Rapid Mixing of Dynamic Graphs with Local Evolution Rules
Laurent Massoulié, Rémi Varloot

To cite this version:
hal-02427853

HAL Id: hal-02427853
https://hal.science/hal-02427853
Submitted on 19 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
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 $\gamma$-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 $S'$ in $V$. It is an expander if it is a $\gamma$-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 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 $S'$ in $V$. It is an expander if it is a $\gamma$-expander for some $\gamma \geq \Omega(1)$.

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 $[N] = \{1, \ldots, N\}$, 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 setup. 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 $E_0 = \{(n, n + 1) : i \in [N]\}$, with $N + 1 \equiv 1$. Each node $n \in [N]$ furthermore maintains two pointers, one blue and one red, with respective destinations $b_n$ and $r_n$ 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 $E_b = \{(n, b_n) : n \in [N]\}$ and $E_r = \{(n, r_n) : n \in [N]\}$. 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 $G_b$ constituted of the edges in both $E_0$ and $E_b$. Note that $G_b$ is a 4-regular multigraph.

The dynamics for a blue phase are defined as follows. Each edge $e = (i, j)$ of $G_b$ 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 $E_0$ and $E_b$ is denoted $G_r$. To clarify notations, we denote $E_f$ the fixed edge set during phase $f$, i.e. $E_b$ if $f$ is a red phase, $E_r$ otherwise, and $G_f$ the fixed graph, i.e. containing the edges in $E_0$ and $E_f$. We finally write $E^0 = E_f \cup E_0$.

Our main result is then as follows.

**Theorem 1.** Let $T = \ln^8 N$ where $a > 8$ is a constant. After $F = \lfloor \log_2 N \rfloor$ 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 $8\tau$. 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 $16\tau$ is then, by Chernoff’s bound for deviations of Poisson random variables from their mean, bounded by

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

where $h(x) := x \ln(x) - x + 1$ is the Cramér transform of a unit mean Poisson random variable. Since $\tau$ 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 $\tau$ 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 a directed multi-graph with maximum degree $\Delta$, 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 $\{\pi_t(t)\}_{t \in [N]}$ 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:

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

where

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

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

**Remark 1.** The quantity $\lambda_2^*$ is of the same form as the lower bound on the spectral gap $\lambda_2$ of the Laplacian that the
where $\gamma \geq \delta$. 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 $d_{\text{var}}(\pi(t), \mathcal{U}([N])) \leq \sqrt{Ne^{-\lambda^2 t}}$, 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

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

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 $(\gamma, c)$-expander if, for all $k \leq N/2$, $\phi_k(G) \geq \min(\gamma, c/k)$.

Note that a graph is a $\gamma$-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 $\beta$ be a constant such that $1 < \beta < (a - 4)/4$. Such $\beta$ exists by our assumption that $a > 8$. Let $\gamma = \ln^{-2} N$. We show the following

**Lemma 1.** If, for a given phase $f$, the fixed graph $G_f$ is a $(\gamma, c)$-expander for some integer $c$, then with probability at least $1 - o(1/N)$, $G_{f+1}$ (the 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 $(\gamma, c)$-expander, then for all $S \subset [N]$ with $|S| \leq N/2$,

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

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 $F = \lfloor \log_2 N \rfloor$ phases, with high probability both $G_b$ and $G_r$ 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

$$|E^i(S, \overline{S})| \geq 2 \geq \min(\gamma|S|, 2).$$
Denote by $E_f$ the event that $G_f$ is a $(\gamma, 2^f)$-expander. Thus we have just established that event $E_1$ holds with certainty, and Lemma 1 entails that, for all $f \geq 1$,
$$\Pr[E_{f+1} | E_f] \leq o(1/N).$$

Thus
$$\Pr[E_f] = \Pr[E_{f+1} | E_f] \Pr[E_f] = \Pr[E_{f+1}] \Pr[E_f] \leq o(1/N) + \Pr[E_f].$$

By induction on $f$, this yields
$$\Pr[E_f] \leq o(f/N).$$

For $F = \lfloor \log_2 N \rfloor$, 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 as follows:

**Proof.** By Corollary 2, after $F = \lfloor \log_2 N \rfloor$ phases, the Cheeger constants $\phi_{N/2}(G_b)$ and $\phi_{N/2}(G_r)$ are at least $\gamma$. 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 space $S$, denoting $\pi_s(t)$ its distribution at time $t$ conditional on the initial state being $s \in S$, and $\pi(\infty)$ its stationary distribution, its mixing time $T_{\text{mix}}$ is defined as
$$T_{\text{mix}} = \inf\{t > 0 : \sup_{s \in S} ||\pi_s(t) - \pi(\infty)||_{TV} \leq 1/4\},$$
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 $\gamma_{ij}$ on $G$ connecting $i$ and $j$, and let $\Gamma = \{\gamma_{ij} : i, j \in [N]\}$. Denote $\Upsilon$ the length of the longest path in $\Gamma$, and $K$ the supremum over the edges $e$ of $G$ of the number of paths in $\Gamma$ passing through $e$.

The continuous time interchange process on $G$ will have mixed in time $T$ provided
$$T \geq 8 \ln(N) \Upsilon K / N. \quad (5)$$

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 $\ln(N)^4/\gamma^2$, i.e provided $\ln(N)^4 = \omega((\ln(N)^4+4\beta))$. This condition holds since $\beta < (a - 4)/4$.

**Lemma 3.** Let $G$ be a $d$-regular 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 $\gamma_{ij}$ each have length at most $\Upsilon = 2 \ln(N)^2 / \gamma^2$, 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 $\Upsilon := 2 \ln(N)^2 / \gamma^2$ 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 $\Upsilon$ is at most $1/(2N)$. Consider then the following randomized construction. For each $i$, create $5N \ln(N)$ independent walks of length $\Upsilon$ 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,
$$\mathbb{E} \sum_{i \in [N]} \sum_{t \leq 5N \ln(N)} X_i(t) = 5N \ln(N) \sum_{i \in [N]} \sum_{\ell = 0}^{\Upsilon} p_i^{(\ell)}.$$
where \( P_{iu}^{(t)} \) denotes the transition probability from \( i \) to \( u \) in \( \ell \) steps of the walk. However the walk is symmetric, so that \( P_{iu}^{(t)} = P_{ui}^{(t)} \). The above expection thus reads
\[
E \sum_{i \in [N]} \sum_{t \leq 5N \ln(N)} X_i(t) = 5N \ln(N)(\Upsilon + 1).
\]
Let \( Z = \sum_{i \in [N]} \sum_{t \leq 5N \ln(N)} X_i(t) \) denote the total number of visits to \( u \) by all paths. For any \( C > 0 \), Hoeffding's inequality then gives
\[
\Pr(Z \geq E(Z) + CN \Upsilon \ln(N)) \leq \exp\left(-\frac{C^2N^2\Upsilon^2 \ln(N)^2}{\Upsilon^2N^2 \ln(N)}\right) = e^{-C^2 \ln(N)/5}.
\]
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.

4 From bounds in expectation to bounds with high probability

4.1 Proof of Lemma 1
Assume that \( G_f \) is a \((\gamma,c)\)-expander. By Lemma 2, for each fixed set \( S \subset [N] \) with \(|S| \leq N/2\), we have that
\[
E|E_{f+1}(S, \mathcal{S})| \geq \frac{1}{2\gamma} \min(\gamma|S|, 2c).
\]
Fix \( k \leq N/2 \). We further restrict ourselves to \( k > 2/\gamma \), since \(|E_0(S, \mathcal{S})| \geq 2\) and therefore one always has that \( \phi_k \geq \gamma \) for \( k \leq 2/\gamma \).

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 \((S, E_0 \cap S^2)\). Clearly \(|E_0(S, \mathcal{S})| = 2\ell\), and therefore
\[
|E_{f+1}^g(S, \mathcal{S})| = |E_{f+1}(S, \mathcal{S})| + 2\ell.
\]
Recall (see e.g. Dubashi and Ranjan [19], and Borcea et al. [20]) that a set of random variables \((X_i)_{i \in \mathbb{I}}\) is said to be negatively associated if for any two functions \(f, g : \mathbb{R}^\mathbb{I} \to \mathbb{R}\) that are non-decreasing in each of their coordinates, and depend on disjoint sets of variables \(X_i\), the two random variables \(f((X_i)_{i \in \mathbb{I}})\) and \(g((X_i)_{i \in \mathbb{I}})\) are negatively correlated, i.e.
\[
E[f((X_i)_{i \in \mathbb{I}})]g((X_i)_{i \in \mathbb{I}})] \leq E[f((X_i)_{i \in \mathbb{I}})]E[g((X_i)_{i \in \mathbb{I}})].
\]
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 \(|E_{f+1}(S, \mathcal{S})|\) consists of the sum of negatively associated Bernoulli random variables. Consequently, for any \( r \in (0,1) \), it holds that
\[
P \left(|E_{f+1}(S, \mathcal{S})| \leq r(2\gamma)^{-1} \min(\gamma|S|, 2c)\right) \leq e^{-(2\gamma)^{-1} \min(\gamma|S|, 2c) m h(r)},
\]
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
\[
|E_{f+1}(S, \mathcal{S})| = \sum_{i \in S} \xi_i(T).
\]
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 \(|E_{f+1}(S, \mathcal{S})|\) 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 \( m \leq M := E(|E_{f+1}(S, \mathcal{S})|) \leq \xi(0) \), that
\[
P(|E_{f+1}(S, \mathcal{S})| \leq m \mid \xi(0)) \leq e^{-M h(m/M)}{,}
\]
where \( h(x) := x \ln(x) - x + 1 \). Indeed, the right-hand side in the above inequality is the Chernoff bound of \( \Pr(\text{Poisson}(M) \leq m) \), 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 \( t \geq 0 \) 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.

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

It is also upper-bounded by \( N^{k-1}(N-k-1)_{\ell-1} \).

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 \( 2\ell \) numbers in \([N]\), which characterize \( S \), hence the upper bound \( N^{2k/\ell} \).

To obtain the other upper bound, note that the number of strictly positive sequences of \( \ell \) integers \( x_1, \ldots, x_\ell \) such
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 \( y_1, \ldots, y_\ell \) 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 \( \ell \) distinct intervals, let \( z \in \{0, \ldots, N - 1\} \) 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, \( x_1, \ldots, x_\ell \), in clockwise order, together with the lengths of the segments separating its own segments, \( y_1, \ldots, y_\ell \). Since this construction forms an injection from the collection of considered sets \( S \) to a set of size \( N^{(k-1)/\ell} (N - k - 1) \), 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 \( |E_{\ell+1}(S, \overline{S})| \geq \min(\gamma k, 2c) \):
\[
p_k \leq \sum_{\ell=1}^{k} N^{2\ell} \mathbb{P}(|E_{\ell+1}(S, \overline{S})| \leq \min(\gamma k, 2c) - 2\ell).
\]

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

**Case 1:** \( \gamma k \leq 2c \). We then have, by (6):
\[
p_k \leq \sum_{\ell=1}^{\gamma k/2} N^{2\ell} \exp\left(-2(\gamma)\gamma k h\left(\frac{\gamma k - 2\ell}{(2\gamma)^{\ell}-2\gamma k}\right)\right)
\]
\[
\leq N \exp\left(\gamma k \ln(N) - 2(\gamma)\gamma k h(o(1))\right)
\]
\[
= \exp\left(\ln(N)[1 + \gamma k - \gamma k (1/2) \ln(N)^{\beta-1} h(o(1))\right)].
\]

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

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 \( F = [\log_2(N)] \) phases is an expander, i.e., an \( \epsilon \)-expander for some fixed \( \epsilon \geq \Omega(1) \), thereby strengthening the statement that it is a \( \gamma \)-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
\[
\sum_{\ell=1}^{k} \xi_\ell,
\]
for Bernoulli random variables \( \xi_\ell \) 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 \( r \leq 1 - k/N \),
\[
\mathbb{P}\left(\sum_{\ell=1}^{k} \xi_\ell \leq r k\right) \leq e^{-kD(r||1-k/N)}
\]
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 \( |E_{\ell+1}(S, \overline{S})| < \epsilon k \):
\[
p_k \leq \sum_{\ell=1}^{k} N^{2\ell} \mathbb{P}(|E_{\ell+1}(S, \overline{S})| \leq \epsilon k - 2\ell)
\]
\[
\leq \sum_{\ell=1}^{[\epsilon k/2]-1} N^{2\ell} e^{-kD(\epsilon - 2\ell/k||1-k/N)}.
\]

Since \( D(\epsilon - 2\ell/k||1 - k/N) \) increases with \( \ell \), we may upper-bound each term in this last summation by \( N^{k\epsilon/2} e^{-kD(\epsilon - 2\ell/k||1-k/N)} \). Its logarithm \( C \) reads
\[
C := \epsilon k \ln(N) - kD(\epsilon||1-k/N)
\]
\[
= \epsilon k \ln(N) - k \epsilon \ln\left(\frac{1 - \epsilon}{1 - k/N}\right)
\]
\[
- (k(1 - \epsilon)) \ln\left(\frac{N^{1 - \epsilon/k}}{k}\right).
\]

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 \(-r_k \ln(N)\), where \( r_k = \frac{1 - 3\epsilon}{4\epsilon} > 0 \). This yields for all \( k \leq N^{1/3} \),
\[
p_k \leq ke^{-r_k \ln(N)}.
\]

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

For \( |S| = k \in [N^{1/3}, N/2] \), we use the second upper bound of Lemma 5 on the probability of size \( k \) sets made of \( \ell \) segments. Since this bound increases with \( \ell \) for \( \ell \leq \epsilon k/2 \), we obtain
\[
p_k \leq kN \left(\frac{k - 1}{ke/2}\right) \left|N - k - 1 - ke/2\right| e^{-kD(\epsilon||1-k/N)}.
\]
Stirling’s formula implies that the logarithm of this upper bound is no larger than
\[
\ln(kN) + \frac{ek}{2} \ln \left( \frac{k - 1}{ke/2} \right) + (k - 1 - ke/2) \ln \left( \frac{k - 1}{k - 1 - ke/2} \right) + ke/2 \ln \left( \frac{N - k - 1}{k - 1 - ke/2} \right) + (N - k - 1 - ke/2) \ln \left( \frac{N - k - 1}{N - k - 1 - ke/2} \right) - ke \ln \left( 1 - \frac{\epsilon}{k - 1} \right) - (k(1 - \epsilon)) \ln \left( \frac{N - k - 1}{k - 1} \right).
\]

The first term is at most \(2 \ln(N)\) since \(k \leq N/2\). Using the inequality \(\ln(u) \leq u - 1\), the fifth is no larger than \(ke/2\). It follows that the sum of the second, third, fifth and sixth terms is upper-bounded by \(kf(\epsilon)\) for a function \(f(\epsilon)\) such that \(f(\epsilon) = O(\epsilon \ln(1/\epsilon))\). Thus:
\[
\ln(pk) \leq 2 \ln(N) + kf(\epsilon) + ke/2 \ln \left( \frac{N - k - 1}{k - 1 - ke/2} \right)
- (k(1 - \epsilon)) \ln \left( \frac{N - k - 1}{k - 1} \right)
= 2 \ln(N) + kf(\epsilon) + ke/2 \ln \left( \frac{1 - (k + 1)/N}{\epsilon/2} \right)
+ ke/2 \ln(N/k) - k(1 - \epsilon) \ln(1 - \epsilon)
- k(1 - \epsilon) \ln(N/k)
= 2 \ln(N) + kg(\epsilon) - k(1 - 3/2e) \ln(N/k),
\]
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 \(k \in [N^{1/3}, N/2]\), \(pk \leq e^{-s_\epsilon k}\). The corresponding sum is \(o(1)\). Therefore, graph \(G_f\) is with high probability an \(\epsilon\)-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 \((\gamma, c)\)-expander:
\[
\forall k \leq N/2, \quad \phi_k(G_f) \geq \min(\gamma, c/k).
\]

Our goal is to prove that for any fixed set \(S\) of size \(k \leq N/2\), by the end of phase \(f\) (i.e. after \(T\) time steps), the expected number of pointers connecting \(S\) to \(S\) satisfies
\[
\mathbb{E}[E_{f+1}(S, S)] \geq \frac{1}{2\gamma} \min(\gamma k, 2c).
\]

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 \(k \leq k_c\) (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 \(k \leq k_c\), 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
\[
\phi_{3k}(G_f) \geq \min(\gamma, c/(3k)) \geq \min(\gamma, c/(3k_c)) = \gamma/12.
\]

By Theorem 2, one therefore has
\[
\mathbb{P}(X_i(T) \in S) \leq \frac{|S|}{3k} + \frac{1}{e^{\lambda_i^2 T}},
\]
where \(\lambda_i^2 = \phi_{3k}(G_f)^2/(2\Delta) \geq \gamma^2/1152\).

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
\[
\mathbb{P}(X_i(T) \in S) \leq \frac{1}{3} + \frac{1}{e^{\gamma^2/1152}} \leq 1/2.
\]

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
\[
\mathbb{E}[E_{f+1}(S, S)] \geq k/2 \geq \frac{1}{2\gamma} \min(\gamma k, 2c).
\]

5.2 Large sets
Consider a fixed set \(S\) of size \(k\) such that \(k_c < k \leq N/2\), and define \(\pi_t(i)\) 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 \(\pi_t(i)\) denote the \(i\)-th largest value in \(\pi_t(j) : j \in [N]\), and \(\pi_{[m]}(t) := \sum_{i \in [m]} \pi_t(i)\) 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
\[
\pi_t(i) = 1/k \quad \text{for } i \in [k].
\]

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

Lemma 6. Under the assumptions of Lemma 1 that \(\phi_{m}(G_f) \geq \min(\gamma, c/m)\) for all \(m \in [N]\), one has the inequalities
\[
\frac{d}{dt} \pi_{[m]}(t) \leq -4 \sum_{j=1}^{c_m} \left( \pi_{[m-j+1]} - \pi_{[m-j+1+c_m]} \right),
\]
where \(c_m = \lfloor \min(\gamma m, c) \rfloor/4\).

Proof. Assume to simplify notation that the permutation which sorts nodes \(i\) in \([N]\) in decreasing order of \(\pi_t(i)\) is the identity, so that \(\pi_t(i) = \pi_t(i)\). The time derivative of \(\pi_{[m]}\) then reads
\[
\frac{d}{dt} \pi_{[m]}(t) = \sum_{i \in [m]} \sum_{j > m} \mathbb{1}_{(i, j) \in E_j^t} (\pi_j - \pi_i).
\]

Indeed, changes in the mass \(\pi_{[m]}\) result from interchange of pointer extremities \(i, j\) with \(i < m\) and \(j > m\), which occur at unit rate for \((i, j) \in E_j^t\). When 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 $\pi_i$ 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 $I := \{i : k - (2/3)k_c \leq i \leq k + (2/3)k_c \}$. We now introduce an auxiliary process $\nu_i(t) \forall i \in [N], t > 0$ defined via:

$$
\nu_i(t) = \frac{1}{k} \mathbb{1}_{i \in [k]}, \quad i \in [N],
$$

$$
\frac{d}{dt} \nu_i(t) = 4\mathbb{1}_{i \in I} \left( (\nu_{i-c'}(t) - \nu_i(t)) \mathbb{1}_{i-c' \in I} + (\nu_{i+c'}(t) - \nu_i(t)) \mathbb{1}_{i+c' \in I} \right), \quad i \in [N], \; t > 0.
$$

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

$$
\pi[m](t) \leq \nu[m](t).
$$

**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

$$
\frac{d}{dt} \pi[m](t) \leq -4 \sum_{j=1}^{c'} \mathbb{1}_{m-j+1 \in I} \mathbb{1}_{m-j+1+c' \in I} \times \left( \pi[m-j+1] - \pi[m-j+1+c'] \right).
$$

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

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

In turn this implies that $\gamma m \geq c$, 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

$$
\frac{d}{dt} \nu[m](t) = -4 \sum_{j=1}^{c'} \mathbb{1}_{m-j+1 \in I} \mathbb{1}_{m-j+1+c' \in I} \times \left( \nu[m-j+1] - \nu[m-j+1+c'] \right).
$$

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

$$
\frac{d}{dt} \pi[m](t) \leq -4 \left( 2\pi[m] - \pi[m-i_m] - \pi[m+i_m] \right),
$$

$$
\frac{d}{dt} \nu[m](t) = -4 \left( 2\nu[m] - \nu[m-i_m] - \nu[m+i_m] \right),
$$

so that

$$
\frac{d}{dt} \delta_m \leq -4 \left( 2\delta_m - \delta_{m-i_m} - \delta_{m+i_m} \right).
$$

In the above, as is easily seen, necessarily $i_1 = 0$, so that we have the boundary condition $\delta_1 \leq 0$. 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 $\delta_m$ 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 $i \in [k - (2/3)k_c, k]$, the number of sites it can visit is of the order of $(4/3)k_c c'/\Theta(\ln(N)^3)$. 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 $O(\ln(N)^3)$, because $2\beta < a$. We therefore have the following evaluation:

$$
\pi[k](T) \leq \nu[k](T) \leq 1 - \frac{(2/3)k_c}{k} \frac{(1/2 - o(1))}{k_c}.
$$

The expected number $\mathbb{E}[E_{|S\in X}(S, \mathcal{S})]$ is then lower-bounded by

$$
\mathbb{E}[E_{|S\in X}(S, \mathcal{S})] \geq \frac{(2/3)k_c}{k_c} \frac{(1/2 - o(1))}{k_c} = \frac{1}{3} - \frac{o(1)}{4c/\gamma}
$$

$$
\geq \frac{1}{2\gamma}(2\epsilon).
$$

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 $\lambda_j$ 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 $i \neq j$, either $\pi_i(t) \neq \pi_j(t)$ except perhaps for finitely many $t$, or else $\pi_i(0) = \pi_j(0)$ for all $t \geq 0$.

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

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

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...
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 \( t \in I^j(\lambda) \), \( \pi(i)(t) = \pi(\sigma(i)(t)) \).

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:

\[
\pi(i)(t) \leq \nu_i \wedge (k+1)(t), \quad i \in [N], \ t \geq 0. \tag{11}
\]

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 \( \lambda_2^j \) given in (2), where crucially \( \Delta \) is the largest node degree in \( G \), not in \( G^j \).

This readily implies the \( L^2 \) control

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

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

\[
\sum_{i \in [k+1]} \left| \nu_i(t) - 1/(k+1) \right| \leq \sqrt{k+1} e^{-2\lambda_2^j t}.
\]

Together, these two results entail that for all \( s \leq k \),

\[
\sum_{i \in [s]} \pi(i)(t) \leq \frac{s}{k+1} + \sqrt{k+1} e^{-2\lambda_2^j t}, \tag{12}
\]

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 \( I^j(\lambda) \). Let us consider one such interval, and assume that the property holds at its left end. For notational simplicity we will assume that \( \sigma(i) \) 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

\[
\frac{d}{dt} \pi_i = \sum_{j \in [k]} (\pi_j - \pi_i) + \sum_{j \notin [k]} (\pi_j - \pi_i), \quad i \in [N],
\]

\[
\frac{d}{dt} \nu_i = \sum_{j \in [k]} (\nu_j - \nu_i) + \sum_{j \notin [k]} (\nu_j - \nu_i), \quad i \in [k],
\]

\[
\frac{d}{dt} \nu_k = \sum_{i \notin [k]} \sum_{j \in [k]} (\nu_i - \nu_{k+1}).
\]

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

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

Note that, because the values \( \pi_i \) are sorted, for all \( j \notin [k] \), \( \pi_j - \pi_{k+1} \leq 0 \). This together with the expression for the time derivative of \( \pi_k \) yield

\[
\frac{d}{dt} \pi_k \leq \sum_{j \in [k], j \sim k+1} (\pi_j - \pi_{k+1}).
\]

Thus

\[
\frac{d}{dt} \delta_k \leq \sum_{j \in [k]} (\pi_j - \pi_{k+1}) - \sum_{i \notin [k]} (\nu_j - \nu_{k+1}) = \sum_{j \in [k]} (\delta_j - \delta_{k+1}) - \sum_{i \notin [k]} (\nu_j - \nu_{k+1}). \tag{14}
\]

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 \( \pi_j \) are sorted in decreasing order, one also has \( \delta(t) \leq \sup_{j \in [k]} \delta_j(t) \).

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 \( i \in [k+1] \), \( \delta_i(t) = \delta_{i+1}(t) \), and moreover \( \delta(a) = 0 \), \( \delta(t) = 0 \), \( t \in (a, b) \).

Assume that \( i \in [k] \). From expression (13), we see that on \( J \),

\[
\frac{d}{dt} \delta_i \leq \sum_{j \in [N], j \sim i} (\delta_j - \delta_i) \leq 0.
\]

Assume then that \( i = k+1 \). Then on \( J \) one has that, for all \( j \in [k] \), since the \( \pi_j \) are sorted, \( \nu_{k+1} = \pi_{k+1} - \pi_{k+1} \leq \pi_j \leq \nu_j + \delta_j \leq \nu_j + \delta_{k+1} \).

Thus for all \( j \in [k] \), \( \nu_{k+1} + \delta_{k+1} \leq \delta_{k+1} \leq 0 \). It then follows from (14) that

\[
\frac{d}{dt} \delta_{k+1} \leq 0 + \alpha \delta_{k+1},
\]

where \( \alpha = \sum_{i \notin [k]} \sum_{j \in [k]} \delta_j - \delta_{k+1} \). Gronwall’s lemma (see e.g. [24]) then implies that \( \delta_{k+1} \leq 0 \) on \( J \), a contradiction. \( \square \)

**Remark 2.** When we move from interval \( I^j(\lambda) \) to \( I^{j+1}(\lambda) \) one can check that the meaning of distribution \( \nu \) is preserved: we may change the permutation sorting the entries \( \pi_i \), which results in a change in the graph used to define the evolution of \( \nu \), but while the vertex to which \( \nu_i \) 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 \( \Delta \) 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 \( \lambda_2^j \geq \lambda_2^j \), where \( \lambda_2^j := \phi_k(G)^2/2\Delta \).

**Proof.** Without loss of generality we assume nodes \( k+1, \ldots, N \) of \( G \) have been collapsed into node \( k+1 \) of \( G' \). Let \( f \) be an eigenvector of \( L \) associated with its second smallest eigenvalue \( \lambda_2 \). We can always choose \( f \) such that \( f_{k+1} \leq 0 \).

Define \( g_v := \max(f_v, 0), v \in [k+1] \), and thus \( g_{k+1} = 0 \).

Let \( W = \{v \in [k+1] : g_v > 0\} \). 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

$$
\lambda_2 \geq \langle \lambda_2, g \rangle =: K.
$$

On the other hand,

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

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

By Cauchy-Schwarz inequality,

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

Combined, these bounds give

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

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

$$
M := \sum_{(uv) \in E} a_{uv}|g_u^2 - g_v^2| \\
= \sum_{i=1}^m (uv) \in E, g_v < g_u = t_i \\
= \sum_{i=1}^m \sum_{v:g_v = t_j} a_{uv}(t_j^2 - t_i^2 - 1 + \cdots - t_{j+1}^2) \\
= \sum_{i=1}^m \sum_{v \in V_i} a_{uv}(t_i^2 - t_{i-1}^2) \\
= \phi_k(G) \sum_{i=1}^m |V_i|(t_i^2 - t_{i-1}^2) \\
= \phi_k(G) \sum_{i=1}^m t_i^2 (|V_i| - |V_{i+1}|) \\
= \phi_k(G) \langle g, g \rangle.
$$

Combined, these results yield

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

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


