Random walks on randomly evolving graphs*

Leran Cai, Thomas Sauerwald, and Luca Zanetti

University of Cambridge, Cambridge, UK
{lc647,tms41,lz381}@cam.ac.uk

Abstract. A random walk is a basic stochastic process on graphs and a key primitive in the design of distributed algorithms. One of the most important features of random walks is that, under mild conditions, they converge to a stationary distribution in time that is at most polynomial in the size of the graph. This fundamental property, however, only holds if the graph does not change over time; on the other hand, many distributed networks are inherently dynamic, and their topology is subjected to potentially drastic changes.
In this work we study the mixing (i.e., convergence) properties of random walks on graphs subjected to random changes over time. Specifically, we consider the edge-Markovian random graph model: for each edge slot, there is a two-state Markov chain with transition probabilities p (add a non-existing edge) and q (remove an existing edge). We derive several positive and negative results that depend on both the density of the graph and the speed by which the graph changes.

Keywords: Random Walks · Evolving Graphs · Mixing Times

1 Introduction

A random walk on a network is a simple stochastic process, defined as follows. Given an undirected graph $G = (V, E)$, the walk starts at a fixed vertex. Then, at each step, the random walk moves to a randomly chosen neighbor [1]. Due to their simplicity and locality, random walks are very useful algorithmic primitive, especially in the design of distributed algorithms. In contrast to topology-driven algorithms, algorithms based on random walks usually benefit from a strong robustness against structural changes in the network.

Random walks and related works have found various applications such as routing, information spreading, opinion dynamics, and graph exploration [3, 9]. One key property of random walks is that, under mild assumptions on the underlying network, they converge to a stationary distribution – an equilibrium state in which every vertex is visited proportionally to its degree. The time for this convergence to happen is called *mixing time*, and understanding this time is

* The second and third author acknowledge support by the ERC Starting Grant "Dynamic March".

[1] In case of a *lazy* random walk, the walk would remain at the current location with probability 1/2, and otherwise move to a neighbor chosen uniformly at random.

crucial for many sampling and exploration related tasks. In particular, whenever a graph has a small *mixing time*, also its *cover time* (the expected time to visit all vertices of the graph) is small as well.

While most of the classical work devoted to understanding random walks has focused on static graphs, many networks today are subject to dramatic changes over time. Hence understanding the theoretical power and limitations of dynamic graphs has become one of the key challenges in computer science [17]. Several recent works have indeed considered this problem and studied the behavior of random walks [2, 3, 10, 15, 21, 22, 24] or similar processes [4, 5, 8, 11, 14] on such dynamic graphs, and their applications to distributed computing [2, 14, 24].

In this work, we study the popular *evolving graph model*. That is, we consider sequences of graphs $G_1, G_2, \ldots$ over the same set of vertices but with a varying set of edges. This model has been studied in, for example, [3, 15, 25]. Both [3] and later [25] proved a collection of positive and negative results about the mixing time (and related parameters), and they assume a worst-case scenario where the changes to the graph are dictated by an oblivious, non-adaptive adversary. For example, [3] proved the following remarkable dichotomy. First, even if all graphs $G_1, G_2, \ldots$ are connected, small (but adversarial) changes to the stationary distribution can cause exponential mixing (and hitting) times. Secondly, if the sequence of connected graphs share the same stationary distribution, i.e., the degrees (or relative degrees) of vertices are time-invariant, then mixing and hitting times are polynomial. This assumption about a time-invariant stationary distribution is crucial in the majority of the positive results in [3, 25].

In contrast to [3, 25], we do not impose such assumptions, but instead study a model with incremental changes. Specifically, we consider a setting where the evolving graph model changes *randomly* and study the so-called *edge-Markovian random graph* $\mathcal{G}(n, p, q)$, which is defined as follows (see Definition 2.5 for a more formal description). For each edge slot, there is a two-state Markov chain that switches from off to on with probability p and from on to off with probability q. This model can be seen as a dynamic version of the Erdős-Rényi random graph, and has been studied in the context of information spreading and flooding [5–7]. While these results demonstrate that information disseminates very quickly on these dynamic graphs, analysing the convergence properties of a random walk seems to require new techniques, since degree fluctuations make the use of any "inductive" argument very difficult – from one step to another, the distribution of the walk could become "worse", whereas the set of informed (or reachable) nodes can never decrease.

In this work, we will investigate the mixing time of a random walk on such evolving graphs. It turns out that, as our results demonstrate, the mixing time depends crucially on the density as well as on the speed by which the graph changes. We remark that deriving bounds on the mixing time on $\mathcal{G}(n, p, q)$ poses some unique challenges, which are not present in the positive results of [3, 25]. The main difficulty is that in $\mathcal{G}(n, p, q)$, due to the changing degrees of the vertices, there is no time-invariant stationary distribution, and the traditional notion of mixing time must be adapted to our dynamic setting. Informally, what

we ask, then, is how many steps the walk needs to take before the distance to a *time-dependent* stationary distribution becomes *small enough*. Furthermore, in contrast to static graphs, where the distance between the distribution of the walk and the stationary distribution can only decrease, in dynamic graphs the distance to the time-dependent stationary distribution might increase with time. For this reason, we also ask that the distribution of the walk remains close to a time-dependent stationary distribution for a *long enough* interval of time (for a precise definition of our notion of mixing time, see Definition 2.7). We believe this requirement is necessary for our definition of mixing time to be useful in potential applications.

Further Related Work. Recently, [15] analysed the cover time of so-called "Edge-Uniform Stochastically-Evolving Graphs", that include our model as a special case (i.e., the history is $k = 1$). Their focus is on a process called "Random Walk with a Delay", where at each step the walk picks a (possible) neighbor and then waits until the edge becomes present. In [15, Theorem 4], the authors also relate this process to the standard random walk, and prove a worst-case upper bound on the cover time. However, one of the key differences to [15] is that we will study the *mixing time* instead of the *cover time*.

In [26], the authors analysed a continuous-time version of the edge-Markovian random graph. However, unlike the standard random walk, they consider a slightly different process: when the random walk tries to make a transition from a vertex u, it picks one of the $n - 1$ other vertices and moves there *only if* the edge is present; otherwise it remains in place. For this process, they were able to derive very tight bounds on the mixing time and establish a cutoff phenomenon. The same random walk was also analysed on a dynamic graph model of the d-dimensional grid in [19, 20] and, more generally, in [12].

1.1 Main Results

We study the mixing properties of random walks on edge-Markovian random graphs $\mathcal{G}(n, p, q)$. In particular, we consider six different settings of parameters p and q, which separates edge-Markovian models based on how fast graphs change over time (slowly vs. fast changing), and how dense graphs in the dynamic sequence are (sparse vs. semi-sparse vs dense).

As noted in previous works (see, e.g., [5]), a dynamic sequence sampled from $\mathcal{G}(n, p, q)$ will eventually converge to an Erdős-Rényi random graph $\mathcal{G}(n, \tilde{p})$ where $\tilde{p} = \frac{p}{p+q}$ (for the sake of completeness, we give a proof of this fact in Appendix A). We use the expected degree in such random graph, which is equal to $d = (n-1)\tilde{p}$, to separate edge-Markovian models according to their density as follows:

1. **Sparse** $d = o(\log n)$
2. **Semi-sparse** $d = \Theta(\log n)$
3. **Dense** $d = \omega(\log n)$.

Notice that the sparse regime corresponds to random graphs with density below the connectivity threshold of Erdős-Rényi random graphs.

We further separate edge-Markovian models based on how fast they change over time. Let $\delta = \binom{n}{2}\tilde{p}q + \binom{n}{2}(1-\tilde{p})p$ be the expected number of changes at each step, when starting from a stationary initial graph $G_0 \sim \mathcal{G}(n, \tilde{p})$. We consider the following two opposite regimes.

1. **Fast-changing** $\delta = \Theta(dn)$.
2. **Slowly-changing** $\delta = O(\log n)$

Notice that the fast-changing regime corresponds to graphs for which a constant fraction of edges change at each step in expectation.

	Fast-changing $\delta = \Theta(dn)$	**Slowly-changing** $\delta = O(\log n)$
Sparse $d \in [1, o(\log n)]$	$t_{\text{mix}} = \infty$ Thm 1.1	$t_{\text{mix}} = \Omega(n)$ Proposition 1.4
Semi-sparse $d = \Theta(\log n)$	Coarse mixing[2] in $O(\log n)$ Prop 3.2	$t_{\text{mix}} = O(\log n)$, Thm 1.3
Dense $d \in [\omega(\log n), n/2]$	$t_{\text{mix}} = O(\log n)$ Thm 1.2	

Table 1.1. Summary of our main results (informal). See referenced theorems for the precise and complete statements.

The main results of our work are presented in Table 1.1. Here, we assume G_0 is sampled from the stationary graph distribution $\mathcal{G}(n, \tilde{p})$. In the fast-changing regime, as highlighted in Remark A.1, this is without loss of generality. For slowly-changing models, instead, different choices of G_0 can result in drastically different outcomes with regard to the mixing time. For ease of presentation, we assume in Table 1.1 that $G_0 \sim \mathcal{G}(n, \tilde{p})$, but this assumption can usually be relaxed, and we refer to the full statement of the corresponding results for our actual assumptions on G_0.

Next, we formally state the four main results of our work. The formal definitions of mixing time for random walk on dynamic graphs will be presented in Section 2.1 (see in particular Definition 2.7 and Definition 2.8). The first theorem is a negative result that tells us that, for fast-changing and sparse edge-Markovian graphs, random walks don't have finite mixing time. Its proof will be presented in Section 3.1.

[2] In this regime we are not able to prove finite mixing time. However, we show that the distribution of the walk will "flatten out" after $O(\log n)$ steps. We refer to this behavior as *coarse* mixing.

Theorem 1.1 (Fast-changing and sparse, no mixing). *Let $p = \Theta(1/n)$ and $q = \Omega(1)$. Then, $t_{\text{mix}}(\mathcal{G}(n,p,q)) = \infty$.*

The following theorem is a positive result that establishes fast mixing time in the dense and fast-changing regime. Its proof is presented in Section 3.2.

Theorem 1.2 (Fast-changing and dense, fast mixing). *Let $p = \omega(\log n/n)$ and $q = \Omega(1)$. Then, $t_{\text{mix}}(\mathcal{G}(n,p,q)) = O(\log n)$.*

The only case missing in the fast-changing regime is the semi-sparse case, where nodes have average degree $d = \Theta(\log n)$. We do not have a definitive answer on the mixing time of random walks in such case, however, we do have a partial result that guarantees at least that random walk distributions will be "well spread" over a large support after $O(\log n)$ steps (we call this behavior *coarse mixing*). This statement can be made formal by considering the ℓ_2-norm of the distribution of the walk. Because of its technical nature, we defer the formal statement to Section 3.2 and Proposition 3.2.

We now turn our attention to the slowly-changing regime, where at most $\delta = O(\log n)$ edges are created and removed at each step. Unlike the results for the fast-changing regime, where the choice of the starting graph G_0 does not really affect the mixing time of a random walk (see Appendix A and Remark A.1 for a discussion), in the slowly-changing regime the choice of G_0 will affect the properties of G_t for a large number of steps t.

The following theorem shows that in the slowly-changing and dense regime, under mild conditions on the starting graph $G_0 = (V, E_0)$ (which are satisfied for G_0 drawn from the limiting distribution of dense $\mathcal{G}(n,p,q)$), random walks will mix relatively fast. We use $E_0(S, V \setminus S)$ to indicate the set of edges in G_0 between a subset of vertices $S \subset V$ and its complement, and Φ_{G_0} to indicate the minimum conductance of G_0 (see Definition 2.2).

Theorem 1.3 (Slowly-changing and dense, fast mixing). *Let $d = \Omega(\log n)$, $p = O(\log n/n^2)$, and $q = O(\log n/(dn))$. Let the following assumptions on the starting graph $G_0 = (V, E_0)$ be satisfied for large enough constants $c_1, c_2, c_3 > 0$.*

(1) $\deg_0(x) = \Theta(d)$ *for any* $x \in V$*;*
(2) $|E_0(S, V \setminus S)| \geq c_2 \log n |S|$*, for any* $S \subset V$ *with* $|S| \leq c_1 \log n$*;*
(3) $\Phi_{G_0} \geq c_3 \log d/d$*.*

Then, $t_{\text{mix}}(\mathcal{G}(n,p,q)) = O(\log n/\Phi^2_{G_0})$.

Let us briefly discuss the assumptions and results of Theorem 1.3. First of all notice that the parameters p and q are defined so that the average degree is $d = \Omega(\log(n))$ and the number of changes in the graph at each step is $\delta = O(\log(n))$. Assumption (1) just requires the degree of the vertices in G_0 to be of the same order as the degree of the vertices in the limiting graph $\mathcal{G}(n,\tilde{p})$. Assumption (2) guarantees that for any small set S there are enough edges going from S to the rest of the graph. Assumption (3) is a mild condition on the conductance of G_0. The last two assumptions ensure that the conductance

of G_t will not be much lower than the conductance of G_0 for a large number of steps t. Finally, notice that $O(\log n/\Phi^2_{G_0})$ is a classic bound for the mixing time of a *static* random walk on G_0. Theorem 1.3 essentially states that, if the three assumptions are satisfied, the mixing time of a random walk on $\mathcal{G}(n,p,q)$ will not be much larger. In particular, all the three assumptions are satisfied for a starting graph $G_0 \sim \mathcal{G}(n,\tilde{p})$ with $\tilde{p} = p/(p+q)$. Furthermore, in such case $t_{\text{mix}}(\mathcal{G}(n,p,q)) = O(\log n)$. The proof of this theorem can be found in Section 4.1.

We conclude this section by stating our result in the slowly-changing and dense regime. We prove a negative result: we show that the mixing time of $\mathcal{G}(n,p,q)$ is at least linear in n.

Proposition 1.4 (Slowly-changing and sparse, slow mixing). *Let $p = O(1/n^2)$ and $q = \omega(1/(n\log n))$. Consider a random walk on $\mathcal{G}(n,p,q)$ with starting graph $G_0 \sim \mathcal{G}(n,\tilde{p})$ with $\tilde{p} = p/(p+q)$. Then, $t_{\text{mix}}(\mathcal{G}(n,p,q)) = \Omega(n)$.*

2 Notation and Definitions

2.1 Random Walk and Conductance

In this section we introduce the relevant notation and basic results about Markov chains that we will use throughout the paper. For more background on Markov chains and random walks we defer the reader to [16].

Let $\mathscr{G} = (G_t)_{t\in\mathbb{N}}$ be a sequence of undirected and unweighted graphs defined on the same vertex set V, with $|V| = n$, but with potentially different edge-sets E_t ($t \in \mathbb{N}$). We study (lazy) random walks on $\mathscr{G}$: suppose that at a time $t \geq 0$ a particle occupies a vertex $u \in V$. At step $t+1$ the particle will remain at the same vertex u with probability $1/2$, or will move to a random neighbor of u in G_t. In other words, it will perform a single random walk step according to a transition matrix P_t, which is the transition matrix of a lazy random walk on G_t: $P_t(u,u) = 1/2$, $P_t(u,v) = 1/(2\deg_t(u))$ (where $\deg_t(u)$ is the degree of u in G_t) if there is an edge between u and v in G_t , or $P_t(u,v) = 0$ otherwise.

Given an initial probability distribution $\mu_0 \colon V \to [0,1]$, which is the distribution of the initial position of the walk, the t-step distribution of a random walk on $\mathscr{G}$ is equal to $\mu_t = \mu_0 P_1 \cdot P_2 \cdot \ldots \cdot P_t$. In particular, we use μ^x_t to denote the t-step distribution of the random walk starting at a vertex $x \in V$. Hence $\mu^x_0(x) = 1$ and $\mu^x_0(y) = 0$ for $x \neq y \in V$. Furthermore, we use π_t to denote the probability distribution with entries equal to $\pi_t(x) = \deg_t(x)/(2|E_t|)$ for any $x \in V$. This distribution is stationary for P_t (i.e, it satisfies $\pi_t P_t = \pi_t$) and, if G_t is connected, it is the unique stationary distribution of P_t. If G_t is disconnected, P_t will have multiple stationary distribution. However, unless stated otherwise, we will consider only the "canonical" stationary distribution π_t. Finally, while any individual P_t is *time-reversible* (it satisfies $\pi_t(x)P_t(x,y) = \pi_t(y)P_t(y,x)$ for any $x, y \in V$), a random walk on $\mathcal{G}$ may not. [3]

[3] For example, it might happen that $P_1 \cdots P_t(x,y) > 0$ while $P_1 \cdots P_t(y,x) = 0$. This cannot happen in the "static" case where $P_1 = \cdots = P_t = P$ with P reversible.

Recall that if P is a transition matrix of a reversible Markov chain, it has n real eigenvalues, which we denote with $-1 \leq \lambda_n(P) \leq \cdots \leq \lambda_1(P) = 1$. If P is the transition matrix of a lazy random walk on a graph G, it holds that $\lambda_n(P) \geq 0$. Moreover, $\lambda_1(P) < 1$ if and only if G is connected

For two probability distributions $f, g \colon V \to [0,1]$, the *total variation distance* between f and g is defined as $\|f-g\|_{TV} := \frac{1}{2}\sum_{x\in V} |f(x)-g(x)|$. We denote with $\|f\|_2 = \left(\sum_{x\in V} f^2(x)\right)^{1/2}$ and $\|f\|_\infty = \max_{x\in V}|f(x)|$ the standard ℓ_2 and ℓ_∞ norms of f. Given a probability distribution $\pi \colon V \to \mathbb{R}_+$, we also define the $\ell_2(\pi)$-norm as $\|f\|_{2,\pi} := \sqrt{\sum_{x\in V} f^2(x)\pi(x)}$. By Jensen's inequality, it holds for any f, g that $2\cdot\|f-g\|_{TV} \leq \|f-g\|_{2,\pi}$. The lemma below relates the decrease in the distance to stationarity after one random walk step to the spectral properties of its transition matrix.

Lemma 2.1 (Lemma 1.13 in [18], rephrased). *Let P be the transition matrix of a lazy random walk on a graph $G = (V, E)$ with stationary distribution π. Then, for any $f : V \to \mathbb{R}$, we have that*

$$\left\| \frac{fP}{\pi} - \mathbf{1} \right\|_{2,\pi}^2 \leq \lambda_2(P)^2 \left\| \frac{f}{\pi} - \mathbf{1} \right\|_{2,\pi}^2 .$$

In the lemma above and throughout the paper, a division between two functions is to be understood entry-wise, while $\mathbf{1}$ refers to a function always equal to one. An important quantity which can be used to obtain bounds on $\lambda_2(P)$ is the *conductance* of G, which is defined as follows.

Definition 2.2. *The conductance of a non-empty set $S \subseteq V$ in a graph G is defined as:*

$$\Phi_G(S) := \frac{|E(S, V \setminus S)|}{\mathrm{vol}(S)},$$

where $\mathrm{vol}(S) := \sum_{x\in V} \deg(x)$ *and $E(S, V \setminus S)$ is the set of edges between S and $V \setminus S$. The conductance of the entire graph G is defined as*

$$\Phi_G := \min_{\substack{S\subset V:\\ 1\leq \mathrm{vol}(S)\leq \mathrm{vol}(V)/2}} \frac{|E(S, V \setminus S)|}{\mathrm{vol}(S)}.$$

The conductance of G and the second largest eigenvalue of the transition matrix P of a lazy random walk in G are related by the so-called discrete Cheeger inequality [1], which we state below.

Theorem 2.3 (Cheeger inequality). *Let P be the transition matrix of a lazy random walk on a graph G. Then, it holds that*

$$1 - \lambda_2(P) \leq \Phi_G \leq 2\sqrt{1-\lambda_2(P)}.$$

Finally, we use the notation $o_n(1)$ to denote any function $f : \mathbb{N} \to \mathbb{R}$ such that $\lim_{n\to+\infty} f(n) = 0$. We often drop the subscript n.

2.2 Dynamic graph models

In this section we formally introduce the random models of (dynamic) graphs that are the focus of this work. We start by recalling the definition of the Erdős-Rényi model of (static) random graphs.

Definition 2.4 (Erdős-Rényi model). *$G = (V, E) \sim \mathcal{G}(n, p)$ is a random graph such that $|V| = \{1, \dots, n\}$ and the $\binom{n}{2}$ possible edges appear independently, each with probability p.*

We now introduce the *edge-Markovian* model of dynamic random graphs, which has been studied both in the context of information spreading in networks [5,6] and random walks [15]. This model is the focus of our work.

Definition 2.5 (edge-Markovian model). *Given a starting graph G_0, we denote with $(G_t)_{t\in\mathbb{N}} \sim \mathcal{G}(n, p, q)$ a sequence of graphs such that $G_t = (V, E_t)$, where $V = \{1, \dots, n\}$ and, for each $t \in \mathbb{N}$, any pair of distinct vertices $u, v \in V$ will be connected by an edge in G_t independently at random with the following probability:*

$$\mathbb{P}\left[\{u, v\} \in E_{t+1} \mid G_t\right] = \begin{cases} 1 - q & \text{if } \{u, v\} \in E_t \\ p & \text{if } \{u, v\} \notin E_t. \end{cases}$$

Notice that different choices of a starting graph G_0 will induce different probability distributions over $(G_t)_{t\in\mathbb{N}}$. In general, we try to study $\mathcal{G}(n, p, q)$ by making the fewest possible assumptions on our choice of G_0. Moreover, as pointed out for example in [15], $(G_t)_{t\in\mathbb{N}} \sim \mathcal{G}(n, p, q)$ converges to $\mathcal{G}(n, \tilde{p})$ with $\tilde{p} = p/(p + q)$. We leave considerations about the speed of this convergence and how this affects our choice of G_0 to Appendix A and, in particular, Remark A.1 .

2.3 Mixing time of random walks on dynamic graphs

One of the most studied quantities in the literature about time-homogeneous (i.e., static) Markov chains (random walks included) is the mixing time, i.e., the time it takes for the distribution of the chain to become close to stationarity. Formally, it is defined as follows.

Definition 2.6 (Mixing time for time-homogeneous Markov chains). *Let μ_t^x be the t-step distribution of a Markov chain with state space V starting from $x \in V$. Let π be its stationary distribution. For any $\epsilon > 0$, the ϵ-mixing time is defined as*

$$t_{\text{mix}}(\epsilon) := \min\{t \in \mathbb{N} : \max_{x\in V} \|\mu_t^x - \pi\|_{TV} \leq \epsilon\}.$$

A basic fact in random walk theory states that a lazy random walk on a connected undirected graph $G = (V, E)$ has always a finite mixing time. In particular, if $|V| = n$, $t_{\text{mix}}(1/4) = O(n^3)$. Moreover, considering a different ϵ does not significantly change the mixing time: for any $\epsilon > 0$, $t_{\text{mix}}(\epsilon) =$

$O(t_{\text{mix}}(1/4)\log(1/\epsilon))$ (see, e.g., [16]). Also, it is a well-known fact that $\|\mu_t^x - \pi\|_{TV}$ is non-increasing.

However, in the case of random walks on dynamic graphs, convergence to a time-invariant stationary distribution does not, in general, happen. For this reason, other works have studied alternative notions of mixing for dynamic graphs, such as merging [23], which happens when a random walk "forgets" the vertex where it started. In this work, instead, we focus on a different approach that we believe best translates the classical notion of mixing from the static to the dynamic case. More precisely, let us consider a dynamic sequence of graphs $(G_t)_{t\in\mathbb{N}}$ with corresponding stationary distributions $(\pi_t)_{t\in\mathbb{N}}$. Our goal is to establish if there exists a time t such that the distribution μ_t of the walk at time t is close to π_t. Moreover, to make this notion of mixing useful in possible applications, we require that μ_s remains close to π_s for a reasonably large number of steps $s \geq t$. Formally, we introduce the following definition of mixing time for dynamic graph sequences.

Definition 2.7 (Mixing time for dynamic graph sequences). *Let $\mathcal{G} = (G_t)_{t\in\mathbb{N}}$ be a dynamic graph sequence on a vertex set V, $|V| = n$. The mixing time of a random walk in $\mathcal{G}$ is defined as*

$$t_{\text{mix}}(\mathcal{G}) = \min\left\{t \in \mathbb{N} \colon \forall s \in [t, t+\sqrt{n}),\ \forall x \in V,\ \ \|\mu_s^x - \pi_s\|_{TV} = o_n(1)\right\},$$

where π_s is the stationary distribution of a random walk in G_s, and μ_s^x is the s-step distribution of a random walk in $\mathcal{G}$ that started from $x \in V$.

First observe we require that the total variation distance between μ_s and π_s goes to zero as the number of vertices increases.[4] This is motivated by the fact that the distance to stationarity, unlike in the static case, might not tend to zero as the number of steps t goes to infinity. However, we ask that the distance to stationarity is smaller than a threshold which decreases for larger sized graphs. Secondly, we require that such distance remains small for $\sqrt{n}$ steps (recall n is the number of vertices in the graph). This is due to the fact that, for all dynamic graph models we consider, we cannot hope for such distance to stay small arbitrarily long. However, we believe that $\sqrt{n}$ steps is a long enough period of time for mixing properties to be useful in applications.

Since our goal is to study the mixing property of $\mathcal{G}(n,p,q)$, we now introduce a definition of mixing time for edge-Markovian models that takes into account the probabilistic nature of such graph sequences. Essentially, we say that the mixing time of $\mathcal{G}(n,p,q)$ is t if a random walk on a dynamic sequence of graphs sampled from $\mathcal{G}(n,p,q)$ mixes (according to the previous definition) in t steps with high probability over the sampled dynamic graph sequence.

Definition 2.8 (Mixing time for edge-Markovian models). *Given an edge-Markovian model $\mathcal{G}(n,p,q)$, its mixing time is defined as*

$$t_{\text{mix}}(\mathcal{G}(n,p,q)) = \min\left\{t \in \mathbb{N}\colon\ \mathbb{P}_{\mathcal{G}\sim\mathcal{G}(n,p,q)}\left[t_{\text{mix}}(\mathscr{G}) \leq t\right] \geq 1 - o_n(1)\right\}.$$

[4] We are implicitly assuming there is an infinite family of dynamic graph sequences with increasing n.

Finally, we remark that, while in static graphs connectivity is a necessary prerequisite to mixing, random walks on sequences of disconnected dynamic graphs might nonetheless exhibit mixing properties. Examples of this behavior were studied in [25].

3 Results for the fast-changing case

3.1 Negative result for mixing in the sparse and fast-changing case

In this section we consider random walks on sparse and fast-changing edge-Markovian graphs. In particular, we study $\mathcal{G}(n,p,q)$ with $0 < q = \Omega(1)$ and $p = \frac{1}{n}$. Since $\Omega(1)$, by Remark A.1, we can restrict ourselves to consider the case where $G_0 \sim \mathcal{G}(n,\tilde{p})$ with $\tilde{p} = p/(p+q)$. We prove the following theorem.

Theorem 1.1 (Fast-changing and sparse, no mixing). *Let $p = \Theta(1/n)$ and $q = \Omega(1)$. Then, $t_{\text{mix}}(\mathcal{G}(n,p,q)) = \infty$.*

The key idea behind this result is that, due to the fast-changing nature of graphs in this model, the degrees of the nodes also change rapidly. In particular, for a linear number of nodes such as u, there is at least one neighbor $v_{\min}$ in the neighbors of u whose degree may change from one constant in round t to basically any other constant (this also makes use of the assumption on p, ensuring that the graph is sparse). The proof then exploits that, due to the "unpredictable" nature of this change, the probability mass received by $v_{\min}$ in round $t+1$ is likely to cause a significant difference between $\mu_{t+1}(u)$ and $\pi_{t+1}(u)$. Since this holds for a linear number of nodes u, we obtain a sufficiently large lower bound on the total variation distance, and the theorem is established. The complete proof will appear in the full version of the paper.

3.2 Positive result for mixing in the dense and fast-changing case

In this section we analyse the mixing properties of $\mathcal{G}(n,p,q)$ for $p = \Omega(\log n/n)$ and $q = \Omega(1)$. Since q is large, for simplicity we will assume throughout this section that $G_0 \sim \mathcal{G}(n,\tilde{p})$, where $\tilde{p} = \frac{p}{p+q}$ (see Remark A.1 for an explanation of why this is not a restriction). The following theorem is the main result.

Theorem 1.2 (Fast-changing and dense, fast mixing). *Let $p = \omega(\log n/n)$ and $q = \Omega(1)$. Then, $t_{\text{mix}}(\mathcal{G}(n,p,q)) = O(\log n)$.*

While in this paper we study for simplicity only lazy random walks on graphs, to prove Theorem 1.2, however, we need to introduce *simple* random walks on graphs: given a graph $G = (V,E)$, a simple random walk on G has transition matrix Q such that, for any $x, y \in V$, $Q(x,y) = 1/\deg(x)$ if $\{x,y\} \in E$, $Q(x,y) = 0$ otherwise. The following lemma, whose proof is the main technical part of the section, shows that if the *simple* random walk on a sequence of graphs $\mathcal{G} = (G_t)_{t\in\mathbb{N}}$ exhibits strong expansion properties, and the time-varying stationary

distribution is always close to uniform, then a *lazy* random walk on $\mathcal{G}$ will be close to the stationary distribution of G_t for any t large enough. Note that a strong expansion condition on lazy random walks can never be satisfied; luckily, we just need this strong expansion condition to hold for their simple counterpart.

Lemma 3.1. *Let $(G_t)_{t\in\mathbb{N}}$ be a sequence of graphs, and $(P_t)_{t\in\mathbb{N}}$ (resp. $(Q_t)_{t\in\mathbb{N}}$) the corresponding sequence of transition matrices for a* lazy *(resp.* non-lazy*) random walk. Assume there exists $1 < C = O(1)$ such that, for any $t \geq 1$ and any $x \in V$, $1/(C\cdot n) \leq \pi_t(x) \leq C/n$. Moreover, also assume that, for any $t \in \mathbb{N}$, $\max\{|\lambda_2(Q_t)|, |\lambda_n(Q_t)|\} \leq \lambda = o(1)$. Then, there exists an absolute constant C' such that, w.h.p., for any $t \geq C' \log n$ and any starting distribution μ_0,*

$$\left\| \frac{\mu_t}{\pi_t} - \mathbf{1} \right\|_{2,\pi_t}^2 \leq 10C^2(C-1)^2,$$

where $\mu_t = \mu_0 P_1 \cdots P_t$.

We now show how it can be used to derive Theorem 1.2. First recall that since we are assuming $G_0 \sim \mathcal{G}(n,\tilde{p})$, all graphs in the sequence $(G_t)_{t\in\mathbb{N}}$ are sampled (non-independently) from $\mathcal{G}(n,\tilde{p})$ (see Appendix A). Furthermore, for any $t \in \mathbb{N}$, the assumptions of Theorem 1.2 on $\lambda_2(Q_t)$ and $\lambda_n(Q_t)$ are satisfied with probability $1 - o(1/n^2)$ for any graph sampled from $\mathcal{G}(n,\tilde{p})$ with $\tilde{p} > 2\log n/n$ by [13, Theorem 1.1]. Moreover, for $\tilde{p} = \omega(\log n/n)$, by a standard Chernoff bound argument we can show that, with probability $1 - o(1/n^2)$, all vertices of a graph sampled from $\mathcal{G}(n,\tilde{p})$ have degree $(1 + o_n(1))n\tilde{p}$. This implies that, for any t, w.h.p, the stationary distribution of G_t satisfies the assumptions of Lemma 3.1 with $C = 1 + o(1)$, which yields Theorem 1.2.

It is natural to ask if we can relax the condition on p. Assume for example that p, q are such that $\tilde{p} = p/(p+q) > 2\log n$. By [13, Theorem 1.1], the conditions on λ are still satisfied. However, it only holds that $C = \Theta(1)$. Therefore, Lemma 3.1 can only establish that the $\ell_2(\pi_t)$-distance to stationarity is a constant (potentially larger than 1). This, unfortunately, does not give us any meaningful bound on the total variation distance. However, if the ℓ_2-distance between two distributions μ and π is small, $\mu(x)$ cannot be much larger than $\pi(x)$. In a sense, this result can be interpreted as a *coarse* mixing property. This is summarised in the following proposition.

Proposition 3.2. *Let $(G_t)_{t\in\mathbb{N}} \sim \mathcal{G}(n,p,q)$ with $p/(p+q) > 2\log n/n$ and $q = \Omega(1)$. Let π_t be the stationary distribution of G_t. Then, there exists absolute constants $c_1, c_2 > 0$ such that, for any starting distribution μ_0 and any $c_1 \log n \leq t \leq \sqrt{n} + c_1 \log n$, it holds that*

$$\mathbb{P}\left[\left\| \frac{\mu_t}{\pi_t} - 1 \right\|_{2,\pi_t}^2 \leq c_2 \right] \geq 1 - o_n(1).$$

4 Results for the slowly-changing case

4.1 Positive result for mixing in the dense and slowly-changing case

The aim of this section is to prove the following theorem.

Theorem 1.3 (Slowly-changing and dense, fast mixing). *Let $d = \Omega(\log n)$, $p = O(\log n/n^2)$, and $q = O(\log n/(dn))$. Let the following assumptions on the starting graph $G_0 = (V, E_0)$ be satisfied for large enough constants $c_1, c_2, c_3 > 0$.*

(1) $\deg_0(x) = \Theta(d)$ for any $x \in V$;
(2) $|E_0(S, V \setminus S)| \geq c_2 \log n|S|$, for any $S \subset V$ with $|S| \leq c_1 \log n$;
(3) $\Phi_{G_0} \geq c_3 \log d/d$.

Then, $t_{\text{mix}}(\mathcal{G}(n, p, q)) = O(\log n/\Phi^2_{G_0})$.

We start by stating that, if the three assumptions of Theorem 1.3 are satisfied, then, for any $t = O(nd \log n)$, the conductance of G_t is not much worse than the conductance of G_0 (with high probability).

Lemma 4.1 (Conductance lower bound). *Let $d = \Omega(\log n)$, $p = O(\log n/n^2)$, and $q = O(\log n/(dn))$. Assume that G_0 satisfies assumptions (1),(2),(3) of Theorem 1.3. Then, there exists a constant $c > 0$ such that, for any $t = O(nd \log n)$ and any vertex $v \in V$,*

$$\mathbb{P}\left[\deg_t(v) \leq \frac{1}{2}\deg_0(v)\right] = O(n^{-4})$$

and

$$\mathbb{P}\left[\Phi_{G_t} \geq c \cdot \Phi_{G_0}\right] = 1 - O(n^{-4}).$$

The proof of this lemma proceeds as follows: for any $S \subset V$, when an edge is randomly added or removed from the graph, we show that the probability that $|E_t(S, V \setminus S)|$ increases is usually larger than the probability it decreases. Therefore, we model $|E_t(S, V \setminus S)|$ as a random walk on $\mathbb{N}$ with a bias towards large values of $|E_t(S, V \setminus S)|$, i.e., a *birth-and-death* chain. Using standard arguments about birth-and-death chains, we show it is very unlikely that $|E_t(S, V \setminus S)|$ becomes much smaller than $|E_0(S, V \setminus S)|$. By a similar argument we also show that the degrees of all nodes in S are approximately the same as their original degrees in G_0. This ensures that the conductance of a single set S is preserved after $t = O(dn \log n)$ steps. We then use a union bound argument to show that, with high probability, the conductance of the entire graph is preserved. For certain values of d, however, we cannot afford to use a union bound on *all* possible sets of vertices. To overcome this, we show that only applying the union bound for connected sets S would suffice. By bounding the number of such sets with respect to the maximum degree in G_0, we establish the lemma.

We can now give an outline of the proof of Theorem 1.3. The idea is to show that $\left\|\frac{\mu_{t+1}}{\pi_{t+1}} - \mathbf{1}\right\|_{2,\pi_{t+1}}$ is smaller than $\left\|\frac{\mu_t}{\pi_t} - \mathbf{1}\right\|_{2,\pi_t}$ (unless the latter is already

very small). We do this by first relating $\left\|\frac{\mu_t}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}$ with $\left\|\frac{\mu_{t+1}}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}$. More precisely, we can use Lemma 2.1 and Lemma 4.1 to show that the latter is smaller than the former by a multiplicative factor that depends on Φ_{G_0}. Then, we bound the difference between $\left\|\frac{\mu_{t+1}}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}$ and $\left\|\frac{\mu_{t+1}}{\pi_{t+1}}-\mathbf{1}\right\|_{2,\pi_{t+1}}$. In particular, by exploiting the fact that at each step only $O(\log n)$ random edges can be deleted with high probability, we are able to show that $\left\|\frac{\mu_{t+1}}{\pi_{t+1}}-\mathbf{1}\right\|_{2,\pi_{t+1}}$ is not much larger than $\left\|\frac{\mu_{t+1}}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}$. Finally, by putting together all these argument, we show that $\left\|\frac{\mu_t}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}$ is monotonically decreasing in t, at least until the walk is mixed. This establishes the theorem.

Proof of Theorem 1.3. We establish the theorem by showing that, unless $\left\|\frac{\mu_t}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}$ is already small, $\left\|\frac{\mu_t}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}$ will significantly decrease at each step. In particular we relate $\left\|\frac{\mu_t}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}$ to $\left\|\frac{\mu_{t+1}}{\pi_{t+1}}-\mathbf{1}\right\|_{2,\pi_{t+1}}$ in two steps:

(1) We lower bound the change between $\left\|\frac{\mu_t}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}$ and $\left\|\frac{\mu_{t+1}}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}$;
(2) We upper bound the difference between $\left\|\frac{\mu_{t+1}}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}$ and $\left\|\frac{\mu_{t+1}}{\pi_{t+1}}-\mathbf{1}\right\|_{2,\pi_{t+1}}$.

Step 1: The first step follows from a simple spectral argument. Indeed, by Lemma 2.1, we have that

$$\left\|\frac{\mu_{t+1}}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}^2 \leq \lambda_2^2(P_t)\left\|\frac{\mu_t}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}^2,$$

where $\lambda_2(P_t)$ is the second largest eigenvalue of P_t, the transition matrix of G_t.

Step 2: We now upper bound the expected difference between $\left\|\frac{\mu_{t+1}}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}$ and $\left\|\frac{\mu_{t+1}}{\pi_{t+1}}-\mathbf{1}\right\|_{2,\pi_{t+1}}$. In the following analysis we condition on the event that at any time t, $|E_t| \in [(1-o(1))nd, (1+o(1))nd]$ where $d=(n-1)\tilde{p}$. This event happens with probability $1-o(1)$ by Lemma 4.1. Recall that

$$\left\|\frac{\mu_t}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}^2 = \sum_{y\in V}\pi_t(y)\left(\frac{\mu(y)}{\pi_t(y)}-1\right)^2 = \left(\sum_{y\in V}\frac{\mu_t^2(y)}{\pi_t(y)}\right)-1.$$

Hence, we have that

$$\begin{aligned}
&\mathbb{E}\left[\left\|\frac{\mu_{t+1}}{\pi_{t+1}}-\mathbf{1}\right\|_{2,\pi_{t+1}}^2 - \left\|\frac{\mu_{t+1}}{\pi_t}-\mathbf{1}\right\|_{2,\pi_t}^2\right] \\
&= \sum_{y\in V}\mathbb{E}\left[\mu_{t+1}^2(y)\left(\frac{1}{\pi_{t+1}(y)}-\frac{1}{\pi_t(y)}\right)\right] \\
&= \sum_{y\in V}\mathbb{E}\left[\mu_{t+1}^2(y)\left(\frac{2|E_{t+1}|}{\deg_{t+1}(y)}-\frac{2|E_t|}{\deg_t(y)}\right)\right]
\end{aligned}$$

$$\leq 2(1+o(1))|E| \sum_{y\in V} \mu_{t+1}^2(y) \mathbb{E}\left[\left(\frac{1}{\deg_{t+1}(y)} - \frac{1}{\deg_t(y)}\right)\right] \tag{4.1}$$

$$\leq 2(1+o(1))|E| \sum_{y\in V} \mu_{t+1}^2(y) \frac{(1-\frac{1}{2})\deg_t(y)}{\frac{1}{2}\deg_t(y)\cdot \deg_t(y)} (1-(1-q)^{\deg_t(y)}) \tag{4.2}$$

$$\leq \frac{2(1+o(1))}{1-o(1)} \sum_{y\in V} \cdot \frac{\mu_{t+1}^2(y)}{\deg_t(y)/((1-o(1))|E|)} (1-(1-q)^{\deg_t(y)})$$

$$\leq \frac{2(1+o(1))}{1-o(1)} \cdot (1-(1-q)^{\deg_t(y)}) \sum_{y\in V} \frac{\mu_{t+1}^2(y)}{\pi_t(y)}$$

$$\leq O\left(\frac{\log n}{n}\right)\left(\left\|\frac{\mu_{t+1}}{\pi_t} - \mathbf{1}\right\|_{2,\pi_t}^2 + 1\right) \tag{4.3}$$

where $|E| = nd$ and $d = (n-1)\tilde{p}$. From line (4.1) to line (4.2) we upper bound the expectation by only considering the cases where the difference is positive, i.e., $\deg_t(y) \geq \deg_{t+1}(y)$. In line (4.2), by Lemma 4.1 we know $\deg_{t+1}(y)$ will not be smaller than $\frac{1}{2}\cdot\deg_t(y)$ with probability $1-O(n^{-4})$. Moreover, the probability $1-(1-q)^{\deg_t(y)}$ is the probability that at least one of the edges connected to y at time t changes at $t+1$. In line (4.3), we hide unimportant constants in the O-notation and we use the inequality $(1-q)^{\deg_t(y)} \geq 1 - q\cdot\deg_t(y)$. Since $q = O(\log n/(dn))$ by assumption, we get $O(\log n/n)$ in line (4.3).

By combining the two steps above we have

$$\begin{aligned}
&\mathbb{E}\left[\left\|\frac{\mu_{t+1}}{\pi_{t+1}} - \mathbf{1}\right\|_{2,\pi_{t+1}}^2 - \left\|\frac{\mu_t}{\pi_t} - \mathbf{1}\right\|_{2,\pi_t}^2\right] \\
&\leq O\left(\frac{\log n}{n}\right)\left(\left\|\frac{\mu_{t+1}}{\pi_t} - \mathbf{1}\right\|_{2,\pi_t}^2 + 1\right) - (1-\lambda_2^2(P_t))\left\|\frac{\mu_t}{\pi_t} - \mathbf{1}\right\|_{2,\pi_t}^2 \\
&\leq O\left(\frac{\log n}{n}\right)\left(\lambda_2^2(P_t)\left\|\frac{\mu_t}{\pi_t} - \mathbf{1}\right\|_{2,\pi_t}^2 + 1\right) - (1-\lambda_2^2(P_t))\left\|\frac{\mu_t}{\pi_t} - \mathbf{1}\right\|_{2,\pi_t}^2 \\
&\leq \left(\frac{n+\log n}{n}\cdot\lambda_2^2(P_t) - 1\right)\left\|\frac{\mu_t}{\pi_t} - \mathbf{1}\right\|_{2,\pi_t}^2 + O\left(\frac{\log n}{n}\right)
\end{aligned}$$

Therefore, it holds that

$$\mathbb{E}\left[\left\|\frac{\mu_{t+1}}{\pi_{t+1}} - \mathbf{1}\right\|_{2,\pi_{t+1}}^2\right] \leq \left(\frac{n+\log n}{n}\right)\lambda_2^2(P_t)\cdot\left\|\frac{\mu_t}{\pi_t} - \mathbf{1}\right\|_{2,\pi_t}^2 + O\left(\frac{\log n}{n}\right).$$

By Theorem 2.3 and the laziness of the walk,

$$\frac{\Phi_{G_t}^2}{2} \leq 1 - \lambda_2(P_t) \leq 2\Phi_{G_t}.$$

Since we assume the conductance is lower bounded by $O(\log d/d)$, we have $\lambda_2(P_t) \leq 1 - O(\log^2 d/d^2)$ and hence $((n+\log n)/n)\lambda_2^2(P_t) \leq 1$. Therefore,

in expectation, the ℓ_2 distance shrinks by a constant factor (unless it's already small in the first place). Therefore, by standard arguments, after $O(\log n/\Phi^2_{G_0})$ rounds the expected distance to π_t is at most $O(\sqrt{\log n/n})$. By Lemma 4.1, we know this holds for $poly(n)$ time steps. Finally, it suffices to apply Markov's inequality and a union bound to show the expected distance is small with probability $1-O(n^{-4})$ on a polynomially long time interval as required by Definition 2.7. □

4.2 Negative result for mixing in the sparse and slowly changing case

Proposition 1.4 (Slowly-changing and sparse, slow mixing). *Let $p = O(1/n^2)$ and $q = \omega(1/(n\log n))$. Consider a random walk on $\mathcal{G}(n,p,q)$ with starting graph $G_0 \sim \mathcal{G}(n,\tilde{p})$ with $\tilde{p} = p/(p+q)$. Then, $t_{\text{mix}}(\mathcal{G}(n,p,q)) = \Omega(n)$.*

Proof. Consider the graph $G_0 \sim \mathcal{G}(n,\tilde{p})$. Notice that $\tilde{p} = o(\log n/n)$ is well below the connectivity threshold of Erdős-Rényi random graphs. Therefore, with high probability, there is at least one isolated vertex in G_0; call this vertex u and assume the random walk starts from that vertex. The probability that u remains isolated in the steps $1, 2, \ldots, t$ is at least

$$(1-p)^{(n-1)\cdot t} \geq (1-O(1/n^2))^{(n-1)\cdot t} \geq 1 - O(t/n).$$

Therefore, with at least constant nonzero probability, there exists a constant $c > 0$ such that, for any $t \leq c \cdot n$, $\mu^u_t(u) = 1$. Since $\pi_t(u) = 0$, this implies that $\|\mu^u_t - \pi^t\|_{TV} = 1$. □

Actually the proof reveals a stronger "non-mixing" property; if the random walk starts from a vertex that is isolated in G_0, then this vertex will remain isolated for $\Theta(1/(np))$ rounds in expectation, and in this case the random walk did not move at all!

5 Conclusion

In this work we investigated the mixing time of random walks on the edge-Markovian random graph model. Our results cover a wide range of different densities and speeds by which the graph changes. On a high level, these findings provide some evidence to the intuition that both "high density" and "slow change" correlate with fast mixing.

For further work, one interesting setting that is not fully understood is the semi-sparse ($d = \Theta(\log n)$) and fast-changing ($q = \Omega(1) > 0$) case. While we proved that the random walk achieves some coarse mixing in $O(\log n)$, we conjecture that strong mixing is not possible. Another possible direction for future work is, given the bounds on the mixing time at hand, to derive tight bounds on the cover time. Finally, it would be also interesting to study the mixing time in a dynamic random graph model where not all edge slots are present (similar to the models studied in [12, 15], where the graph at each step is a random subgraph of a fixed, possibly sparse, network).

A Mixing times for the graph chain of edge-Markovian models

It is well known that the edge-Markovian graph model $\mathcal{G}(n,p,q)$ converges to an Erdős-Rényi model $\mathcal{G}(n,\tilde{p})$ where $\tilde{p}=\frac{p}{p+q}$, which is the stationary distribution of the original edge-Markovian model. The mixing time of the graph chain has not been proven formally in previous works. Hence, we provide a proof for the sake of completeness. We remark that since an edge-Markovian model is a time-homogeneous (i.e., static) Markov chain, the classical definition of mixing time (Definition 2.6) applies.

Theorem A.1 (Graph chain mixing time). *For an edge-Markovian model $\mathcal{G}(n,p,q)$, the graph distribution converges to the graph distribution of the random graph model $\mathcal{G}(n,\tilde{p})$ where $\tilde{p}=\frac{p}{p+q}$. For any $\epsilon\in(0,1)$, the mixing time of the graph chain $\mathcal{G}(n,p,q)$ is $t_{\text{mix}}(\epsilon)=O\left(\frac{\log(n/\epsilon)}{\log(1/|1-p-q|)}\right)$ for $p+q\neq 1$, and $t_{\text{mix}}(\epsilon)=1$ if $p+q=1$.*

Proof. Every edge slot can be represented by a two-state (close/open) Markov chain with transition matrix

$$P=\begin{pmatrix}1-p & p\\ q & 1-q\end{pmatrix}$$

and stationary distribution $\left(\frac{q}{p+q},\frac{p}{p+q}\right)$. By using standard Markov chain arguments (see, e.g., [16, Chapter 1]), the distance to the stationary distribution shrinks at each step by a factor of $|1-p-q|$, i.e.,

$$\|\mu_{t+1}-\pi\|_{TV}\leq|1-p-q|\,\|\mu_t-\pi\|_{TV}\,.$$

Therefore, when $p+q\neq 1$, the mixing time $t_{\text{mix}}(\epsilon)$ of this two-state Markov chain is $O\left(\frac{\log(1/\epsilon)}{\log(|1-p-q|)}\right)$ where $\epsilon<1$. For all the $\binom{n}{2}$ edge slots, the time that all of them mix is $O\left(\frac{\log\binom{n}{2}+\log(1/\epsilon)}{\log(|1-p-q|)}\right)$. When $p+q=1$, instead, the graph mixes immediately, which confirms the fact that in this regime the graph model is equivalent to a sequence of independent graphs from $\mathcal{G}(n,\tilde{p})$. □

Remark A.1. Theorem A.1 essentially tells us that, whenever at least one between p and q is large (e.g., $\Omega(1)$), the graph chain quickly converges to $\mathcal{G}(n,\tilde{p})$ with $\tilde{p}=\frac{p}{p+q}$. This suggests that for a fast-changing edge-Markovian model $\mathcal{G}(n,p,q)$ with $q=\Omega(1)$, we can consider w.l.o.g. the starting graph G_0 as sampled from $\mathcal{G}(n,\tilde{p})$.

References

1. Alon, N., Milman, V.D.: λ_1, isoperimetric inequalities for graphs, and superconcentrators. J. Combin. Theory Ser. B **38**(1), 73–88 (1985)
2. Augustine, J., Pandurangan, G., Robinson, P.: Distributed algorithmic foundations of dynamic networks. SIGACT News **47**(1), 69–98 (2016)
3. Avin, C., Koucký, M., Lotker, Z.: Cover time and mixing time of random walks on dynamic graphs. Random Structures & Algorithms **52**(4), 576–596 (2018)
4. Berenbrink, P., Giakkoupis, G., Kermarrec, A., Mallmann-Trenn, F.: Bounds on the voter model in dynamic networks. In: 43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016. LIPIcs, vol. 55, pp. 146:1–146:15 (2016)
5. Clementi, A., Crescenzi, P., Doerr, C., Fraigniaud, P., Pasquale, F., Silvestri, R.: Rumor spreading in random evolving graphs. Random Structures Algorithms **48**(2), 290–312 (2016)
6. Clementi, A., Monti, A., Pasquale, F., Silvestri, R.: Information spreading in stationary Markovian evolving graphs. IEEE Transactions on Parallel and Distributed Systems **22**(9), 1425–1432 (2011)
7. Clementi, A.E.F., Macci, C., Monti, A., Pasquale, F., Silvestri, R.: Flooding time of edge-markovian evolving graphs. SIAM J. Discrete Math. **24**(4), 1694–1712 (2010)
8. Clementi, A.E.F., Silvestri, R., Trevisan, L.: Information spreading in dynamic graphs. Distributed Computing **28**(1), 55–73 (2015)
9. Cooper, C.: Random walks, interacting particles, dynamic networks: Randomness can be helpful. In: Kosowski, A., Yamashita, M. (eds.) Structural Information and Communication Complexity - 18th International Colloquium, SIROCCO 2011. vol. 6796, pp. 1–14. Springer (2011)
10. Denysyuk, O., Rodrigues, L.E.T.: Random walks on evolving graphs with recurring topologies. In: Distributed Computing - 28th International Symposium, DISC 2014, Austin, TX, USA, October 12-15, 2014. Proceedings. pp. 333–345 (2014)
11. Giakkoupis, G., Sauerwald, T., Stauffer, A.: Randomized rumor spreading in dynamic graphs. In: Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014. Lecture Notes in Computer Science, vol. 8573, pp. 495–507 (2014)
12. Hermon, J., Sousi, P.: Random walk on dynamical percolation. arXiv preprint arXiv:1902.02770 (2019)
13. Hoffman, C., Kahle, M., Paquette, E.: Spectral Gaps of Random Graphs and Applications. International Mathematics Research Notices (05 2019)
14. Kuhn, F., Oshman, R.: Dynamic networks: models and algorithms. SIGACT News **42**(1), 82–96 (2011)
15. Lamprou, I., Martin, R., Spirakis, P.: Cover time in edge-uniform stochastically-evolving graphs. Algorithms **11**(10), Paper No. 149, 15 (2018)
16. Levin, D.A., Peres, Y.: Markov chains and mixing times. American Mathematical Society, Providence, RI (2017)
17. Michail, O., Spirakis, P.G.: Elements of the theory of dynamic networks. Commun. ACM **61**(2), 72 (2018)
18. Montenegro, R., Tetali, P.: Mathematical aspects of mixing times in Markov chains. Found. Trends Theor. Comput. Sci. **1**(3), x+121 (2006)
19. Peres, Y., Sousi, P., Steif, J.: Mixing time for random walk on supercritical dynamical percolation. Prob. Theory and Related Fields (2019), (to appear)

20. Peres, Y., Stauffer, A., Steif, J.E.: Random walks on dynamical percolation: mixing times, mean squared displacement and hitting times. Probab. Theory Related Fields **162**(3-4), 487–530 (2015)
21. Saloff-Coste, L., Zúñiga, J.: Merging for time inhomogeneous finite Markov chains. I. Singular values and stability. Electron. J. Probab. **14**, 1456–1494 (2009)
22. Saloff-Coste, L., Zúñiga, J.: Merging for inhomogeneous finite Markov chains, Part II: Nash and log-Sobolev inequalities. Ann. Probab. **39**(3), 1161–1203 (2011)
23. Saloff-Coste, L., Zúñiga, J.: Merging and stability for time inhomogeneous finite Markov chains. In: Surveys in stochastic processes, pp. 127–151. EMS Ser. Congr. Rep., Eur. Math. Soc., Zürich (2011)
24. Sarma, A.D., Molla, A.R., Pandurangan, G.: Distributed computation in dynamic networks via random walks. Theor. Comput. Sci. **581**, 45–66 (2015)
25. Sauerwald, T., Zanetti, L.: Random walks on dynamic graphs: Mixing times, hitting times, and return probabilities. In: 46th International Colloquium on Automata, Languages, and Programming, ICALP 2019. LIPIcs, vol. 132, pp. 93:1–93:15 (2019)
26. Sousi, P., Thomas, S.: Cutoff for random walk on dynamical Erdos-Renyi graph. arXiv preprint arXiv:1807.04719 (2018)