Rapid Mixing of Dynamic Graphs with Local Evolution Rules

Laurent Massoulié, Rémi Varloot

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TNSE.2019.2926887

Rapid Mixing of Dynamic Graphs with Local Evolution Rules

Laurent Massoulié, *MSR-Inria Joint Center*, and Rémi Varloot, *Nokia Bell Labs France*

Abstract—Dynamic graphs arise naturally in many contexts. In peer-to-peer networks, for instance, a participating peer may replace an existing connection with one neighbor by a new connection with a neighbor of that neighbor. Several such local rewiring rules have been proposed to ensure that peer-to-peer networks achieve good connectivity properties (e.g. high expansion) at equilibrium. However, the question of whether there exists such a rule that converges rapidly to equilibrium has remained open. In this work, we provide an affirmative answer: we exhibit a local rewiring rule that converges to equilibrium after each participating node has undergone only a number of changes that is at most poly-logarithmic in the system size. As a byproduct, we derive new results for random walks on graphs, bounding the spread of their law throughout the transient phase, i.e. prior to mixing. These rely on an extension of Cheeger's inequality, based on generalized isoperimetric constants, and may be of independent interest.

Index Terms—Markov chains, Mixing time, Dynamic graphs, Local dynamics, Bottleneck ratio

1 Introduction

Distributed systems typically cannot operate efficiently unless their constituting parts are interconnected via a network with suitable properties. In the context of peer-to-peer systems, desirable properties of the interconnection graph between peers include having a small diameter, small node degrees, and requiring many failures to disconnect a sizeable part of the network. Yet another useful property is the ability to obtain at low cost and via a distributed algorithm uniform samples of nodes in the system.

It turns out that all such properties follow if the interconnection graph is an expander (see [1] for a general reference). By definition, the unoriented graph $G = (V, E)$ is a γ-expander for some constant $\gamma > 0$ if each set $S \subset V$ of vertices with size $|S|$ no larger than $|V|/2$ is such that at least $\gamma|S|$ distinct edges in E connect S to its complement $\overline{S}$ in V. It is an expander if it is a γ-expander for some $\gamma \geq \Omega(1)$.

For instance, when the graph is an expander, uniform samples are obtained in a distributed manner and at low cost by running a random walk on the peer-to-peer graph: It being an expander then ensures that only few steps (on the order of the logarithm of the system size, measured in number of nodes) of the walk suffice.

Because peer-to-peer systems are volatile, i.e. subject to node arrivals and departures, it is not possible to determine once and for all an expander graph to interconnect participating peers. Instead the graph must constantly evolve, with the aim to preserve or restore the desired expander property. Moreover, the graph evolution must rely on local adjustments, since by design no central controller has knowledge of the whole graph.

This has prompted research on dynamics for continuous modification of graphs that would:

1) Rely only on local modifications of the current graph;
2) Require minimal computation and storage capabilities per node;
3) Produce expanders at equilibrium;
4) Quickly reach equilibrium.

More precisely, considering graphs on a set of N nodes, by quickly we mean requiring a number of modifications per node that scales poly-logarithmically in N before equilibrium is attained.

The main contribution of this paper is to propose a new graph dynamics together with the proof that it meets these four requirements. Its organization is as follows. Section 2 describes our proposed dynamics together with the main result, Theorem 1. It also highlights Theorem 2, our technical result controlling the spread of laws of random walks on graphs at short times. Section 3 explains the proof strategy. Section 4 explains how to deduce bounds with high probability on isoperimetric ratios from bounds on corresponding expectations, using negative dependence properties. Section 5 derives the necessary bounds on expectations, leveraging in particular Theorem 2. The proof of the latter constitutes Section 6. A global outline of the proof is given in Figure 3.

We now review relevant prior work.

Related Work

Markovian local graph dynamics for peer-to-peer systems have been considered in [2], [3], [4], [5]. In all these papers the stationary regime for the proposed dynamics has been identified; in the last three references, loose bounds on the mixing time (defined below), or time to achieve equilibrium, have been obtained. Similarly, in [6], the authors consider the SKIP+ local graph dynamic, specifically tuned to rapidly update the skip graph data structure.

The best known bound on the time before local graph dynamics produce an expander graph with high probability is $O(\ln^2 N)$ for the SKIP+ dynamic [6]. The context differs from our approach, however, as this dynamic is specifically tailored to construct a predefined topology, rather than converge to a given equilibrium. Furthermore, it does not fully satisfy our 2nd requirement in that it requires that each node store various state variables (including its own state, that of its neighbors, whether each outgoing edge is stable or temporary, etc.), with a complex update scheme based on these variables.

Among the remaining papers, the tightest bound available prior to the present article was obtained in [7]. Specifically, it is shown in [7] that a discrete time Markov chain on the set of connected d-regular graphs on N vertices creates with high probability an expander graph after a time of $O(N^2 d^2 \sqrt{\ln N})$ with d of order $O(\ln N)$. This implies the realization of this property after each node has performed a number of updates of order $O(N \ln^{5/2}(N))$, i.e. a number that is quasi-linear in the system size N.

Graph dynamics have also been considered in different contexts. [8] considers local dynamics for producing so-called cladograms uniformly at random, and bounds their mixing time. [9] considers dynamics of matchings in bipartite graphs and controls their mixing time. The motivation of [9] is the estimation of graph descriptors using a Markov chain Monte-Carlo approach. Finally, non-local graph dynamics together with their mixing time have been considered in [10] in order to sample from so-called exponential random graph distributions.

2 Main results

In the sequel, we consider graphs over vertex set $[N] = \{1, \dots, N\}$, where N is a positive integer. Asymptotic results will be with respect to N. We write $\text{polylog}(N)$ to represent $O(\ln^c N)$ for some positive constant c.

Consider the following setting. The vertices in $[N]$ are connected by edges of three distinct types: a fixed cycle, blue edges and red edges. The cycle is constituted of a fixed set of edges $E_\circ = \{(n, n+1) : i \in [N]\}$, with $N+1 \equiv 1$. Each node $n \in [N]$ furthermore maintains two pointers, one blue and one red, with respective destinations b_n, r_n in $[N]$. The destinations of the pointers are such that (b_n) and (r_n) are permutations: each node n is the destination of exactly one blue pointer and one red pointer. The blue edges and red edge sets are respectively $E_{\mathbf{b}} = \{(n, b_n) : n \in [N]\}$ and $E_{\mathbf{r}} = \{(n, r_n) : n \in [N]\}$. All edges are considered to be unoriented.

We now consider the following continuous-time dynamics. The graph evolves through alternating blue and red phases. During each phase, only the edges of a given color evolve, while those of the other color are kept fixed. During a blue phase, for example, blue pointers are swapped along graph $G_{\mathbf{r}}$ constituted of the edges in both $E_\circ$ and $E_{\mathbf{r}}$. Note that $G_{\mathbf{r}}$ is a 4-regular multigraph.

The dynamics for a blue phase are defined as follows. Each edge $e = (i, j)$ of $G_{\mathbf{r}}$ maintains an internal clock, in which the time between ticks are exponentially distributed with mean 1 and all independent. At every tick, the two nodes $n, m \in [N]$ such that $b_n = i$ and $b_m = j$ swap their pointers. This effectively boils down to transposing i and j in the permutation $(b_n)_{n \in [N]}$. Such a process has been studied in the literature, where it is known as the *interchange process*. See for instance Jonasson [11] or N. Berestycki [12], where the discrete time version of this process is analyzed.

For the red phases, the roles of blue and red pointers are swapped; the graph containing the edges in $E_\circ$ and $E_{\mathbf{b}}$ is denoted $G_{\mathbf{b}}$. To clarify notations, we denote E_f the fixed edge set during phase f, i.e. $E_{\mathbf{b}}$ if f is a red phase, $E_{\mathbf{r}}$ otherwise, and G_f the fixed graph, i.e. containing the edges in $E_\circ$ and E_f. We finally write $E_f^\circ = E_f \cup E_\circ$.

Our main result is then as follows

Theorem 1. *Let $T = \ln^a N$ where $a > 8$ is a constant. After $F = \lceil \log_2 N \rceil$ phases of length T, the variation distance between the joint distribution of the sets of blue and red pointers and that of two independent permutations uniformly distributed over the symmetric group S_N is $o(1)$.*

Corollary 1. *After each node has undergone a number of local connectivity modifications that is polylogarithmic in N, the process has produced an expander with high probability.*

Proof. By time $\tau := FT$, a given node $n \in [N]$ has seen under these dynamics a number of connectivity modifications M_n that is at most a Poisson random variable with mean 8τ. Indeed, since G_f has degree 4, the rate at which the inbound and outbound pointers at n move are both 4, for a total transition rate of 8.

The probability that M_n exceeds 16τ is then, by Chernoff's bound for deviations of Poisson random variables from their mean, bounded by

$$\mathbb{P}(M_n \geq 16\tau) \leq e^{-8\tau h(16\tau/(8\tau))} = e^{-8\tau h(2)},$$

where $h(x) := x \ln(x) - x + 1$ is the Cramér transform of a unit mean Poisson random variable. Since τ is at least of order $\ln^{a+1} N$ with $a > 0$, the last term is $o(1/N)$. Thus the probability that at least one node $n \in [N]$ undergoes more than $16\tau = \text{polylog}(N)$ local modifications by time τ is, by the union bound, no more than $N o(1/N) = o(1)$.

The fact that the resulting graph G_F is an expander will be shown in Section 4, in which we introduce the necessary technical lemmas. Note that results in [14] establish for very similar (although not identical) random graph models that these are expanders with high probability. □

We now state another result, which will be instrumental in Section 5, and proven in Section 6, but which we believe could be of independent interest. For this, recall that the Laplacian matrix L of a multi-graph G with adjacency matrix A is by definition $L = D - A$, where D is the diagonal matrix $\text{Diag}(\{d_i\})$, d_i being the degree of node i (see [15] for background on graph Laplacians).

Theorem 2. *Let $G = ([N], E)$ be an undirected multi-graph with maximum degree Δ, and $\{X_t\}$ the continuous time random walk on G, i.e. the Markov jump process on $[N]$ with jump rates q_{ij} equal to the multiplicity of (i, j) in E. The infinitesimal generator of $\{X_t\}$ is $-L$, where L is the Laplacian matrix of G. Let $\{\pi_i(t)\}_{i \in [N]}$ denote its law at time t. For an arbitrary initial distribution of the random walk, for any $k \leq N/2$ and $S \subset [N]$ such that $|S| \leq k$ and any $t \geq 0$, one has:*

$$\sum_{i \in S} \pi_i(t) \leq \frac{|S|}{k+1} + \sqrt{k+1} e^{-\lambda_2^* t}, \tag{1}$$

where

$$\lambda_2^* = \frac{\phi_k(G)^2}{2\Delta}, \tag{2}$$

and $\phi_k(G)$ is defined in (3).

Remark 1. The quantity λ_2^* is of the same form as the lower bound on the spectral gap λ_2 of the Laplacian that the

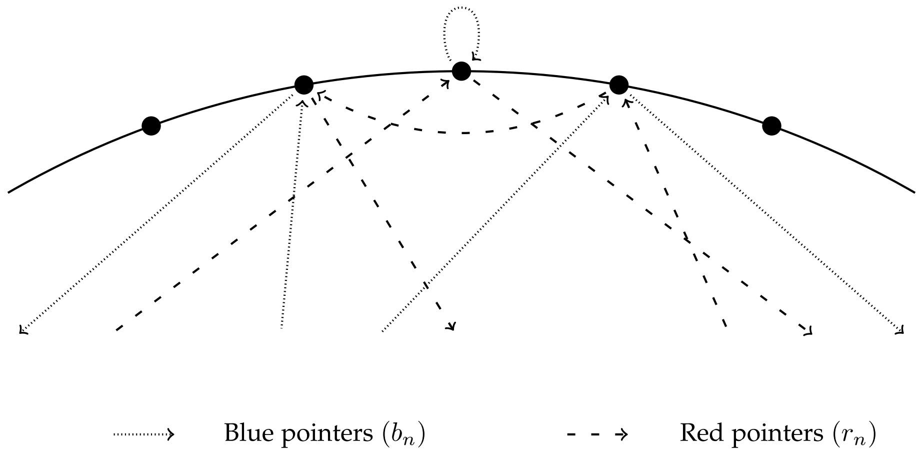


Fig. 1. Bi-color pointer interchange model

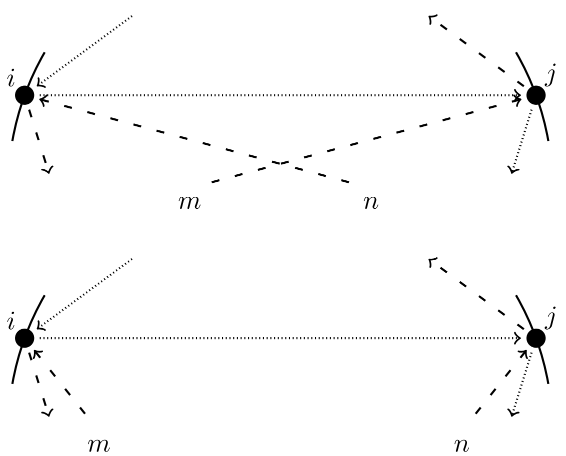


Fig. 2. Swapping two red pointers along a blue edge (red phase)

Cheeger inequality gives when $k = N/2$. In fact for $k = N/2$, the expression (3) of $\phi_k(G)$ coincides with this lower bound. In this classical situation, instead of (1) one has the conclusion that $d_{var}(\pi(t), \mathcal{U}([N])) \leq \sqrt{N}e^{-\lambda_2^* t}$, see e.g. [16].

3 Proof Strategy

To proceed, we first introduce some definitions.

Definition 1. For each $k \in [N/2]$, the k-th isoperimetric constant $\phi_k(G)$ of a graph G with vertex set $V(G) = [N]$ and edge set E is defined as

$$\phi_k(G) := \min_{S \subset [N], |S| \leq k} \frac{|E(S, \overline{S})|}{|S|}, \tag{3}$$

where $\overline{S}$ denotes the complement $[N] \setminus S$ of a set S, $E(S, \overline{S})$ denotes the set of edges in G between S and its complement, and $|\cdot|$ denotes the cardinality of a set.

Note that the particular value $\phi_{N/2}(G)$ is often referred to as the Cheeger constant, and plays an important role in Lemma 3 below.

Definition 2. The collection $\{\phi_k(G)\}_{k\in[N/2]}$ of isoperimetric constants of graph G constitutes its *isoperimetric profile*.

The graph is said to be a (γ, c)-*expander* if, for all $k \leq N/2$, $\phi_k(G) \geq \min(\gamma, c/k)$.

Note that a graph is a γ-expander according to the classical notion if it is a $(\gamma, N/2)$-expander according to the above definition.

Other generalizations of the isoperimetric constant exist, including average conductance [17] and the higher order Cheeger inequalities [18].

Our proof consists of controlling the evolution of the isoperimetric profile of the graph along which pointers move from one phase to the next, establishing lower bounds on this profile in an iterative manner.

Let β be a constant such that $1 < \beta < (a-4)/4$. Such β exists by our assumption that $a > 8$. Let $\gamma = \ln^{-\beta} N$. We show the following

Lemma 1. *If, for a given phase f, the fixed graph G_f is a (γ, c)-expander for some integer c, then with probability at least $1 - o(1/N)$, G_{f+1} (th fixed graph in the following phase) is a $(\gamma, 2c)$-expander.*

Notice that if G_f contained blue edges, G_{f+1} contains red edges, and vice versa.

To prove this, we first show a stronger lower bound on the average number of pointers leaving any given set S:

Lemma 2. *If, for a given phase f and integer c, G_f is a (γ, c)-expander, then for all $S \subset [N]$ with $|S| \leq N/2$,*

$$\mathbb{E}\left|E_{f+1}(S, \overline{S})\right| \geq \frac{1}{2\gamma} \min(\gamma|S|, 2c). \tag{4}$$

Lemma 1 is then deduced from Lemma 2 by invoking some concentration inequalities together with union bounds. Details are given in Section 4.2.

An easy consequence of Lemma 1 is the following:

Corollary 2. *After $F = \lceil \log_2 N \rceil$ phases, with high probability both $G_{\mathbf{b}}$ and $G_{\mathbf{r}}$ are $(\gamma, N/2)$-expanders.*

Proof. Clearly, G_1 is a $(\gamma, 2)$-expander. Indeed, any subset $S \subset [N]$ of size $|S| \leq N/2$ is connected by at least two edges (that come from the cycle) to its complement $\overline{S}$, so that

$$|E_1^\circ(S, \overline{S})| \geq 2 \geq \min(\gamma|S|, 2).$$

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TNSE.2019.2926887

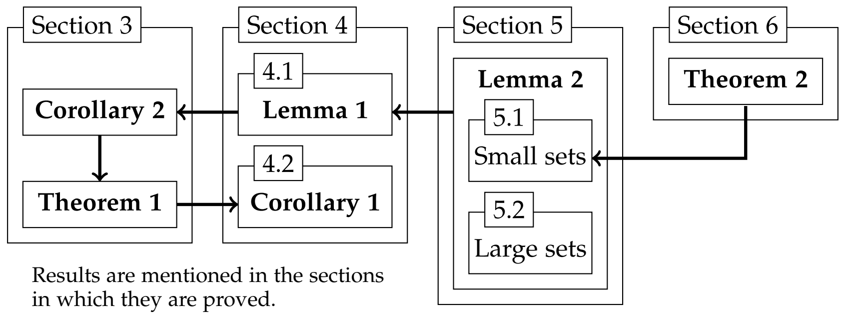


Fig. 3. Outline of proof

Denote by $\mathcal{E}_f$ the event that G_f is a $(\gamma, 2^f)$-expander. Thus we have just established that event $\mathcal{E}_1$ holds with certainty, and Lemma 1 entails that, for all $f \geq 1$,

$$\mathbb{P}(\overline{\mathcal{E}}_{f+1} | \mathcal{E}_f) \leq o(1/N).$$

Thus

$$\begin{aligned}\mathbb{P}(\overline{\mathcal{E}}_{f+1}) &= \mathbb{P}(\overline{\mathcal{E}}_{f+1} | \mathcal{E}_f)\mathbb{P}(\mathcal{E}_f) + \mathbb{P}(\overline{\mathcal{E}}_{f+1} | \overline{\mathcal{E}_f})\mathbb{P}(\overline{\mathcal{E}_f}) \\ &\leq o(1/N) + \mathbb{P}(\overline{\mathcal{E}}_f).\end{aligned}$$

By induction on f, this yields

$$\mathbb{P}(\overline{\mathcal{E}}_{f+1}) \leq o(f/N).$$

For $F = \lceil \log_2 N \rceil$, the right-hand side of this expression is $o(1)$, so that with high probability, the graphs G_{F-1} and G_F are $(\gamma, N/2)$-expanders. As the color of these two graphs differ, this concludes the proof. □

The proof of Theorem 1 is then concluded as follows:

Proof. By Corollary 2, after $F = \lceil \log_2 N \rceil$ phases, the Cheeger constants $\phi_{N/2}(G_{\mathbf{b}})$ and $\phi_{N/2}(G_{\mathbf{r}})$ are at least γ. To, proceed, recall the following definition (see e.g. Levin et al. [16] or [12]).

Definition 3. For a Markov process on some discrete state $\mathcal{S}$, denoting $\pi_s(t)$ its distribution at time t conditional on the initial state being $s \in \mathcal{S}$, and $\pi(\infty)$ its stationary distribution, its mixing time T_{mix} is defined as

$$T_{mix} = \inf\{t > 0 : \sup_{s \in \mathcal{S}} ||\pi_s(t) - \pi(\infty)||_{TV} \leq 1/4\},$$

where $|| \cdot ||_{TV}$ denotes total variation distance.

We show that, for any fixed graph G_f, $\phi_{N/2}(G_f) \geq \gamma$ implies that the interchange processes on G_f mixes in less than T steps, so that in two more phases, the total variation distance between the state of our process and its equilibrium distribution will be $o(1)$.

Our main tool to this end is Theorem 4.3, p. 39 in Berestycki [12], which gives a sufficient condition for the discrete time interchange process on a graph G_f to mix in time T. The proof provided in [12] is a direct application of the so-called method of distinguished paths, a classical technique for bounding mixing times, reviewed for instance in [12] and [16].

The continuous time analogous result reads

Theorem 3 (Theorem 4.3, p. 39, [12])**.** *For each pair of nodes $i, j \in [N]$, define a path γ_{ij} on G connecting i and j, and let $\Gamma = \{\gamma_{ij} : i, j \in [N]\}$. Denote Υ the length of the longest path in Γ, and K the supremum over the edges e of G of the number of paths in Γ passing through e.*

The continuous time interchange process on G will have mixed in time T provided

$$T \geq 8 \ln(N) \Upsilon K / N. \tag{5}$$

According to Lemma 3 below, for a d-regular $(\gamma, N/2)$-expander with d constant, we can choose these paths such that $\Upsilon = O(\ln N/\gamma^2)$ and $K = O(N \ln^2 N/\gamma^2)$. Plugged into (5), these evaluations imply that mixing has occurred by time T provided T is large compared to $\ln(N)^4/\gamma^4$, i.e provided $\ln(N)^a = \omega(\ln(N)^{4+4\beta})$. This condition holds since $\beta < (a-4)/4$. □

Lemma 3. *Let G be a d-regular graph with vertex set $[N]$, such that $\phi_{N/2}(G) \geq \gamma$. One can construct a set of paths $\Gamma = \{\gamma_{ij} : i, j \in [N]\}$ such that the γ_{ij} each have length at most $\Upsilon = 2\ln(N)d^2/\gamma^2$, and such that each edge e of G is crossed by at most $18N \ln^2(N)d^2/\gamma^2$ paths.*

Proof. Cheeger's inequality (see e.g. Berestycki [12] Theorem 3.5, p. 30) ensures that the spectral gap for the discrete time random walk on a d-regular graph G with $\phi_{N/2}(G) \geq \gamma$ is at least $\gamma^2/(2d^2)$. Thus the total variation distance between the distribution of the random walk at time $\Upsilon := 2d^2 \ln(N)/\gamma^2$ and the uniform distribution on G is $o(1/N)$ (this follows e.g. by Theorem 2.2, p. 18 in [12]). As a result, for any $i \in [N]$, the probability that the walk started at i hits j at time Υ is at least $1/(2N)$. Consider then the following randomized construction. For each i, create $5N \ln(N)$ independent walks of length Υ started at i. The probability that for some particular $j \in [N]$, no such walk issued from i hits j is then at most

$$(1 - 1/2N)^{5N \ln(N)} \leq e^{-5\ln(N)/2} = o(N^{-2}).$$

Using the union bound, we can therefore conclude that, with high probability, the collection of paths thus created joins every node i to every node j.

Let us now evaluate the number of times a given edge $e = (u, v)$ of G is traversed by this collection of paths. This is no larger than the number of times these paths visit node u. For $t \leq 5N \ln(N)$, denote by $X_i(t)$ the number of visits to u by the t-th path sampled with starting point i. Clearly, $X_i(t) \leq \Upsilon$. Also,

$$\mathbb{E} \sum_{i \in [N]} \sum_{t \leq 5N \ln(N)} X_i(t) = 5N \ln(N) \sum_{i \in [N]} \sum_{\ell=0}^{\Upsilon} P_{iu}^{(\ell)},$$

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TNSE.2019.2926887

where $P_{iu}^{(\ell)}$ denotes the transition probability from i to u in ℓ steps of the walk. However the walk is symmetric, so that $P_{iu}^{(\ell)} = P_{ui}^{(\ell)}$. The above expection thus reads

$$\mathbb{E} \sum_{i\in[N]} \sum_{t\le 5N\ln(N)} X_i(t) = 5N\ln(N)(\Upsilon+1).$$

Let $Z = \sum_{i\in[N]} \sum_{t\le 5N\ln(N)} X_i(t)$ denote the total number of visits to u by all paths. For any $C > 0$, Hoeffding's inequality then gives

$$\mathbb{P}(Z \ge \mathbb{E}(Z) + CN\Upsilon\ln(N)) \le \exp\left(-\frac{C^2N^2\Upsilon^2\ln(N)^2}{\Upsilon^2 5N^2\ln(N)}\right) = e^{-C^2\ln(N)/5}.$$

Taking $C = 3$, the right-hand side is $o(N^{-1})$. Thus, by the union bound, with high probability, no node u is visited more than $9N\ln(N)\Upsilon = 18N\ln^2(N)d^2/\gamma^2$ times by the collection of constructed paths. $\square$

4 From bounds in expectation to bounds with high probability

4.1 Proof of Lemma 1

Assume that G_f is a (γ, c)-expander. By Lemma 2, for each fixed set $S \subset [N]$ with $|S| \le N/2$, we have that

$$\mathbb{E}|E_{f+1}(S,\overline{S})| \ge \frac{1}{2\gamma}\min(\gamma|S|, 2c).$$

Fix $k \le N/2$. We further restrict ourselves to $k > 2/\gamma$, since $|E_\circ(S,\overline{S})| \ge 2$ and therefore one always has that $\phi_k \ge \gamma$ for $k \le 2/\gamma$.

For some set S of size k, let $\ell \in [k]$ be the number of contiguous portions of the cycle it is made of, i.e. the number of connected subgraphs in the graph $(S, E_\circ \cap S^2)$. Clearly $|E_\circ(S,\overline{S})| = 2\ell$, and therefore

$$|E^\circ_{f+1}(S,\overline{S})| = |E_{f+1}(S,\overline{S})| + 2\ell.$$

Recall (see e.g. Dubashi and Ranjan [19], and Borcea et al. [20]) that a set of random variables $(X_i)_{i\in I}$ is said to be negatively associated if for any two functions $f, g : \mathbb{R}^I \to \mathbb{R}$ that are non-decreasing in each of their coordinates, and depend on disjoint sets of variables X_i, the two random variables $f((X_i)_{i\in I})$ and $g((X_i)_{i\in I})$ are negatively correlated, i.e.

$$\mathbb{E}\left[f((X_i)_{i\in I})g((X_i)_{i\in I})\right] \le \mathbb{E}\left[f((X_i)_{i\in I})\right]\mathbb{E}\left[g((X_i)_{i\in I})\right].$$

We will need the following two results.

Lemma 4. *Conditionally on the pointer configuration at the beginning of the considered phase f, the random variable $|E_{f+1}(S,\overline{S})|$ consists of the sum of negatively associated Bernoulli random variables. Consequently, for any $r \in (0,1)$, it holds that*

$$\mathbb{P}\left(|E_{f+1}(S,\overline{S})| \le r(2\gamma)^{-1}\min(\gamma|S|, 2c)\right) \le e^{-(2\gamma)^{-1}\min(\gamma|S|,2c)h(r)}, \quad (6)$$

where $h(r) := r\ln(r) - r + 1$.

Proof. Consider a blue phase and a given set S. Represent the collection of termination points of pointers through the binary variables $\xi_i \in \{0,1\}$, $i \in [N]$ where $\xi_i = 1$ if and only if one pointer issued from S points towards i, i.e. there exists $j \in S$ such that $b_j = i$. Note that

$$|E_{f+1}(S,\overline{S})| = \sum_{i\in\overline{S}} \xi_i(T).$$

We shall show that the variables $\xi_i(T)$, $i \in [N]$ are, conditionally on their initial values $\xi(0) = \{\xi_i(0)\}_{i\in[N]}$, negatively associated. This will imply, by the results of Dubhashi and Ranjan [19] that $|E_{f+1}(S,\overline{S})|$ satisfies, conditionally on $\xi(0)$, the same Chernoff bounds that it would if the $\xi_i(T)$ were mutually independent. In turn, this guarantees, for all $m < M := \mathbb{E}(|E_{f+1}(S,\overline{S})| \mid \xi(0))$, that

$$\mathbb{P}(|E_{f+1}(S,\overline{S})| \le m \mid \xi(0)) \le e^{-Mh(m/M)}, \quad (7)$$

where $h(x) := x\ln(x) - x + 1$. Indeed, the right-hand side in the above inequality is the Chernoff bound of $\mathbb{P}(\text{Poisson}(M) \le m)$, and a standard argument shows that Chernoff bounds of sums of independent Bernoulli random variables are tighter than the Chernoff bound where that sum is replaced by a Poisson random variable with the same mean. This property is in fact a special case of the so-called Bennett inequality, see e.g. [21], Theorem 2.9 p. 35 (alternatively, see [22], proof of Corollary A.1.7, p. 310). The announced result (6) then follows by taking $m = r(2\gamma)^{-1}\min(\gamma|S|, 2c)$ in (7), remarking that the right-hand side of (7) decreases with M, and thus replacing M by its previously established lower bound, i.e. $(2\gamma)^{-1}\min(\gamma|S|, 2c)$.

It thus remains to prove negative association of $\{\xi_i(T)\}_{i\in[N]}$ conditionally on $\xi(0)$. We shall in fact establish that a stronger form of negative correlation is satisfied, namely the *strong Rayleigh property*. For a precise definition, and the fact that strong Rayleigh property implies negative association, we refer the reader to Liggett [23].

To that end, recall that the symmetric exclusion process consists of particles located at nodes of a graph, that each perform independent random walks along the edges of the graph, except that transitions that would lead to multiple particles at the same site are not allowed. The variables $\{\xi_i(t)\}$ then evolve, under the interchange process dynamics, as a symmetric exclusion process where the transition rates of one particle along any edge are all equal to 1.

Conditionally on $\xi(0)$, the variables $\xi_i(0)$ are deterministic, and thus trivially satisfy the strong Rayleigh property (see Borcea et al. [20] for a proof). Proposition 5.1 in [20] establishes that the symmetric exclusion process preserves the strong Rayleigh property in the following sense: the collection of indicator variables $(\xi_i(t))_{i\in[N]}$ satisfies the strong Rayleigh property for all $t \ge 0$ provided that $(\xi_i(0))_{i\in[N]}$ does. This holds in particular for $(\xi_i(T))_{i\in[N]}$, which concludes the proof of the Lemma. $\square$

Lemma 5. *The number of sets $S \subset [N]$ of size k that consist of ℓ contiguous portions of the cycle is at most $N^{2\ell}$.*

It is also upper-bounded by $N\binom{k-1}{\ell-1}\binom{N-k-1}{\ell-1}$.

Proof. We may enumerate such sets S by scanning the cycle $[N]$ starting from 1, and identifying the first time we find a starting point of an interval in S, then the end point of that interval, and so on. Clearly this will produce 2ℓ numbers in $[N]$, which characterize S, hence the upper bound $N^{2\ell}$.

To obtain the other upper bound, note that the number of strictly positive sequences of ℓ integers $x_1, \ldots, x_k$ such

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TNSE.2019.2926887

that $x_1 + \cdots + x_\ell = k$ equals the number of non-negative sequences of such integers such that $x_1 + \cdots + x_\ell = k - \ell$, and this number is well known to equal

$$\binom{k-\ell+\ell-1}{\ell-1}.$$

Similarly, the number of strictly positive sequences of integers $y_1, \dots, y_\ell$ such that $y_1 + \cdots + y_\ell = N - k$ equals $\binom{N-k-1}{\ell-1}$. Given a set $S \subset [N]$ of size k and made of ℓ distinct intervals, let $z \in \{0, \dots, N-1\}$ be the smallest number of clockwise rotations of the set such that $i = 1$ corresponds exactly to the beginning (in clockwise order) of a contiguous segment of the set. The set is then fully specified by the lengths of its constituting segments, $x_1, \dots, x_\ell$, in clockwise order, together with the lengths of the segments separating its own segments, $y_1, \dots, y_\ell$. Since this construction forms an injection from the collection of considered sets S to a set of size $N\binom{k-1}{\ell-1}\binom{N-k-1}{\ell-1}$, the announced upper bound follows. □

The union bound gives us, using the first upper bound in this last lemma, the following bound on the probability p_k that for some set S of size k, one does not have the desired property $|E^\circ_{f+1}(S, \overline{S})| \geq \min(\gamma k, 2c)$:

$$p_k \leq \sum_{\ell=1}^{k} N^{2\ell} \mathbb{P}\left(|E_{f+1}(S, \overline{S})| \leq \min(\gamma k, 2c) - 2\ell\right).$$

We now distinguish according to whether $\gamma k \leq 2c$ or not.

Case 1: $\gamma k \leq 2c$. We then have, by (6):

$$\begin{aligned} p_k &\leq \sum_{\ell=1}^{\gamma k/2} N^{2\ell} \exp\left(-(2\gamma)^{-1}\gamma k h\left(\frac{\gamma k - 2\ell}{(2\gamma)^{-1}\gamma k}\right)\right) \\ &\leq N \exp\left(\gamma k \ln(N) - (2\gamma)^{-1}\gamma k h(o(1))\right) \\ &= \exp\left(\ln(N)[1 + \gamma k - \gamma k(1/2)\ln(N)^{\beta-1} h(o(1))]\right). \end{aligned}$$

The term in square brackets is asymptotically equivalent to $-\gamma k(1/2)\ln(N)^{\beta-1}$, because $h(o(1))$ tends to 1 and $\beta > 1$. Moreover, since $\gamma k > 2$, the whole exponent is large compared to $\ln(N)$. Thus $p_k = o(N^{-r})$ for any fixed $r > 0$.

Case 2: $\gamma k > 2c$. We then have

$$\begin{aligned} p_k &\leq \sum_{\ell=1}^{c} N^{2\ell} \exp\left(-(2\gamma)^{-1} 2ch\left(\frac{2c - 2\ell}{(2\gamma)^{-1}2c}\right)\right) \\ &\leq N \exp\left(2c\ln(N) - (2\gamma)^{-1}2ch(o(1))\right) \\ &= \exp\left(\ln(N)[1 + 2c - 2c(1/2)\ln(N)^{\beta-1} h(o(1))]\right). \end{aligned}$$

We can then conclude as in the previous case.

4.2 Proof of expansion at equilibrium

To complete the proof of Corollary 1, we now show that the graph G_F obtained after $F = \lceil \log_2(N) \rceil$ phases is an expander, i.e. an ϵ-expander for some fixed $\epsilon \geq \Omega(1)$, thereby strengthening the statement that it is a γ-expander. The total variation distance between the distribution of the graph G_F and that of a cycle plus a uniform random permutation is $o(1)$. To show that G_F is with high probability an expander, it thus suffices to show that a cycle plus a uniform random permutation is an expander. This is shown as follows.

For a set $S \subset [N]$ of size k, the number of edges from the permutation starting at nodes in S and ending at nodes in $\overline{S}$ reads $\sum_{i=1}^k \xi_i$, for Bernoulli random variables ξ_i with mean $1 - k/N$. Moreover, these random variables are negatively associated, as follows from [19]. Chernoff bound on their deviation from the mean is then stronger than the corresponding bound obtained assuming they are independent. This entails that, for $r \leq 1 - k/N$,

$$\mathbb{P}\left(\sum_{i=1}^k \xi_i \leq rk\right) \leq e^{-kD(r||1-k/N)},$$

where $D(r||s) := r\ln(r/s) + (1-r)\ln((1-r)/(1-s))$ is the Kullback-Leibler divergence between Bernoulli distributions with parameters r and s.

Fix $\epsilon > 0$ a small positive constant, and let $k \leq N^{1/3}$. In particular, one has that $k/N \leq 1 - \epsilon$. We thus have, in view of the first bound in Lemma 5, the upper-bound on the probability p_k that there exists some set $S \subset [N]$ of size k such that $|E^\circ_F(S, \overline{S})| < \epsilon k$:

$$\begin{aligned} p_k &\leq \sum_{\ell=1}^{k} N^{2\ell} \mathbb{P}(|E_f(S, \overline{S})| < \epsilon k - 2\ell) \\ &\leq \sum_{\ell=1}^{\lceil \epsilon k/2 \rceil - 1} N^{2\ell} e^{-kD(\epsilon - 2\ell/k||1-k/N)}. \end{aligned}$$

Since $D(\epsilon - 2\ell/k||1 - k/N)$ increases with ℓ, we may upper-bound each term in this last summation by $N^{k\epsilon} e^{-kD(\epsilon||1-k/N)}$. Its logarithm C reads

$$\begin{aligned} C &:= \epsilon k \ln(N) - kD(\epsilon||1 - k/N) \\ &= \epsilon k \ln(N) - k\epsilon \ln\left(\frac{\epsilon}{1 - k/N}\right) \\ &\quad - (k(1-\epsilon))\ln\left(N\frac{1-\epsilon}{k}\right). \end{aligned}$$

The second term is $O(k)$, while the third term is for large enough N no larger than

$$-k(1-\epsilon)\ln(\sqrt{N}) = -(k(1-\epsilon)/2)\ln N.$$

It follows that

$$C \leq k\{O(1) + \ln(N)[\epsilon - (1-\epsilon)/2]\},$$

and, assuming $\epsilon < 1/3$, this is no larger than $-r_\epsilon \ln(N)$, where $r_\epsilon = \frac{1-3\epsilon}{4} > 0$. This yields for all $k \leq N^{1/3}$:

$$p_k \leq ke^{-r_\epsilon k \ln(N)}.$$

As $\sum_{k=1}^{N^{1/3}} ke^{-r_\epsilon k \ln(N)} = o(1)$, with high probability no subset $S \subset [N]$ of size $|S| \leq N^{1/3}$ is such that $|E^\circ_F(S, \overline{S})| < \epsilon|S|$.

For $|S| = k \in [N^{1/3}, N/2]$, we use the second upper bound of Lemma 5 on the number of size k sets made of ℓ segments. Since this bound increases with ℓ for $\ell \leq \epsilon k/2$, we obtain

$$p_k \leq kN\binom{k-1}{k\epsilon/2}\binom{N-k-1}{k\epsilon/2} e^{-kD(\epsilon||1-k/N)}.$$

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TNSE.2019.2926887

Stirling's formula implies that the logarithm of this upper bound is no larger than

$$\ln(kN) + \frac{\epsilon k}{2}\ln\left(\frac{k-1}{k\epsilon/2}\right)$$
$$+ (k-1-k\epsilon/2)\ln\left(\frac{k-1}{k-1-k\epsilon/2}\right)$$
$$+ k\epsilon/2\ln\left(\frac{N-k-1}{k\epsilon/2}\right)$$
$$+ (N-k-1-k\epsilon/2)\ln\left(\frac{N-k-1}{N-k-1-k\epsilon/2}\right)$$
$$- k\epsilon\ln\left(\frac{\epsilon}{1-k/N}\right) - (k(1-\epsilon))\ln\left(N\frac{1-\epsilon}{k}\right).$$

The first term is at most $2\ln(N)$ since $k \le N/2$. Using the inequality $\ln(u) \le u-1$, the fifth is no larger than $k\epsilon/2$. It follows that the sum of the second, third, fifth and sixth terms is upper-bounded by $kf(\epsilon)$ for a function f of ϵ such that $f(\epsilon) = O(\epsilon\ln(1/\epsilon))$. Thus:

$$\begin{aligned}\ln(p_k) &\le 2\ln(N) + kf(\epsilon) + k\epsilon/2\ln\left(\frac{N-k-1}{k\epsilon/2}\right)\\ &\quad - (k(1-\epsilon))\ln\left(N\frac{1-\epsilon}{k}\right)\\ &= 2\ln(N) + kf(\epsilon) + k\epsilon/2\ln\left(\frac{1-(k+1)/N}{\epsilon/2}\right)\\ &\quad + k\epsilon/2\ln(N/k) - k(1-\epsilon)\ln(1-\epsilon)\\ &\quad - k(1-\epsilon)\ln(N/k)\\ &= 2\ln(N) + kg(\epsilon) - k(1-3\epsilon/2)\ln(N/k),\end{aligned}$$

for some function $g(\epsilon)$ such that $g(\epsilon) = O(\epsilon\ln(1/\epsilon))$. This readily implies that for small enough $\epsilon > 0$, there exists a constant $s_\epsilon > 0$ such that, for $k \in [N^{1/3}, N/2]$, $p_k \le e^{-s_\epsilon k}$. The corresponding sum is $o(1)$. Therefore, graph G_F is with high probability an ϵ-expander for some fixed constant $\epsilon > 0$.

5 Controlling the Mean

The goal of this Section is to prove Lemma 2. Let G_f be the static graph during phase f. G_f is a 4-regular undirected graph on $[N]$, and we assume it is a (γ, c)-expander:

$$\forall k \le N/2,\ \phi_k(G_f) \ge \min(\gamma, c/k).$$

Our goal is to prove that for any fixed set S of size $k \le N/2$, by the end of phase f (i.e. after T time steps), the expected number of pointers connecting S to $\overline{S}$ satisfies

$$\mathbb{E}|E_{f+1}(S,\overline{S})| \ge \frac{1}{2\gamma}\min(\gamma k, 2c).$$

The proof is divided into two parts, arguing differently depending on the size k of S. Let $k_c = 4c/\gamma$; sets of size k with $k \le k_c$ (respectively $k > k_c$) will be referred to as small sets (respectively large sets).

5.1 Small sets: from partial expansion to partial spread

Let us now use Theorem 2 to prove the conclusion of Lemma 2 for small values of k.

For a fixed set S of size $k \le k_c$, and a fixed node $i \in S$, let $X_i(t)$ denote the location of the pointer issued from i at time t. Under the dynamics we consider, $X_i(t)$ corresponds to an ordinary random walk on the graph G_f. Moreover, the assumptions of Lemma 2 guarantee that the graph G_f satisfies

$$\phi_{3k}(G_f) \ge \min(\gamma, c/(3k)) \ge \min(\gamma, c/(3k_c)) = \gamma/12.$$

By Theorem 2, one therefore has

$$\mathbb{P}(X_i(T) \in S) \le \frac{|S|}{3k} + \sqrt{3k+1}e^{-\lambda_2^* T},$$

where $\lambda_2^* = \phi_{3k}(G_f)^2/(2\Delta) \ge \gamma^2/1152$.

Recall that $T = \ln(N)^a$ and that $\gamma = \ln(N)^{-\beta}$. Furthermore, $1 < \beta < (a-4)/4$, implying that $a - 2\beta > 1$. We then have

$$\begin{aligned}\mathbb{P}(X_i(T) \in S) &\le \frac{1}{3} + \sqrt{3k+1}\exp\left(-\ln(N)^{a-2\beta}/1152\right)\\ &\le 1/2.\end{aligned}$$

Summing over $i \in S$, we obtain that the expected number of pointers issued from S that point into S at the end of the phase is no larger than $k/2$, and therefore that

$$\mathbb{E}|E_{f+1}(S,\overline{S})| \ge k/2 \ge \frac{1}{2\gamma}\min(\gamma k, 2c).$$

5.2 Large sets

Consider a fixed set S of size k such that $k_c < k \le N/2$, and define $\pi_i(t)$ to be $1/k$ times the probability that a pointer issued from S targets i, conditionally on the initial configuration of these pointers at the beginning of the phase (corresponding to $t = 0$). Let $\pi_{(i)}(t)$ denote the i-th largest value in $(\pi_j(t) : j \in [N])$, and $\pi_{[m]}(t) := \sum_{i\in[m]}\pi_{(i)}(t)$ denote the cumulative mass that the probability distribution $\pi(t)$ puts on the m nodes where its mass is the largest.

One clearly has that

$$\pi_{(i)}(0) = \frac{1}{k}\mathbf{1}_{i\in[k]}.$$

We now establish a property of the time derivative $\frac{\mathrm{d}}{\mathrm{d}t}\pi_{[m]}(t)$:

Lemma 6. *Under the assumptions of Lemma 1 that $\phi_m(G_f) \ge \min(\gamma, c/m)$ for all $m \in [N]$, one has the inequalities*

$$\frac{\mathrm{d}}{\mathrm{d}t}\pi_{[m]}(t) \le -4\sum_{j=1}^{c_m}\left(\pi_{(m-j+1)} - \pi_{(m-j+1+c_m)}\right), \quad (8)$$

where $c_m = \lfloor \min(\gamma m, c)/4 \rfloor$.

Proof. Assume to simplify notation that the permutation which sorts nodes i in $[N]$ in decreasing order of π_i is the identity, so that $\pi_i(t) = \pi_{(i)}(t)$. The time derivative of $\pi_{[m]}$ then reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\pi_{[m]}(t) = \sum_{i\in[m]}\sum_{j>m}\mathbf{1}_{(i,j)\in E_f^\circ}(\pi_j - \pi_i).$$

Indeed, changes in the mass $\pi_{[m]}$ result from interchange of pointer extremities i, j with $i \le m$ and $j > m$, which occur at unit rate for $(i,j) \in E_f^\circ$; when one such interchange occurs, the expected change to $\pi_{[m]}$ is precisely $\pi_j - \pi_i$. Now the number of such edges is by assumption at least $\min(\gamma m, c)$. Moreover, the number of such edges adjacent to any node is at most 4, because the graph has degree bounded by 4.

The value of the right-hand side in the above equation, because the π_i are sorted in decreasing order, is minimized when the edges crossing the cut between $[m]$ are adjacent to nodes with index closest to m. The degree constraint then implies the upper bound (8). □

Let $c' := \lfloor c/4 \rfloor$ and $I := \{i : k-(2/3)k_c \le i \le k+(2/3)k_c\}$. We now introduce an auxiliary process $\{\nu_i(t)\}_{i\in[N],t>0}$ defined via:

$$\nu_i(0) = \frac{1}{k}\mathbf{1}_{i\in[k]}, \qquad i \in [N],$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\nu_i(t) = 4\mathbf{1}_{i\in I}\Big[(\nu_{i-c'}(t)-\nu_i(t))\mathbf{1}_{i-c'\in I} + (\nu_{i+c'}(t)-\nu_i(t))\mathbf{1}_{i+c'\in I}\Big], \quad i\in[N],\ t>0.$$

The probability distribution $\nu(t)$ is readily interpreted as the law at time t of a random walk started with uniform distribution on $[k]$, that jumps from i to $i+c'$ (resp., $i-c'$) at rate 4, provided both i and the destination $i\pm c'$ lie in I.

Denoting $\nu_{[m]}(t) := \sum_{j\in[m]}\nu_j(t)$ for all $m\in[N]$, we then have the following

Lemma 7. *For all $t>0$, $m\in[N]$, it holds that*

$$\pi_{[m]}(t) \le \nu_{[m]}(t).$$

Proof. Introduce the differences $\delta_m(t) := \pi_{[m]}(t) - \nu_{[m]}(t)$. It is readily seen that $\delta_m(0)=0$ for all $m\in[N]$. Inequality (8) of Lemma 6 implies that

$$\frac{\mathrm{d}}{\mathrm{d}t}\pi_{[m]}(t) \le -4\sum_{j=1}^{c'}\Big[\mathbf{1}_{m-j+1\in I}\mathbf{1}_{m-j+1+c'\in I} \times \left(\pi_{(m-j+1)}-\pi_{(m-j+1+c')}\right)\Big]. \quad (9)$$

Indeed, each term in the summation of the right-hand side of (8) is non-negative. The j-th term in the summation in the right-hand side of (9) is included only if $m-j+1\in I$ and $m-j+1+c'\in I$. The first condition implies that

$$m-j+1 \ge k - 2\frac{k_c}{3} \ge \frac{k_c}{3} = \frac{4c}{3\gamma}.$$

In turn this implies that $\gamma m \ge c$, so that $c_m = c'$. Thus the summation in the right-hand side of (9) runs over a subset of indices in the summation in the right-hand side of (8), and (9) follows from (8).

By definition of $\nu_i(t)$, one has

$$\frac{\mathrm{d}}{\mathrm{d}t}\nu_{[m]}(t) = -4\sum_{j=1}^{c'}\Big[\mathbf{1}_{m-j+1\in I}\mathbf{1}_{m-j+1+c'\in I} \times \left(\nu_{(m-j+1)}-\nu_{(m-j+1+c')}\right)\Big]. \quad (10)$$

For all $m\in[N]$, there thus exists an integer $i_m\ge 0$ such that $m-i_m\ge 1$, $m+i_m\le N$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}\pi_{[m]}(t) \le -4\left(2\pi_{[m]}-\pi_{[m-i_m]}-\pi_{[m+i_m]}\right),$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\nu_{[m]}(t) = -4\left(2\nu_{[m]}-\nu_{[m-i_m]}-\nu_{[m+i_m]}\right),$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta_m \le -4\left(2\delta_m-\delta_{m-i_m}-\delta_{m+i_m}\right).$$

In the above, as is easily seen, necessarily $i_1=0$, so that we have the boundary condition $\delta_1\le 0$. Also, since $\pi_{[N]} = \nu_{[N]} = 1$, we have $\delta_N = 0$. The previous equation then implies that necessarily, the supremum over $m\in[N]$ of δ_m cannot become positive, because its derivative is always non-positive. □

By the previous lemma, an upper bound on $\pi_{[k]}(T)$ is provided by $\nu_{[k]}(T)$. However the latter quantity is simpler to analyze. It can be interpreted as $1/k$ times the average number of points of $(2/3)k_c$ random walks initialized at each point in $[k-(2/3)k_c, k]$ which fall within $[k]$ at time T. These walks proceed with jumps of size $\pm c'$ at rate 4, constrained to not leave interval $I = [k-(2/3)k_c, k+(2/3)k_c]$.

For a given initial condition $i\in[k-(2/3)k_c]$, the number of sites it can visit is of the order of $(4/3)k_c/c' = \Theta(\ln(N)^\beta)$. Classical results on the nearest neighbor random walk on an interval $[M]$ state that it mixes in time of the order of M^2 [16]. Thus each of the random walks just introduced mix in time $O(\ln(N)^{2\beta}) = o(T)$, because $2\beta < a$. We therefore have the following evaluation:

$$\pi_{[k]}(T) \le \nu_{[k]}(T) \le 1 - \frac{(2/3)k_c}{k}\left(1/2 - o(1)\right).$$

The expected number $\mathbb{E}|E_{f+1}(S,\overline{S})|$ is then lower-bounded by

$$\begin{aligned}\mathbb{E}|E_{f+1}(S,\overline{S})| &\ge (2/3)k_c(1/2-o(1))\\ &= [1/3-o(1)]4c/\gamma\\ &\ge \frac{1}{2\gamma}(2c).\end{aligned}$$

The announced result follows.

6 Proof of Theorem 2

Proof. In vector form the law $\pi(t)$ of the random walk on G at time t reads $\pi(t) = e^{-tL}\pi(0)$, where L is the graph's Laplacian. Its entries $\pi_i(t)$ are thus linear combinations of n functions of the form $e^{-\lambda_j t}$, where λ_j are the eigenvalues of L, and so is the difference $\pi_i(t)-\pi_j(t)$. It can be shown by induction on N that such linear combinations of N distinct exponential functions are either identically zero in t, or admit at most $N-1$ distinct roots in t. Thus for any $i\ne j$, either $\pi_i(t)\ne\pi_j(t)$ except perhaps for finitely many t, or else $\pi_i(t)=\pi_j(t)$ for all $t\ge 0$.

We can thus split $\mathbb{R}_+$ into finitely many intervals $I^{(1)} = [0,t_1)$, $I^{(2)} = [t_1,t_2), \ldots$, and on each interval $I^{(j)}$ determine a particular permutation $\sigma^{(j)}$ of $[N]$ such that for all j, and all $t\in I_j$, one has

$$\pi_{\sigma^{(j)}(1)}(t) \ge \pi_{\sigma^{(j)}(2)}(t) \ge \cdots \ge \pi_{\sigma^{(j)}(N)}(t).$$

For t in any given interval $I^{(j)}$, we will maintain an auxiliary probability distribution on $[k+1]$, denoted $\{\nu_i(t)\}_{i\in[k+1]}$. This distribution can be interpreted as that of a random walk on a graph $G^{(j)}$ with node set $[k+1]$, obtained from G as follows. We identify node $\sigma^{(j)}(i)$ in G with node i in $G^{(j)}$ for all $i\in[k]$, and collapse all nodes $\sigma^{(j)}(u)$, $u>k$ to form node

$k+1$. All edges are then preserved, so that the adjacency matrix $a^{(j)}$ of $G^{(j)}$ is given by

$$a^{(j)}_{u,v} = a_{\sigma^{(j)}(u),\sigma^{(j)}(v)}, \qquad u,v \in [k],$$
$$a^{(j)}_{u,k+1} = \sum_{v=k+1}^{N} a_{\sigma^{(j)}(u),\sigma^{(j)}(v)}, \qquad u \in [k],$$

where a is the adjacency matrix of G. For convenience, we denote by $\pi_{(i)}(t)$ the i-th largest entry of distribution $\pi(t)$. Thus for $t \in I^{(j)}$, $\pi_{(i)}(t) = \pi_{\sigma^{(j)}(i)}(t)$.

The result of the theorem will then follow from the combination of two ingredients. We first show in Lemma 8 below that, for all t, one has the following bound:

$$\pi_{(i)}(t) \le \nu_{i\wedge(k+1)}(t), \quad i \in [N],\ t \ge 0. \tag{11}$$

We then establish in Lemma 9 below that for all j, the second smallest eigenvalue $\lambda_2^{(j)}$ of the Laplacian of $G^{(j)}$ is lower-bounded by λ_2^* given in (2), where crucially Δ is the largest node degree in G, not in $G^{(j)}$.

This readily implies the L^2 control

$$\sum_{i\in[k+1]} \left|\nu_i(t) - \frac{1}{k+1}\right|^2 \le e^{-2\lambda_2^* t}.$$

Cauchy-Schwarz inequality then gives the following control on variation distance:

$$\sum_{i\in[k+1]} |\nu_i(t) - 1/(k+1)| \le \sqrt{k+1}e^{-\lambda_2^* t}.$$

Together, these two results entail that for all $s \le k$,

$$\sum_{i\in[s]} \pi_{(i)}(t) \le \frac{s}{k+1} + \sqrt{k+1}e^{-\lambda_2^* t}, \tag{12}$$

which is the announced result. □

Lemma 8. *The distributions $\pi(t)$, $\nu(t)$ verify bound* (11).

Proof. The bound trivially holds at $t=0$. We can establish it by induction on each interval $I^{(j)}$. Let us consider one such interval, and assume that the property holds at its left end. For notational simplicity we will assume that $\sigma^{(j)}$ is the identity, so that on this interval $\pi_i(t) = \pi_{(i)}(t)$. Introduce the notation

$$\delta_i(t) := \pi_i(t) - \nu_{i\wedge(k+1)}(t), \quad i \in [N].$$

For any pair of vertices (i,j) in $[N]$, write $i \sim j$ if i and j are neighbors in G. One has the following time derivatives

$$\frac{\mathrm{d}}{\mathrm{d}t}\pi_i = \sum_{\substack{j\in[k]\\ j\sim i}} (\pi_j - \pi_i) + \sum_{\substack{j\notin[k]\\ j\sim i}} (\pi_j - \pi_i), \qquad i \in [N],$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\nu_i = \sum_{\substack{j\in[k]\\ j\sim i}} (\nu_j - \nu_i) + \sum_{\substack{j\notin[k]\\ j\sim i}} (\nu_{k+1} - \nu_i), \qquad i \in [k],$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\nu_{k+1} = \sum_{i\notin[k]} \sum_{\substack{j\in[k]\\ j\sim i}} (\nu_j - \nu_{k+1}).$$

By the previous display one has for $i \in [k]$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta_i = \sum_{j\in[N], j\sim i} (\delta_j - \delta_i). \tag{13}$$

Note that, because the values π_i are sorted, for all $j \notin [k]$, $\pi_j - \pi_{k+1} \le 0$. This together with the expression for the time derivative of π_{k+1} yield

$$\frac{\mathrm{d}}{\mathrm{d}t}\pi_{k+1} \le \sum_{j\in[k], j\sim k+1} (\pi_j - \pi_{k+1}).$$

Thus

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\delta_{k+1} &\le \sum_{\substack{j\in[k]\\ j\sim k+1}} (\pi_j - \pi_{k+1}) - \sum_{i\notin[k]} \sum_{\substack{j\in[k]\\ j\sim i}} (\nu_j - \nu_{k+1}) \\ &= \sum_{\substack{j\in[k]\\ j\sim k+1}} (\delta_j - \delta_{k+1}) - \sum_{i\notin[k+1]} \sum_{\substack{j\in[k]\\ j\sim i}} (\nu_j - \nu_{k+1}).\end{aligned} \tag{14}$$

Let us argue by contradiction, and assume that there exists $t \in \mathbb{R}_+$ and $i \in [N]$ for which $\delta_i(t) > 0$. Let $\delta(t) := \sup_{j\in[N]} \delta_j(t)$. As the π_j are sorted in decreasing order, one also has $\delta(t) = \sup_{j\in[k+1]} \delta_j(t)$.

Since the $\delta_j(t)$ are linear combinations of finitely many exponentials, we can then identify an interval $J = [a,b]$ such that on J, for some $i \in [k+1]$, $\delta(t) = \delta_i(t)$, and moreover $\delta(a) = 0$, $\delta(t) > 0$, $t \in (a,b]$.

Assume that $i \in [k]$. From expression (13), we see that on J, $\frac{\mathrm{d}}{\mathrm{d}t}\delta = \frac{\mathrm{d}}{\mathrm{d}t}\delta_i \le 0$. This contradicts the fact that $\delta > 0$ on $(a,b]$.

Assume then that $i = k+1$. Then on J one has that, for all $j \in [k]$, since the π_j are sorted,

$$\nu_{k+1} = \pi_{k+1} - \delta_{k+1} \le \pi_{k+1} \le \pi_j = \nu_j + \delta_j \le \nu_j + \delta_{k+1}.$$

Thus for all $j \in [k]$, $\nu_{k+1} - \nu_j \le \delta_{k+1}$. It then follows from (14) that

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta_{k+1} \le 0 + \alpha\delta_{k+1},$$

where $\alpha = \sum_{i\notin[k+1]} \sum_{j\in[k], j\sim i}(1)$. Gronwall's lemma (see e.g. [24]) then implies that $\delta_{k+1} \le 0$ on J, a contradiction. □

Remark 2. When we move from interval $I^{(j)}$ to $I^{(j+1)}$ one can check that the meaning of distribution ν is preserved: we may change the permutation sorting the entries π_i, which results in a change in the graph used to define the evolution of ν, but while the vertex to which ν_i refers may change, in that case the corresponding mass does not change.

Lemma 9. *Given a graph G on vertex set $[N]$ with maximal degree Δ and for fixed $k < n$, associated isoperimetric constant $\phi_k(G)$, consider the graph G' obtained by collapsing $N-k$ nodes into a single node as previously described. Then the resulting Laplacian matrix L has spectral gap at least $\lambda_2 \ge \lambda_2^*$, where $\lambda_2^* := \frac{\phi_k(G)^2}{2\Delta}$.*

Proof. Without loss of generality we assume nodes $k+1,\dots,N$ of G have been collapsed into node $k+1$ of G'. Let f be an eigenvector of L associated with its second smallest eigenvalue λ_2. We can always choose f such that $f_{k+1} \le 0$.

Define $g_v = \max(f_v, 0)$, $v \in [k+1]$, and thus $g_{k+1} = 0$. Let $W = \{v \in [k+1] : f_v > 0\}$. Letting $(a_{uv})_{u,v\in[N]}$ denote the adjacency matrix of graph G and d_u the degree of $u \in [N]$, one has

$$\begin{aligned}
\lambda_2 \sum_{u\in W} f_u^2 &= \sum_{u\in W} (Lf)_u f_u \\
&= \sum_{u\in W} \left[d_u f_u - \sum_{v\in[k+1]} a_{uv} f_v \right] f_u \\
&= \sum_{u\in W} \sum_{v\in[k+1]} a_{uv}[f_u - f_v] f_u \\
&= \sum_{u\in W} \sum_{v\in W} a_{uv}(f_u - f_v) f_u \\
&\quad + \sum_{u\in W} \sum_{v\notin W} a_{uv}(f_u - f_v) f_u \\
&\geq \sum_{u\in W} \sum_{v\in W} a_{uv}(f_u - f_v) f_u + \sum_{u\in W} \sum_{v\notin W} a_{uv} f_u^2 \\
&= \langle Lg, g\rangle
\end{aligned}$$

Thus

$$\lambda_2 \geq \frac{\langle Lg, g\rangle}{\langle g, g\rangle} =: K.$$

On the other hand,

$$\begin{aligned}
\sum_{(uv)\in E} a_{uv}(g_u + g_v)^2 &= 2 \sum_{(uv)\in E} a_{uv}(g_u^2 + g_v^2) \\
&\quad - \sum_{(uv)\in E} a_{uv}(g_u - g_v)^2 \\
&\leq 2 \sum_{v\in V} d_v g_v^2 \\
&\leq 2\Delta \langle g, g\rangle,
\end{aligned}$$

where we have used the fact that $g_{k+1} = 0$ to upper bound each product $d_v g_v^2$ by Δg_v^2.

By Cauchy-Schwarz inequality,

$$\begin{aligned}
&\left(\sum_{(uv)\in E} a_{uv} |g_u^2 - g_v^2| \right)^2 \\
&\quad \leq \left(\sum_{(uv)\in E} a_{uv}(g_u - g_v)^2 \right) \left(\sum_{(uv)\in E} a_{uv}(g_u + g_v)^2 \right).
\end{aligned}$$

Combined, these bounds give

$$\begin{aligned}
K &= \frac{\left(\sum_{(uv)\in E} a_{uv}(g_u - g_v)^2\right)\left(\sum_{(uv)\in E} a_{uv}(g_u + g_v)^2\right)}{\langle g, g\rangle \sum_{(uv)\in E} a_{uv}(g_u + g_v)^2} \\
&\geq \frac{\left(\sum_{(uv)\in E} a_{uv}|g_u^2 - g_v^2|\right)^2}{2\Delta \langle g, g\rangle^2}.
\end{aligned}$$

Let $0 = t_0 < t_1 \cdots < t_m$ be the distinct values taken by the g_v. For $i = 0, \ldots, m$, let $V_i := \{v \in V : g_v \geq t_i\}$. Thus for $i > 0$, $(k+1) \notin V_i$. Let

$$\begin{aligned}
M &:= \sum_{(uv)\in E} a_{uv} |g_u^2 - g_v^2| \\
&= \sum_{i=1}^{m} \sum_{(uv)\in E, g_v < g_u = t_i} a_{uv}(g_u^2 - g_v^2) \\
&= \sum_{i=1}^{m} \sum_{u: g_u = t_i} \sum_{\substack{v: g_v = t_j \\ j<i}} a_{uv}(t_i^2 - t_{i-1}^2 + \cdots \\
&\qquad \cdots - t_{j+1}^2 + t_{j+1}^2 - t_j^2) \\
&= \sum_{i=1}^{m} \sum_{u\in V_i} \sum_{v\notin V_i} a_{uv}(t_i^2 - t_{i-1}^2) \\
&= \sum_{i=1}^{m} e(V_i, \overline{V}_i)(t_i^2 - t_{i-1}^2) \\
&\geq \phi_k(G) \sum_{i=1}^{m} |V_i|(t_i^2 - t_{i-1}^2) \\
&= \phi_k(G) \sum_{i=1}^{m} t_i^2(|V_i| - |V_{i+1}|) \\
&= \phi_k(G)\langle g, g\rangle.
\end{aligned}$$

Combined, these results yield

$$\lambda_2 \geq K \geq \frac{\left(\phi_k(G)\langle g, g\rangle\right)^2}{2\Delta \langle g, g\rangle^2} = \lambda_2^*.$$

□

Acknowledgements

The authors would like to thank George Giakkoupis for bringing the problem studied in this paper to their attention, and to acknowledge stimulating discussions on the topic with both George Giakkoupis and Marc Lelarge.

References

[1] S. Hoory, N. Linial, and A. Wigderson, "Expander graphs and their applications," *Bull. Amer. Math. Soc. (N.S.)*, vol. 43, no. 4, pp. 439–561 (electronic), 2006. [Online]. Available: http://dx.doi.org/10.1090/S0273-0979-06-01126-8

[2] L. Massoulié, A.-M. Kermarrec, and A. Ganesh, "Network awareness and failure resilience in self-organizing overlay networks," *22nd International Symposium on Reliable Distributed Systems*, 2003.

[3] C. Cooper, R. Klasing, and T. Radzik, "A randomized algorithm for the joining protocol in dynamic distributed networks," INRIA, Tech. Rep. RR-5376, Nov. 2004. [Online]. Available: https://hal.inria.fr/inria-00070627

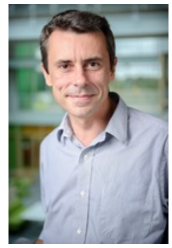
[4] C. Cooper, M. Dyer, and A. J. Handley, "The flip Markov chain and a randomising p2p protocol," in *Proceedings of the 28th ACM Symposium on Principles of Distributed Computing*, ser. PODC '09. New York, NY, USA: ACM, 2009, pp. 141–150. [Online]. Available: http://doi.acm.org/10.1145/1582716.1582742

[5] T. Feder, A. Guetz, M. Mihail, and A. Saberi, "A local switch Markov chain on given degree graphs with application in connectivity of peer-to-peer networks," in *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, ser. FOCS '06. Washington, DC, USA: IEEE Computer Society, 2006, pp. 69–76. [Online]. Available: http://dx.doi.org/10.1109/FOCS.2006.5

[6] R. Jacob, A. Richa, C. Scheideler, S. Schmid, and H. Täubig, "Skip+: A self-stabilizing skip graph," *J. ACM*, vol. 61, no. 6, pp. 36:1–36:26, Dec. 2014. [Online]. Available: http://doi.acm.org/10.1145/2629695

[7] Z. Allen Zhu, A. Bhaskara, S. Lattanzi, V. S. Mirrokni, and L. Orecchia, "Expanders via local edge flips," *CoRR*, vol. abs/1510.07768, 2015.

[8] J. Schweinsberg, "An o(n2) bound for the relaxation time of a Markov chain on cladograms," *Random Struct. Algorithms*, vol. 20, no. 1, pp. 59–70, jan 2002.
[9] M. Jerrum and A. Sinclair, "Approximating the permanent," *SIAM J. Comput.*, vol. 18, no. 6, pp. 1149–1178, dec 1989.
[10] S. Bhamidi, G. Bresler, and A. Sly, "Mixing time of exponential random graphs," *The Annals of Applied Probability*, vol. 21, no. 6, pp. 2146–2170, 12 2011.
[11] J. Jonasson, "Mixing times for the interchange process," *Alea-Latin American Journal of Probability and Mathematical Statistics*, vol. 9, no. 2, pp. 667–683, 2012.
[12] N. Berestycki, "Mixing times of Markov chains: Techniques and examples," *Alea-Latin American Journal of Probability and Mathematical Statistics*, 2016. [Online]. Available: http://www.statslab.cam.ac.uk/~beresty/teach/Mixing/mixing3.pdf
[13] B. Bollobas, *Modern Graph Theory*. Springer, 1998.
[14] J. Friedman, J. Kahn, and E. Szemerédi, "On the second eigenvalue of random regular graphs," in *Proceedings of the Twenty-first Annual ACM Symposium on Theory of Computing*, ser. STOC '89. New York, NY, USA: ACM, 1989, pp. 587–598. [Online]. Available: http://doi.acm.org/10.1145/73007.73063
[15] B. Mohar, "Some applications of laplace eigenvalues of graphs," in *Graph symmetry*. Springer, 1997, pp. 225–275.
[16] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov chains and mixing times*. Providence, R.I. American Mathematical Society, 2009, with a chapter on coupling from the past by James G. Propp and David B. Wilson. [Online]. Available: http://opac.inria.fr/record=b1128575
[17] L. Lovász and R. Kannan, "Faster mixing via average conductance," in *Proceedings of the Thirty-first Annual ACM Symposium on Theory of Computing*, ser. STOC '99. New York, NY, USA: ACM, 1999, pp. 282–287. [Online]. Available: http://doi.acm.org/10.1145/301250.301317
[18] J. R. Lee, S. O. Gharan, and L. Trevisan, "Multiway spectral partitioning and higher-order cheeger inequalities," *J. ACM*, vol. 61, no. 6, pp. 37:1–37:30, Dec. 2014. [Online]. Available: http://doi.acm.org/10.1145/2665063
[19] D. Dubhashi and D. Ranjan, "Balls and bins: a study in negative dependence," *Basic Research in Computer Science*, vol. RS-96-25, 1996.
[20] J. Borcea, P. Brändén, and T. Liggett, "Negative dependence and the geometry of polynomials," *Journal of the American Mathematical Society*, vol. 22, no. 2, pp. 521–567, 2009.
[21] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*. OUP Oxford, 2013. [Online]. Available: https://books.google.fr/books?id=koNqWRluhP0C
[22] N. Alon and J. H. Spencer, *The Probabilistic Method*, 4th ed. Wiley Publishing, 2016.
[23] T. Liggett, "Negative correlations and particle systems," *Markov Processes and Related Fields*, no. 8, pp. 547–564, 2002.
[24] H. K. Khalil, *Nonlinear Systems*. Prentice Hall, 2002.



Laurent Massoulié graduated from the Ecole Polytechnique, Palaiseau, France, in 1991, and received the Ph.D. degree in automatic control from Paris Sud University, Orsay, France, in 1995. He is a researcher at Inria where he leads the Microsoft Research-Inria Joint Centre, and a Professor at the Applied Mathematics Centre of Ecole Polytechnique.

His research focuses on probabilistic modeling and design of algorithms for machine learning as well as "large networks," including P2P and social networks. He has held positions with France Telecom R&D from 1995 to 1999, Microsoft Research, Cambridge, U.K., from 1999 to 2006, and Technicolor, Paris, France, from 2006 to 2012. Dr. Massoulié has served as Associate Editor of Queueing Systems: Theory and Applications from 2000 to 2006, the IEEE/ACM TRANSACTIONS ON NETWORKING in 2008, and the Stochastic Systems Journal from 2011 to the present. He has coauthored the Best Paper Award-winning papers of IEEE INFOCOM 1999, ACM SIGMETRICS 2005, and ACM CoNEXT 2007, been elected a Technicolor Fellow in 2011, and received the "Grand Prix Scientifique" from the Del Duca Foundation in 2017.

Rémi Varloot graduated from the École Normale Supérieure, Paris, France in 2015, where he obtained his B.S., M.S and Ph.D. degrees in 2011, 2013 and 2018, respectively. He is currently a Research Engineer at the IoT-Control department in Nokia Bell Labs France.

His research interests include probabilistic modeling for communication networks, graph theory, sampling algorithms, software defined networking, network automation, wireless networks and distributed systems.