= P(e_{i,A}) P(e_{j,A}) + \Lambda_2 \tag{14}$$

In order to upper bound $\Lambda_2$, we use the following inequality. If $0 \leq b \leq 1$ and $0 \leq 1 - a + b \leq 1$, then, for any integer $k \geq 1$,

$$(1 - a + b)^k - (\max\{1 - a, 0\})^k \leq kb \tag{15}$$

The above inequality can be easily proved by distinguishing the two cases $1 - a \leq 0$ and $1 - a > 0$. In the latter, we apply induction on $k$. Let $\hat{q}(x, y)$ denote $\pi(\Gamma(x) \cap \Gamma(y))$. Observe that

$$q(x, y) = q(x) + q(y) - \hat{q}(x, y)$$

From Inequality 15 with $a = q(x) + q(y)$ and $b = \hat{q}(x, y)$, we get

$$(1 - q(x, y))^{A} - (\max\{1 - q(x) - q(y), 0\})^{A} \leq |A| \hat{q}(x, y)$$

It follows that

$$\Lambda_2 \leq |A| \sum_{(x, y) \in R} \pi(x) \pi(y) \hat{q}(x, y)$$

(it is enough to show that $x, y \notin R$)

$$= |A| \sum_{x, y \in S} \pi(x) \pi(y) \hat{q}(x, y)$$

(since $x, y \notin R$)

$$= |A| \sum_{x, y \in S} \pi(z) C(z, x) C(z, y)$$

$$= |A| \sum_{z \in S} \pi(z) C(z, x) \sum_{y \in S} \pi(y) C(z, y)$$

$$= |A| \sum_{z \in S} \pi(z) q(z)^2$$

$$= |A| P_{NM}^2 \quad \text{(from Fact 4)}$$

$$\leq \eta |A| (P_{NM})^2 \quad \text{(from lemma’s hypothesis)} \tag{16}$$

By combining this with Eq. (14), we obtain
\[ \mathbf{P}(e_{i,A} \cdot e_{j,A}) \leq \mathbf{P}(e_{i,A}) \mathbf{P}(e_{j,A}) + \eta |A| (P_{NM})^2 \]

In order to apply the above two claims we distinguish two cases. We first consider the case where

\[
\sum_{x \in V} \pi(x)q(x) < \frac{1}{2} P_{NM}
\] (17)

Then, from Claim 4.1 and Eq. (17), we get

\[ \mathbf{P}(e_{k,A}) \geq \frac{\sqrt{|A|}}{2} \left( P_{NM} - \sum_{x \in V} \pi(x)q(x) \right) \geq \frac{\sqrt{|A|}}{2} \left( P_{NM} - \frac{1}{2} P_{NM} \right) = \frac{\sqrt{|A|}}{4} P_{NM} \]

From the above inequality and Claim 4.1 we get

\[ \mathbf{P}(e_{i,A} \cdot e_{j,A}) \leq \mathbf{P}(e_{i,A}) \mathbf{P}(e_{j,A}) + \eta |A| (P_{NM})^2 \]

Now, let us consider the case where

\[
\sum_{x \in V} \pi(x)q(x) > \frac{1}{2} P_{NM}
\] (18)

It holds that

\[
\frac{1}{\pi(V)} \left( \sum_{x \in V} \pi(x)q(x) \right)^2 \leq \sum_{x \in V} \pi(x)q(x)^2 \quad \text{(by Jensen’s inequality)}
\]

\[ \leq \sum_{x \in S} \pi(x)q(x)^2 = P_{NM2} \]

\[ \leq \eta (P_{NM})^2 \quad \text{(by lemma’s hypothesis)} \]

\[ < 4\eta \left( \sum_{x \in V} \pi(x)q(x) \right)^2 \quad \text{(by Eq. (18))} \]

It follows that

\[ \pi(V) \geq \frac{1}{4\eta} \]

Thus, thanks to Claim 4.1 we obtain

\[ \mathbf{P}(e_{j,A}) \geq \frac{\sqrt{|A|}}{2} \left( P_{NM} - \sum_{x \in V} \pi(x)q(x) \right) + \frac{1}{2} \pi(V) \geq \frac{1}{8\eta} \]

It follows that

\[ \mathbf{P}(e_{i,A} \cdot e_{j,A}) \leq \mathbf{P}(e_{i,A}) \frac{8\eta}{8\eta} \leq 8\eta \mathbf{P}(e_{i,A}) \mathbf{P}(e_{j,A}) \]

Lemma 12 Let \( NM = NM(n, M, C) \) be a node-MEG such that

\[ P_{NM2} \leq \eta (P_{NM})^2 \quad \text{for some } \eta \geq 1 \] (19)

Moreover, let \( v \) be a product probability distribution such that

\[ \forall i \in [n] \quad ||v_i - \pi||_{TV} \leq \frac{(P_{NM})^2}{2n} \] (20)

Then, for any \( i, j \in [n] \) and for any \( A \subseteq [n] - \{i, j\} \), it holds:

(a) \[ \mathbf{P}(e_{i,j}^\nu) \geq \frac{P_{NM}}{2} \]

(b) \[ \mathbf{P}(e_{i,A}^\nu \cdot e_{j,A}^\nu) \leq 72\eta \mathbf{P}(e_{i,A}) \mathbf{P}(e_{j,A}). \]

Proof For any pair of nodes \( i, j \in [n] \), let us consider the set \( C_{i,j} \) of all global state configurations in which such two nodes are connected (according to the map \( C \)):

\[ \forall i, j \in [n] (i \neq j), \quad C_{i,j} = \left\{ (x_1, \ldots, x_n) \in \prod_{k=1}^{n} S \mid C(x_i, x_j) = 1 \right\} \]

It holds that

\[ \mathbf{P}(e_{i,j}^\nu) = v(C_{i,j}) \]

\[ \geq \pi(C_{i,j}) - \eta (P_{NM})^2 \quad \text{(from Eq. (20) and Lemma 1)} \]

\[ = P_{NM} - \frac{(P_{NM})^2}{2} \geq \frac{P_{NM}}{2} \]

This proves Claim (a).

As for Claim (b), we consider the following variant of the sets \( C_{i,j} \). For any \( k \in [n] \) and for any \( A \subseteq [n] - \{k\} \), let

\[ C_{k,A} = \left\{ (x_1, \ldots, x_n) \in \prod_{i=1}^{n} S \mid \exists h \in A : C(x_k, x_h) = 1 \right\} \]

For any \( i, j \in [n] \) and for any \( A \subseteq [n] - \{i, j\} \), let

\[ C_{i,j,A} = \left\{ (x_1, \ldots, x_n) \in \prod_{i=1}^{n} S \mid \exists h, k \in A : C(x_i, x_h) = 1 \right\} \]

\[ = 1 \land C(x_j, x_k) = 1 \]
It holds that
\[
\begin{align*}
\mathbb{P} \left( e_{i,A}' \cdot e_{j,A}' \right) &= \nu(C_{i,j,A}) \\
&\leq \pi(C_{i,j,A}) + \frac{(P_{\text{NM}})^2}{2n} \quad \text{(from Eq. (20) and Lemma 1)} \\
&= \mathbb{P} \left( e_{i,A} \cdot e_{j,A} \right) + \frac{(P_{\text{NM}})^2}{2} \\
\end{align*}
\]
(21)

Since Hypothesis (19) holds, Lemma 11 ensures that
\[
\mathbb{P} \left( e_{i,A} \cdot e_{j,A} \right) \leq 17\eta \mathbb{P} \left( e_{i,A} \right) \mathbb{P} \left( e_{j,A} \right)
\]
From the above inequality and Eq. (21), we obtain
\[
\begin{align*}
\mathbb{P} \left( e_{i,A}' \cdot e_{j,A}' \right) &\leq 17\eta \mathbb{P} \left( e_{i,A} \right) \mathbb{P} \left( e_{j,A} \right) + \frac{(P_{\text{NM}})^2}{2} \\
&\leq 18\eta \mathbb{P} \left( e_{i,A} \right) \mathbb{P} \left( e_{j,A} \right) \\
&\quad \quad \quad \quad \text{(since } P_{\text{NM}} \leq \mathbb{P} \left( e_{i,A} \right) \mathbb{P} \left( e_{j,A} \right) ) \\
&= 18\eta \pi(C_{i,A}) \pi(C_{j,A}) \\
&\leq 18\eta \pi(C_{i,A}) \pi(C_{j,A}) + \frac{(P_{\text{NM}})^2}{2} (\nu(C_{j,A})) \\
&\quad \quad \quad \quad \text{+} \frac{(P_{\text{NM}})^2}{2} \text{ (from Eq. (20) and L.1)} \\
&\leq 18\eta \mathbb{P} \left( e_{i,A}' \right) + \frac{(P_{\text{NM}})^2}{2} (\mathbb{P} \left( e_{i,A}' \right) + \mathbb{P} \left( e_{j,A}' \right) + \frac{P_{\text{NM}}}{2}) \quad \text{(22)}
\end{align*}
\]
Since
\[
P_{\text{NM}} = \mathbb{P} \left( e_{i,j} \right)
= \pi \left( \left\{ (x_1, \ldots, x_n) \in \prod_{i=1}^n S \mid C(x_i, x_j) = 1 \right\} \right)
\]
from Eq. (20) and Lemma 1 we get that
\[
P_{\text{NM}} \leq \mathbb{P} \left( e_{i,j} \right) + \frac{P_{\text{NM}}^2}{2}
\]
It follows that
\[
\frac{P_{\text{NM}}}{2} \leq P_{\text{NM}} - \frac{P_{\text{NM}}^2}{2} \leq \mathbb{P} \left( e_{i,j}' \right) \leq \mathbb{P} \left( e_{i,A}' \right)
\]
From this and Eq. (22), we get Claim (b)
\[
\begin{align*}
\mathbb{P} \left( e_{i,A}' \cdot e_{j,A}' \right) &\leq 18\eta \left( 2 \mathbb{P} \left( e_{i,A}' \right) \right) \left( 2 \mathbb{P} \left( e_{j,A}' \right) \right) \\
&= 72\eta \mathbb{P} \left( e_{i,A}' \right) \mathbb{P} \left( e_{j,A}' \right)
\end{align*}
\]
4.2 Flooding in classic mobility models
4.2.1 Geometric mobility models

Several mobility models can be represented as special cases of node-MEGs. Many of these are continuous-space models [7, 26] in which nodes move over a subset of \( \mathbb{R}^d \). Since node-MEGs are discrete, we approximate continuous spaces by a suitable discretization. In the simplest and most common case, nodes move over a square of \( \mathbb{R}^2 \) of side length \( L \). The square can be discretized by taking a square grid \( Q \) formed by \( m \times m \) points regularly spaced in the square region, where \( m \) can be arbitrarily chosen.

In the standard random waypoint [7], \( n \) nodes independently move over the square: every node randomly chooses a speed in \( [v_{\text{min}}, v_{\text{max}}] \) where \( v_{\text{max}} = \Theta(v_{\text{min}}) \) and a destination point (waypoint) in the square and moves with the chosen speed on a straight path to this point. Then, it repeats the same process again and again. The destination points are uniformly distributed on the square. At any time two nodes are connected if they are at distance not larger than the transmission radius \( r \).
The formal discrete representation of the random waypoint as a node-MEG can be done by simple and standard arguments [15]. We now provide a description for one of the simplest versions where the node speed is constant (i.e. $v_{\text{min}} = v_{\text{max}}$).

- **Node states.** The generic state of the Markov chain $\mathcal{M}$ must encode the destination point and the current point in the straight point-path the node lies.
- **Matrix of state-transition probabilities.** The transition matrix can be easily defined: when a node is in some internal point of a path the choice of its next state is unique (so there is only one matrix entry having probability set to 1) while, when he arrives at the end of a path (i.e. a waypoint), its next state is randomly chosen by selecting the next destination point (and thus the next path to be followed).
- **Node connectivity.** The connection map $C$ is defined as follows: there is an edge between nodes $u$ and $v$ at time $t$ iff their relative positions (determined by their relative states) at time $t$ are at a distance not larger than the transmission radius $r$.

**Definition 5** The *positional* probability distribution in the stationary phase is defined as the probability that a node is in point $x$ (for any choice of $x$ over the square) when the state of the node is chosen independently at random according to the stationary distribution $\pi$ of $\mathcal{M}$. The density function of this distribution (in short, *positional function*) yielded by the random waypoint over the square will be denoted by $F_{wp}(\cdot)$.

The random waypoint model belongs to a general class of geometric mobility models called the *random trip model* [26] where the mobility space $\mathcal{R}$ can be any bounded connected region of $\mathbb{R}^d$ and the feasible node-trajectories can be any family of continuous curves. Any random trip model can be approximated up to any desired accuracy by a discretized model, by following the procedure described above for the random waypoint model. For this geometric class of node-MEGs, we have the following useful application of Theorem 2.

For any $r \geq 0$, $D(u, r)$ denotes the set of all the points that are within distance $r$ from $u$. For any connected region $B \subset \mathbb{R}^d$, define $B_r = \{u \in B \mid D(u, r) \subset B\}$ and $\text{vol}(B)$ as the volume of region $B$.

**Corollary 1** Let $\text{NM} = N M(n, \mathcal{M}, C)$ be a node-MEG yielded by a sufficiently-refined (in terms of grid resolution and time unit) discretization of a random trip model $\mathcal{T}$ over a bounded connected region $\mathcal{R} \subset \mathbb{R}^d$ with positional function $F_{\mathcal{T}}$. Assume that, for some $\delta \geq 1$ and $\lambda > 0$, it holds that

(a) $\forall u \in \mathcal{R}$, $F_{\mathcal{T}}(u) \leq \frac{\delta}{\text{vol}(\mathcal{R})}$,

(b) $\exists B \subseteq \mathcal{R}$ such that $\text{vol}(B_r) \geq \lambda \text{vol}(\mathcal{R})$ and $\forall u \in B$, $F_{\mathcal{T}}(u) \geq \frac{1}{\delta \text{vol}(\mathcal{R})}$,

(c) $P_{\text{NM}} \geq 1/n^{O(1)}$.

\[
\begin{align*}
\text{Then, with high probability, the flooding time is } &O(T_{\text{mix}}) \\
&\left(\frac{\delta^2 \text{vol}(\mathcal{R})}{\lambda n^{rd}} + \frac{n^6}{\lambda^2} \log^2 n\right), \text{ where } T_{\text{mix}} \text{ is the mixing time of the Markov chain } \mathcal{M}.
\end{align*}
\]

**Proof** Let $\pi$ be the stationary distribution of the Markov chain $\mathcal{M}$ of NM. For any $x \in \mathcal{S}$, let $u(x)$ be the point in $\mathcal{R}$ where a node lies when its state is $x$. By assuming that the states of the nodes are “stationary”, i.e., they are randomly chosen according to $\pi$, let $q(x)$ be the probability that a fixed node is connected to another fixed node being in state $x$. Clearly,

\[
q(x) = \sum_{y \in \mathcal{S} : u(y) \in D(u(x), r)} \pi(y)
\]

Since NM is a sufficiently refined discrete version of the random trip model $\mathcal{T}$, it holds that

\[
\forall x , \ q(x) = \int_{D(u(x), r)} F_{\mathcal{T}}(u)du = o(1)
\]

\[
\text{(in short, } q(x) \approx \int_{D(u(x), r)} F_{\mathcal{T}}(u)du) \tag{26}
\]

In general, with the expression $q(x) \approx f(x)$, we mean that the worst-case difference between the two functions $f$ and $q$ can be made arbitrarily small. From Hypothesis (b), for every $v \in B_r$, we have that

\[
\forall u \in D(v, r) \quad F_{\mathcal{T}}(u) \geq \frac{1}{\delta \text{vol}(\mathcal{R})}
\]

This implies that

\[
\int_{D(v, r)} F_{\mathcal{T}}(u)du \geq \frac{\text{vol}(D(v, r))}{\delta \text{vol}(\mathcal{R})} = \frac{cdr^d}{\delta \text{vol}(\mathcal{R})}, \tag{27}
\]

where $c_d$ is a constant depending only on $d$. By combining Eqs. (26) and (27), we get that, for every $x \in \mathcal{S}$ with $u(x) \in B_r$,

\[
q(x) \geq \frac{cdr^d}{\delta \text{vol}(\mathcal{R})} \tag{28}
\]
It thus holds that
\[
P_{NM} = \sum_{x \in S} \pi(x)q(x) \\
\geq \sum_{x \in S: u(x) \in B_x} \pi(x)q(x) \\
\geq \frac{c_d r^d}{\delta \text{vol}(R)} \sum_{x \in S: u(x) \in B_x} \pi(x) \quad \text{(from Eq. (28))} \\
\geq \frac{c_d r^d}{\delta \text{vol}(R)} \int_{B_x} F_T(u)du \\
\geq \frac{\lambda c_d r^d}{\delta \text{vol}(R)} \quad \text{(from Hyp. (b))}
\] (29)

From Eq. (26) and Hypothesis (a), we have that, for every \( x \in S \),
\[
q(x) \approx \int_{D(u(x), r)} F_T(u)du \leq \frac{\delta \text{vol}(D(u(x), r))}{\text{vol}(R)} = \frac{\delta c_d r^d}{\text{vol}(R)}
\] (30)

As for \( P_{NM2} \), it holds that
\[
P_{NM2} = \sum_{x \in S} \pi(x)q(x)^2 \\
\leq \left( \frac{\delta c_d r^d}{\text{vol}(R)} \right)^2 \sum_{x \in S} \pi(x) \quad \text{(from Eq. (30))} \\
= \frac{\delta^2 c_d^2 r^{2d}}{\text{vol}(R)^2}
\]

From the above inequality and Eq. (29), we get
\[
P_{NM2} \leq \frac{\delta^2 c_d^2 r^{2d}}{\text{vol}(R)^2} = \frac{\delta^6}{\lambda^2} \left( \frac{\lambda c_d r^d}{\delta \text{vol}(R)} \right)^2 \leq \frac{\delta^6}{\lambda^2} (P_{NM})^2
\]

It follows that the node-MEG NM satisfies the hypotheses of Theorem 2 with \( \eta = \delta^6/\lambda^2 \); so, thanks to Eq. (29), with high probability, the flooding time is
\[
O\left( \frac{T_{mix}}{nP_{NM}} + \frac{\delta^6}{\lambda^2} \right)^2 \log^2 n
\]
\[
\leq O\left( \frac{T_{mix}}{\lambda r^d} + \frac{\delta^6}{\lambda^2} \right)^2 \log^2 n
\]

The useful novelty of the above corollary lies in the following fact: the pairwise-independence condition in Theorem 2 is transformed into two mild “uniformity” conditions on the positional function yielded by the mobility model. The latter only refer to the stationary positional distribution of the single node and it is often much easier to verify with respect to the pairwise condition. Indeed, a general method (the Palm Calculus) to derive explicit formulas of such function for random trip models has been introduced in [26]. As for the random waypoint on the square, the explicit positional function \( F_{wp}(\cdot) \) has been derived in [27] and it is easy to verify that the two conditions of the above corollary are satisfied for some absolute constants \( \delta \) and \( \lambda \). Furthermore, the mixing time of the random waypoint over a square of side length \( L \) is \( \Theta(L/\nu_{max}) \) (recall that we are assuming \( \nu_{max} = O(\nu_{min}) \) [1, 31]. We thus obtain the following bound on the flooding time
\[
O\left( \frac{L}{\nu_{max}} \left( \frac{L^2}{nr^2} + 1 \right)^2 \log^2 n \right)
\]

Let us consider the case \( L \sim \sqrt{n}, r = \Omega(1) \); notice that this standard setting yields a stationary mobile network which is w.h.p. sparse and highly disconnected. Then the bound on the flooding time becomes \( O\left( \frac{\sqrt{n}}{\nu_{max}} \log^2 n \right) \) which almost matches the trivial lower bound \( \Omega\left( \frac{\sqrt{n}}{\nu_{max}} \right) \).

4.2.2 Graph mobility models

A natural generalization of random walks over a graph can be defined by considering random paths over a graph. This clearly includes the random waypoint over a graph. At every time step, a node moves along a path instead of on a single edge. More precisely, the model is specified by a graph \( H(V, A) \) (recall that \( H \) is the set of points) and a family \( \mathcal{P} \) of feasible paths in \( H \) satisfying the property: for every path \( h \in \mathcal{P} \), there is a path \( h' \in \mathcal{P} \) such that \( h' \) starts where \( h \) ends. For any point \( u \in V \), let \( \mathcal{P}(u) \) be the set of paths in \( \mathcal{P} \) that starts at point \( u \).

The above mobility model can be represented by a node-MEG in the following way.

- **Node states.** The generic state of the node’s Markov chain \( \mathcal{M} \) must encode the selected path \( h \) and the current point in the path the node lies on.
- **Node mobility.** A node at point \( u \in V \) chooses uniformly at random a path in \( \mathcal{P}(u) \), and then it travels along the path (an edge at the time); when it reaches the end point \( v \), it chooses uniformly at random a path in \( \mathcal{P}(v) \) and travels along that path, and so on. So, the dynamic behaviour of every node can be easily represented by an independent copy of a fixed finite-state Markov Chain (the same for all nodes). Formally, the Markov chain \( \mathcal{M}_{RP} = (S, P) \) is such that
Information spreading in dynamic graphs

\( S = \{ (h, h_i) \mid h \in \mathcal{P}, 2 \leq i \leq \ell(h) \} \) and
\( h_i \) is the \( i \)th point of \( h \).

For any path \( h \), let \( \ell(h) \) denote the number of points of \( h \). Then, the transition probabilities are as follows
\[
\forall h \in \mathcal{P} \quad \forall i : 2 \leq i \leq \ell(h) \quad P((h, h_i), (h, h_{i+1})) = 1
\]
\[
\forall h, h' \in \mathcal{P} : h_{\ell(h)} = h'_{\ell(h)}
\]
\[
P((h, h_{\ell(h)}), (h', h'_{\ell(h)})) = \frac{1}{|\mathcal{P}(h_{\ell(h)})|}
\]
all other transition probabilities are equal to zero;

- **Node connectivity.** We assume that two nodes are connected, at any given time \( t \), if they are in the same point of \( H(V, A) \) at time \( t \). Formally, the connection map is such that \( C_{\mathcal{P}}(h, h_i, (h', h'_i)) = 1 \) if \( h_i = h'_i \). Observe that if \( \mathcal{P} \) is the set of edges of \( H \) then the mobility model is equivalent to the random walk over \( H \).

We now define some natural properties inducing some subclasses of random-path models.

- **Simple and reversible models.** We say that a path \( h \in \mathcal{P} \) passes through a point \( u \) if \( h_i = u \) for some \( 2 \leq i \leq \ell(h) \). For any point \( u \in V \), let \( \#\mathcal{P}(u) \) be the number of paths in \( \mathcal{P} \) that passes through point \( u \). Notice that if \( \mathcal{P} \) is the set of edges of \( H \), then \( \#\mathcal{P}(u) = \deg_H(u) \).

A random-path model \( \mathcal{RP} = (H, \mathcal{P}) \) is said to be simple if every path in \( \mathcal{P} \) does not pass through the same point more than once, but the start and end points may equal.

Moreover, \( \mathcal{RP} \) is reversible if, for every path \( h \in \mathcal{P} \), the reversed path of \( h \) belongs to \( \mathcal{P} \) as well.

- **Regular Models.** Given any positive \( \delta \), we say that a random path model \( (H, \mathcal{P}) \) is \( \delta \)-regular if
\[
\forall u \in V \quad \#\mathcal{P}(u) \leq \delta \sum_{v \in V} \#\mathcal{P}(v) / |V|
\]

Roughly speaking \( \delta \)-regularity ensures that no point is a much busier crossroad than the average.

**Theorem 2** then implies the following useful result.

**Corollary 2** Let \( NM = NM(n, \mathcal{M}, C) \) be a node-MEG yielded by a random path model \( \mathcal{RP} = (H, \mathcal{P}) \) that is simple, reversible, \( \delta \)-regular, and \( |V| \leq n^{O(1)} \). Then, w.h.p the flooding time is \( O \left( T_{\text{mix}} \left( \frac{|V|}{n} + \delta^{3} \right)^{2} \log^{2} n \right) \) where \( T_{\text{mix}} \) is the mixing time of the Markov chain \( \mathcal{M} \).

**Proof** Since \( \mathcal{RP} \) is simple and reversible, Theorem 11 in [15] implies that the stationary distribution \( \pi \) of \( \mathcal{M} \) is uniform. For any state \( x \in S \), let \( u(x) \in V \) be the point where a node lies when its state is \( x \). By assuming that the states of the nodes are randomly chosen according to \( \pi \), let \( q(x) \) be the probability that a fixed node is connected to another fixed node being in state \( x \). Since \( \pi \) is uniform and \( \mathcal{RP} \) is simple, it holds that
\[
q(x) = \sum_{y \in S : C(y, x) = 1} \pi(y) = \frac{1}{|S|} |\{ y \in S \mid u(y) = u(x) \} = \frac{\#\mathcal{P}(u(x))}{|S|}
\]

It follows that
\[
\begin{align*}
P_{ NM} &= \sum_{x \in S} \pi(x) q(x) = \frac{1}{|S|^{2}} \sum_{x \in S} \#\mathcal{P}(u(x)) \\
&= \frac{1}{|S|^{2}} \sum_{u \in V} \#\mathcal{P}(u)^{2},
\end{align*}
\]

where the last equality is derived from \( |\{ y \in S \mid u(y) = u \} | = \#\mathcal{P}(u) \). Thanks to Jensen’s inequality it holds that
\[
\sum_{u \in V} \#\mathcal{P}(u)^{2} \geq \left( \sum_{u \in V} \#\mathcal{P}(u) \right)^{2} = \frac{|S|^{2}}{|V|^{2}}
\]

Thus, from Eq. (32), we have
\[
P_{ NM} \geq \frac{1}{|V|}
\]

From Eq. (31), it holds that
\[
P_{ NM2} = \sum_{x \in S} \pi(x) q(x)^{2} = \frac{1}{|S|^{3}} \sum_{x \in S} \#\mathcal{P}(u(x))^{2} = \frac{1}{|S|^{3}} \sum_{u \in V} \#\mathcal{P}(u)^{3}
\]

Since \( \mathcal{RP} \) is \( \delta \)-regular and Eq. (33) holds, it follows that
\[
P_{ NM2} = \frac{1}{|S|^{3}} \sum_{u \in V} \#\mathcal{P}(u)^{3} \leq \frac{1}{|S|^{3}} \sum_{u \in V} \left( \delta \sum_{v \in V} \#\mathcal{P}(v) / |V| \right)^{3} = \frac{\delta^{3}}{|V|^{2}} \leq \delta^{3} \left( P_{ NM} \right)^{2}
\]

From the above inequality, the hypothesis \( |V| \leq n^{O(1)} \) and Eq. (33), Theorem 2 can be applied with \( \eta = \delta^{3} \), thus obtaining that w.h.p. the flooding time is
\[
O \left( T_{\text{mix}} \left( \frac{|V|}{n} + \delta^{3} \right)^{2} \log^{2} n \right)
\]

\( \Box \)

If the feasible paths between nodes are the shortest one, then it is easy to verify that the mixing time of the relative Markov chain is \( O(D) \) where \( D \) is the diameter of \( H \). Moreover, if the model is \( \delta \)-regular for some \( \delta = \text{polylog}(n) \) and \( |V| = O(n \text{polylog}(n)) \), then the above corollary implies that
the flooding time is $O(D \log(n))$. This is within a polylogarithmic factor from the optimal bound $\Theta(D)$. A basic instance of this case is when $H$ is a grid.

As mentioned in the Introduction, almost tight bounds for the flooding time on the random walk model over grids have been recently obtained in [12,13,29]. In a general graph $H(V, A)$, every node randomly chooses its next position among all points in $V$ that are within $\rho$ hops from its current position. The transmission radius $r$ determines the maximal distance (again in terms of number of hops in $H(V, A)$) within which a message can be successfully transmitted. The most studied setting is $\rho = 1$ and $r = 0$: a node makes at most one hop per time step and it can inform only nodes that lies in the same point. This natural setting in general graphs has been studied in [16]: the flooding time is proved to be $O(T^* \log n)$ where $T^*$ is the meeting time between two independent random walks on $H$.

In what follows we apply our analysis for the random path model to the special case of random walks. The $\delta$-regularity condition over paths is transformed into a simple condition on the degree of the points. Given any $\delta \geq 1$, a graph $H(V, A)$ is said to be $\delta$-regular if $(\max\{\deg(v) \mid v \in V\})/(\min\{\deg(v) \mid v \in V\}) \leq \delta$. Then, we can derive a simple adaptation of Corollary 2.

**Corollary 3** Let $NM = NM(n, M, C)$ be a node-MEG yielded by the random walk over any $\delta$-regular mobility graph $H(V, A)$. Then, w.h.p the flooding time is $O\left(T_{\text{mix}}\left(\frac{\delta^2 |V|}{n} + \delta^3 \right)^2 \log^2 n\right)$ where $T_{\text{mix}}$ is the mixing time of a random walk over $H$.

The above bound improves the result in [16] over an important and general class of mobility graphs. Indeed, given a symmetric graph, the meeting time of two random walks is asymptotically equivalent to the first hitting time and can be much larger than the mixing time of a single random walk [1]. A natural example is that of $k$-augmented grids: take a grid of $s$ points and add an edge between any pair of points whose hop-distance is not larger than $k$. While the meeting time is not smaller than that of a standard grid $\Theta(s \log s)$ [1,29], the mixing time decreases in $k$. For instance, if $s \sim n \log(n)$ then the bound in [16] becomes $O((n \log(n))/k^2)$.

## 5 Generalized edge-MEGs

The link-based dynamic model Edge-MEG has been introduced in [11] and successively studied in [4,12,19]. In this restricted instance of MEGs, edges evolves independently. Every edge of the $n$-node graph can be in two states only: on or off. At any time step, every edge changes its state according to a two-state Markov chain with probabilities $p$ (edge birth-rate) and $q$ (edge death-rate): if at time $t$ an edge $e$ exists then it will die with probability $q$ while if $e$ does not exists then it will come up with probability $p$. In [11], the authors prove an almost tight upper bound on the flooding time:

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

A more refined model with four states has been recently introduced and studied in [5]. Edge-MEGs do not explicitly model node’s mobility, rather they are more suitable to model the link evolution in peer-to-peer networks or faulty networks.

Our main contribution here lies in the fact that Theorem 1 can be applied to the much more general version of edge-MEGs where an arbitrary (hidden) Markov chain $M = (S, P)$ rules the behavior of every edge and an arbitrary map $\chi : S \to \{0, 1\}$ determines, in function of the state, whether the edge exists or not. This generalization of edge-MEG will be denoted as $EM(n, M, \chi)$.

The initial state of each edge $\{i, j\}$ is chosen at random according to a probability distribution $\pi_{\{i,j\}}$ over the set of states $S$. We denote by $\pi$ the global initial probability distribution determined by the product of the probability distributions $\pi_{\{i,j\}}$. The state of an edge $\{i, j\}$ at any time $t$ is a random variable $s_{\{i,j\}}^t$ completely determined by the initial distribution $\pi_{\{i,j\}}$ and the Markov chain $M$.

Any model $EM(n, M, \chi)$ together with an initial probability distribution $\pi$ determines a dynamic graph $G(n, \{E_t\}_{t \geq 0})$ where, for any $t \geq 0$,

$$E_t = \{i, j \mid \chi(s_{\{i,j\}}^t) = 1\}$$

A crucial property of such generalized edge-MEGs is that edges are independent random variables, so it always holds that the $\beta$-independence is satisfied with $\beta = 1$. Then, when the Markov chain $M$ admits a unique stationary distribution $\pi$, Theorem 1 implies that the flooding time is

$$O\left(T_{\text{mix}}\left(\frac{1}{n\alpha} + 1\right)^2 \log n\right),$$

where $T_{\text{mix}}$ is the mixing time of the Markov chain $M$ and $\alpha$ is the probability an edge exists in the stationary regime (i.e., according to $\pi$). For instance, in the basic edge-MEG model with parameters $p$ and $q$ the mixing time is $\Theta(1/(p + q))$ and $\alpha = p/(p + q)$ [11]. We thus get an upper bound

$$O\left(\frac{1}{p + q} \left(\frac{p + q}{np} + 1\right)^2 \log n\right).$$

By comparing our bound to the (almost-tight) one in the bound (34), we get that the former is almost tight whenever $q \geq np$. 
6 Conclusions

We believe that Theorem 1 can be improved along some interesting directions. We suspect that, under mild assumptions on the dynamic graph process, the factor \((\frac{1}{ma} + \beta)^2\) can be improved. A more challenging task is to avoid the dependency of the bound on the mixing time of the graph process. The density and \(\beta\)-independence conditions can be met even at some state in which the graph process is far from the stationary distribution, and so a more refined analysis might be able to bound the flooding time without requiring that the process “wait” for stationarity.

Our method may prove useful in the analysis of more refined communication protocols than flooding. A simple example is a randomized protocol in which, at every step, an informed node transmits to a randomly chosen subset of neighbors. The analysis of such a process can be reduced to the analysis of flooding in a “virtual” dynamic graph in which a subset of the edges are removed. More general communication protocols might also be reduced to flooding by folding the actions of the protocol into the dynamic graph process.

Acknowledgments We thank Francesco Pasquale and Alessandro Pettarin for carefully reading a previous version of our paper and for providing helpful comments. We are also grateful to the anonymous referees that suggested several improvements of the paper: in particular, one of them showed us a simple way to improve our upper bounds by a logarithmic factor.

References

1. Aldous, D., Fill, J.A.: Reversible Markov Chains and Random Walks on Graphs, Chap. 14 (1999). http://www.stat.berkeley.edu/aldous/RWG/book.html
2. Avin, C., Kowcik, M., Lotker, Z.: How to explore a fast-changing world. In: Proceedings of 35th ICALP’08, LNCS, vol. 5125, pp. 121–132 (2008)
3. Azar, Y., Broder, A.Z., Karlin, A.R., Upfal, E.: Balanced allocation. SIAM J. Comput. 29(1), 180–200 (1999)
4. Baumann, H., Crescenzi, P., Fraigniaud, P.: Parsimonious flooding in dynamic graphs. Distrib. Comput. 24(1), 31–44 (2011)
5. Becchetti, L., Clementi, A.E.F., Pasquale, F., Resta, G., Santi, P., Silvestri, R.: Information spreading in opportunistic networks is fast. arXiv:1107.5241v1 (2011)
6. Bettstetter, C., Resta, G., Santi, P.: The node distribution of the random waypoint mobility model for wireless ad hoc networks. IEEE Trans. Mob. Comput. 2, 257–269 (2003)
7. Camp, T., Boleng, J., Davies, V.: A survey of mobility models for ad hoc network research. Wirel. Commun. Mob. Comput. 2(5), 483–502 (2002)
8. Camp, T., Navidi, W., Bauer, N.: Improving the accuracy of random waypoint simulations through steady-state initialization. In: Proceedings of 15th International Conference on Modelling and Simulation, pp. 319–326 (2004)
9. Chierichetti, F., Lattanzi, S., Panconesi, A.: Rumor spreading in social networks. Theor. Comput. Sci. 412(24), 2602–2610 (2011)
10. Clementi, A., Monti, A., Pasquale, F., Silvestri, R.: Broadcasting in dynamic radio networks. J. Comput. Syst. Sci. 75(4), 213–230 (2009)
11. Clementi, A., Macci, C., Monti, A., Pasquale, F., Silvestri, R.: Flooding time of edge-Markovian evolving graphs. SIAM J. Discrete Math. 24(4), 1694–1712 (2010)
12. Clementi, A., Monti, A., Pasquale, F., Silvestri, R.: Information spreading in stationary markovian evolving graphs. IEEE Trans. Parallel Distrib. Syst. 22(9), 1425–1432 (2011)
13. Clementi, A., Pasquale, F., Silvestri, R.: Opportunistic manets: high mobility can make up for low transmission power. IEEE/ACM Trans. Netw. 21(2), 610–620 (2013)
14. Clementi, A., Monti, A., Silvestri, R.: Flooding over Manhattan. Distrib. Comput. 26(1), 25–38 (2013)
15. Clementi, A., Monti, A., Silvestri, R.: Modelling mobility: a discrete revolution. Ad Hoc Netw. 9(6), 998–1014 (2011)
16. Dimitriou, T., Nikoletseas, S., Spirakis, P.: The infection time of graphs. Discrete Appl. Math. 154(8), 2577–2589 (2006)
17. Easley, D., Kleinberg, J.: Networks, Crowds, and Markets. Cambridge University Press, Cambridge (2010)
18. Ghosh, B., Leighton, F.T., Maggs, B.M., Muthukrishnan, S., Plaxton, C.G., Rajaraman, R., Richa, A.W., Tarjan, R.E., Zuckerman, D.J.: Tight analyses of two local load balancing algorithms. SIAM J. Comput. 29(1), 29–64 (1999)
19. Grindrod, P., Higham, D.J.: Evolving graphs: dynamical models, inverse problems and propagation. Proc. R. Soc. A 466(2115), 753–770 (2010)
20. Jacquet, P., Mans, B., Rodolakis, G.: Information propagation speed in mobile and delay tolerant networks. IEEE Trans. Inf. Theory 56, 5001–5015 (2010)
21. Karagiannis, T., Le Boudec, J.-Y., Vojnovic, M.: Power law and exponential decay of inter contact times between mobile devices. In: Proceedings of 13th ACM MOBICOM, pp. 183–194 (2007)
22. Kesten, H., Sidoravicius, V.: The spread of a rumor or infection in a moving population. Ann. Probab. 33(6), 2402–2462 (2005)
23. Khun, F., Lynch, N., Oshman, R.: Distributed computation in dynamic networks. In: Proceedings of 42nd ACM STOC, pp. 513–522 (2010)
24. Khun, F., Oshman, R.: Dynamic networks: models and algorithms. ACM SIGACT News 42(1), 82–96 (2011)
25. Lam, H., Liu, Z., Mitzenmacher, M., Sun, X., Wang, Y.: Information dissemination via random walks in d-dimensional space. In: Proceedings of 23rd ACM-SIAM SODA, pp. 1612–1622 (2012)
26. Le Boudec, J.-Y., Vojnovic, M.: The random trip model: stability, stationary regime, and perfect simulation. IEEE/ACM Trans. Netw. 14(6), 1153–1166 (2006)
27. Le Boudec, J.-Y.: Understanding the simulation of mobility models with palm calculus. Perform. Eval. 64(2), 126–147 (2007)
28. Peres, Y., Sinclair, A., Sousi, P., Stauffer, A.: Mobile geometric graphs: detection, coverage and percolation. In: Proceedings of 22nd ACM-SIAM SODA, pp. 412–428 (2011)
29. Pettarin, A., Pietracaprina, A., Pucci, G., Upfal, E.: Tight bounds on information dissemination in sparse mobile networks. In: Proceedings of 30th ACM PODC, pp. 355–362 (2011)
30. Rabani, Y., Sinclair, A., Wanka, R.: Local divergence of Markov chains and the analysis of iterative load-balancing schemes. In: Proceedings of 39th IEEE FOCS, pp. 694–703 (1998)
31. Spyropoulos, T., Jindal, A., Psounis, K.: An analytical study of fundamental mobility properties for encounter based protocols. Int. J. Auton. Adapt. Comput. Syst. 1(1), 4–40 (2008)
32. Vojnovic, M., Proutier, A.: Hop limited flooding over dynamic networks. In: Proceedings of 30th IEEE INFOCOM, pp. 685–693 (2011)
33. Whitebeck, J., Conan, V., de Amorim, M.D.: Performance of opportunistic epidemic routing on edge-Markovian dynamic graphs. IEEE Trans. Commun. 59(5), 1259–1263 (2011)